Regulation by non-coding RNAs

Volume 1

Edited by

Nicholas Delihas

Printed Edition of the Special Issue Published in *IJMS*

www.mdpi.com/journal/ijms

This book is a reprint of the special issue that appeared in the online open access journal *International Journal of Molecular Sciences (ISSN 1422-0067)* in 2013 (available at: http://www.mdpi.com/journal/ijms/special_issues/regulation-by-non-coding-rnas).

Guest Editor
Nicholas Delihas
Department of Molecular Genetics and Microbiology, School of Medicine, 158 Life Sciences Building, Stony Brook University, Stony Brook, NY 11794, USA

Editorial Office
MDPI AG
Klybeckstrasse 64
Basel, Switzerland

Publisher
Shu-Kun Lin

Managing Editor
Rui Lu

1. Edition 2014

MDPI • Basel, Switzerland

Volume 1 (ISBN 978-3-03842-010-1)
Volume 2 (ISBN 978-3-03842-011-8)

Nicholas Delihas (Ed.)

Regulation by non-coding RNAs

Volume 1

MDPI

Table of Contents

Volume 1

Nicholas Delihas
Preface *Guest Editor* IX

Chapter 1. Methodologies and mechanisms

Michelle L. Stoller, Henry C. Chang and Donna M. Fekete
Bicistronic Gene Transfer Tools for Delivery of miRNAs and Protein Coding Sequences
Reprinted from Int. J. Mol. Sci. 2013, *14(9)*, 18239–18255 1
http://www.mdpi.com/1422-0067/14/9/18239

Florian Mayr and Udo Heinemann
Mechanisms of Lin28-Mediated miRNA and mRNA Regulation—A Structural and Functional Perspective
Reprinted from Int. J. Mol. Sci. 2013, *14(8)*, 16532–16553 19
http://www.mdpi.com/1422-0067/14/8/16532

Epaminondas Doxakis
Principles of miRNA-Target Regulation in Metazoan Models
Reprinted from Int. J. Mol. Sci. 2013, *14(8)*, 16280–16302 41
http://www.mdpi.com/1422-0067/14/8/16280

Hao Zheng, Rongguo Fu, Jin-Tao Wang, Qinyou Liu, Haibin Chen and Shi-Wen Jiang
Advances in the Techniques for the Prediction of microRNA Targets
Reprinted from Int. J. Mol. Sci. 2013, *14(4)*, 8179–8187 65
http://www.mdpi.com/1422-0067/14/4/8179

Anita Quintal Gomes, Sofia Nolasco and Helena Soares
Non-Coding RNAs: Multi-Tasking Molecules in the Cell
Reprinted from Int. J. Mol. Sci. 2013, *14(8)*, 16010–16039 75
http://www.mdpi.com/1422-0067/14/8/16010

Giovanni Bussotti, Cedric Notredame and Anton J. Enright
Detecting and Comparing Non-Coding RNAs in the High-Throughput Era
Reprinted from Int. J. Mol. Sci. 2013, *14(8)*, 15423–15458 105
http://www.mdpi.com/1422-0067/14/8/15423

Changqing Zhang, Guangping Li, Jin Wang, Shinong Zhu and Hailing Li
Cascading *cis*-Cleavage on Transcript from *trans*-Acting siRNA-Producing Locus 3
Reprinted from Int. J. Mol. Sci. 2013, *14(7)*, 14689–14699 141
http://www.mdpi.com/1422-0067/14/7/14689

Sung-Min Kang, Ji-Woong Choi, Su-Hyung Hong and Heon-Jin Lee
Up-Regulation of microRNA* Strands by Their Target Transcripts
Reprinted from Int. J. Mol. Sci. 2013, *14(7)*, 13231–13240 153
http://www.mdpi.com/1422-0067/14/7/13231

Maren Thomas, Kerstin Lange-Grünweller, Dorothee Hartmann, Lara Golde, Julia Schlereth, Dennis Streng, Achim Aigner, Arnold Grünweller and Roland K. Hartmann
Analysis of Transcriptional Regulation of the Human miR-17-92 Cluster; Evidence for Involvement of Pim-1
Reprinted from Int. J. Mol. Sci. 2013, *14(6)*, 12273–12296 163
http://www.mdpi.com/1422-0067/14/6/12273

Yoshiro Nagata, Eigo Shimizu, Naoki Hibio and Kumiko Ui-Tei
Fluctuation of Global Gene Expression by Endogenous miRNA Response to the Introduction of an Exogenous miRNA
Reprinted from Int. J. Mol. Sci. 2013, *14(6)*, 11171–11189 189
http://www.mdpi.com/1422-0067/14/6/11171

Sara Tomaselli, Barbara Bonamassa, Anna Alisi, Valerio Nobili, Franco Locatelli and Angela Gallo
ADAR Enzyme and miRNA Story: A Nucleotide that Can Make the Difference
Reprinted from Int. J. Mol. Sci. 2013, *14*(11), 22796-22816 209
http://www.mdpi.com/1422-0067/14/11/22796

Chapter 2. Role of ncRNAs in disease

Jing Li, Zhenyu Xuan and Changning Liu
Long Non-Coding RNAs and Complex Human Diseases
Reprinted from Int. J. Mol. Sci. 2013, *14(9)*, 18790–18808 229
http://www.mdpi.com/1422-0067/14/9/18790

2.1. Transposable Elements, Non-Coding RNAs and disease formation

Michael Hadjiargyrou and Nicholas Delihas
The Intertwining of Transposable Elements and Non-Coding RNAs
Reprinted from Int. J. Mol. Sci. 2013, *14(7)*, 13307–13328 249
http://www.mdpi.com/1422-0067/14/7/13307

2.2. miRNAs and renal pathophysiology

Jianghui Hou and Dan Zhao
MicroRNA Regulation in Renal Pathophysiology
Reprinted from Int. J. Mol. Sci. 2013, *14(7)*, 13078–13092 271
http://www.mdpi.com/1422-0067/14/7/13078

2.3. Cardiovascular system

Paolo Martini, Gabriele Sales, Enrica Calura, Mattia Brugiolo, Gerolamo Lanfranchi, Chiara Romualdi and Stefano Cagnin
Systems Biology Approach to the Dissection of the Complexity of Regulatory Networks in the *S. scrofa* Cardiocirculatory System
Reprinted from Int. J. Mol. Sci. 2013, *14(11)*, 23160–23187 287
http://www.mdpi.com/1422-0067/14/11/23160

Claudio Iaconetti, Clarice Gareri, Alberto Polimeni and Ciro Indolfi
Non-Coding RNAs: The "Dark Matter" of Cardiovascular Pathophysiology
Reprinted from Int. J. Mol. Sci. 2013, *14(10)*, 19987–20018 315
http://www.mdpi.com/1422-0067/14/10/19987

Tilde V. Eskildsen, Pia L. Jeppesen, Mikael Schneider, Anne Y. Nossent, Maria B. Sandberg, Pernille B. L. Hansen, Charlotte H. Jensen, Maria L. Hansen, Niels Marcussen, Lars M. Rasmussen, Peter Bie, Ditte C. Andersen and Søren P. Sheikh
Angiotensin II Regulates microRNA-132/-212 in Hypertensive Rats and Humans
Reprinted from Int. J. Mol. Sci. 2013, *14(6)*, 11190–11207 347
http://www.mdpi.com/1422-0067/14/6/11190

2.4. ncRNAs in neurological disorders

Chiara Fenoglio, Elisa Ridolfi, Daniela Galimberti and Elio Scarpini
An Emerging Role for Long Non-Coding RNA Dysregulation in Neurological Disorders
Reprinted from Int. J. Mol. Sci. 2013, *14(10)*, 20427–20442 365
http://www.mdpi.com/1422-0067/14/10/20427

2.5. ncRNAs in muscle dystrophies

Daniela Erriquez, Giovanni Perini and Alessandra Ferlini
Non-Coding RNAs in Muscle Dystrophies
Reprinted from Int. J. Mol. Sci. 2013, *14 (10)*, 19681–19704 381
http://www.mdpi.com/1422-0067/14/10/19681

2.6. Role of ncRNAs in cancer

Guorui Deng and Guangchao Sui
Noncoding RNA in Oncogenesis: A New Era of Identifying Key Players
Reprinted from Int. J. Mol. Sci. 2013, *14(9)*, 18319–18349 405
http://www.mdpi.com/1422-0067/14/9/18319

Raymond Wai-Ming Lung, Joanna Hung-Man Tong and Ka-Fai To
Emerging Roles of Small Epstein-Barr Virus Derived Non-Coding RNAs in Epithelial Malignancy Reprinted from Int. J. Mol. Sci. 2013, *14(9)*, 17378–17409 437
http://www.mdpi.com/1422-0067/14/9/17378

Federica Calore, Francesca Lovat and Michela Garofalo
Non-coding RNAs and Cancer
Reprinted from Int. J. Mol. Sci. 2013, *14(8)*, 17085–17110 469
http://www.mdpi.com/1422-0067/14/8/17085

Nina Hauptman and Damjan Glavač
Long Non-Coding RNA in Cancer
Reprinted from Int. J. Mol. Sci. 2013, *14(3)*, 4655–4669 495
http://www.mdpi.com/1422-0067/14/3/4655

Xiangsheng Li, Zhichao Zhang, Ming Yu, Liqi Li, Guangsheng Du, Weidong Xiao and Hua Yang
Involvement of miR-20a in Promoting Gastric Cancer Progression by Targeting Early Growth Response 2 (EGR2)
Reprinted from Int. J. Mol. Sci. 2013, *14(8)*, 16226–16239 511
http://www.mdpi.com/1422-0067/14/8/16226

Toshihiro Nishizawa and Hidekazu Suzuki
The Role of microRNA in Gastric Malignancy
Reprinted from Int. J. Mol. Sci. 2013, *14(5)*, 9487–9496 525
http://www.mdpi.com/1422-0067/14/5/9487

Andoni Garitano-Trojaola, Xabier Agirre, Felipe Prósper and Puri Fortes
Long Non-Coding RNAs in Haematological Malignancies
Reprinted from Int. J. Mol. Sci. 2013, *14(8)*, 15386–15422 535
http://www.mdpi.com/1422-0067/14/8/15386

Daniela Schwarzenbacher, Marija Balic and Martin Pichler
The Role of MicroRNAs in Breast Cancer Stem Cells
Reprinted from Int. J. Mol. Sci. 2013, *14(7)*, 14712–14723 573
http://www.mdpi.com/1422-0067/14/7/14712

Bethany N. Hannafon and Wei-Qun Ding
Intercellular Communication by Exosome-Derived microRNAs in Cancer
Reprinted from Int. J. Mol. Sci. 2013, *14(7)*, 14240–14269 585
http://www.mdpi.com/1422-0067/14/7/14240

Chia-Hsien Lee, Wen-Hong Kuo, Chen-Ching Lin, Yen-Jen Oyang, Hsuan-Cheng Huang and Hsueh-Fen Juan
MicroRNA-Regulated Protein-Protein Interaction Networks and Their Functions in Breast Cancer
Reprinted from Int. J. Mol. Sci. 2013, *14(6)*, 11560–11606 615
http://www.mdpi.com/1422-0067/14/6/11560

Ting Shuang, Chunxue Shi, Shuang Chang, Min Wang and Cui Hong Bai
Downregulation of miR-17~92 Expression Increase Paclitaxel Sensitivity in Human Ovarian Carcinoma SKOV3-TR30 Cells via BIM Instead of PTEN
Reprinted from Int. J. Mol. Sci. 2013, *14(2)*, 3802–3816 663
http://www.mdpi.com/1422-0067/14/2/3802

2.7. snoRNAs and cancer

Annalisa Pacilli, Claudio Ceccarelli, Davide Treré and Lorenzo Montanaro
SnoRNA U50 Levels Are Regulated by Cell Proliferation and rRNA Transcription
Reprinted from Int. J. Mol. Sci. 2013, *14(7)*, 14923–14935 679
http://www.mdpi.com/1422-0067/14/7/14923

2.8. miRNAs and multiple sclerosis

Michael Hecker, Madhan Thamilarasan, Dirk Koczan, Ina Schröder, Kristin Flechtner, Sherry Freiesleben, Georg Füllen, Hans-Jürgen Thiesen and Uwe Klaus Zettl
MicroRNA Expression Changes during Interferon-Beta Treatment in the Peripheral Blood of Multiple Sclerosis Patients
Reprinted from Int. J. Mol. Sci. 2013, *14(8)*, 16087–16110 693
http://www.mdpi.com/1422-0067/14/8/16087

2.9. Immune system

Graziella Curtale and Franca Citarella
Dynamic Nature of Noncoding RNA Regulation of Adaptive Immune Response
Reprinted from Int. J. Mol. Sci. 2013, *14(9)*, 17347–17377 721
http://www.mdpi.com/1422-0067/14/9/17347

Shaoqing Yu, Ruxin Zhang, Chunshen Zhu, Jianqiu Cheng, Hong Wang and Jing Wu
MicroRNA-143 Downregulates Interleukin-13 Receptor Alpha1 in Human Mast Cells
Reprinted from Int. J. Mol. Sci. 2013, *14(8)*, 16958–16969 755
http:// www.mdpi.com/1422-0067/14/8/16958

2.10. Genetic disorders

Emilia Kozlowska, Wlodzimierz J. Krzyzosiak and Edyta Koscianska
Regulation of Huntingtin Gene Expression by miRNA-137, -214, -148a, and Their Respective isomiRs
Reprinted from Int. J. Mol. Sci. 2013, *14(8)*, 16999–167016 767
http://www.mdpi.com/1422-0067/14/8/16999

2.11. miRNAs as risk factors in obesity

Jideng Ma, Shuzhen Yu, Fengjiao Wang, Lin Bai, Jian Xiao, Yanzhi Jiang, Lei Chen, Jinyong Wang, Anan Jiang, Mingzhou Li and Xuewei Li
MicroRNA Transcriptomes Relate Intermuscular Adipose Tissue to Metabolic Risk
Reprinted from Int. J. Mol. Sci. 2013, *14(4)*, 8611–8624 787
http://www.mdpi.com/1422-0067/14/4/8611

2.12. Aortic Aneurysms

Lars Maegdefessel, Joshua M. Spin, Matti Adam, Uwe Raaz, Ryuji Toh, Futoshi Nakagami and Philip S. Tsao
Micromanaging Abdominal Aortic Aneurysms
Reprinted from Int. J. Mol. Sci. 2013, *14(7)*, 14374–14394 801
http://www.mdpi.com/1422-0067/14/7/14374

Volume 2

Nicholas Delihas
Preface *Guest Editor* III

Chapter 3. ncRNAs and hematopoietic and stem cell differentiation

Franck Morceau, Sébastien Chateauvieux, Anthoula Gaigneaux, Mario Dicato and Marc Diederich
Long and short non-coding RNAs as regulators of hematopoietic differentiation
Reprinted from Int. J. Mol. Sci. 2013, *14(7)*, 14744–14770 805
http://www.mdpi.com/1422-0067/14/7/14744

Alessandro Rosa and Ali H. Brivanlou
Regulatory non-coding RNAs in pluripotent stem cells
Reprinted from Int. J. Mol. Sci. 2013, *14(7)*, 14346–14373 833
http://www.mdpi.com/1422-0067/14/7/14346

Marica Battista, Anna Musto, Angelica Navarra, Giuseppina Minopoli, Tommaso Russo and Silvia Parisi
miR-125b Regulates the Early Steps of ESC Differentiation through Dies1 in a TGF-Independent Manner
Reprinted from Int. J. Mol. Sci. 2013, *14(7)*, 13482–13496 861
http://www.mdpi.com/1422-0067/14/7/13482

Chapter 4. Tendon adhesion and siRNAs

Hongjiang Ruan, Shen Liu, Fengfeng Li, Xujun Li and Cunyi Fan
Prevention of Tendon Adhesions by ERK2 Small Interfering RNAs
Reprinted from Int. J. Mol. Sci. **2013**, *14(2)*, 4361–4371 877
http://www.mdpi.com/1422-0067/14/2/4361

Chapter 5. ncRNAs and laser therapy

Toshihiro Kushibiki, Takeshi Hirasawa, Shinpei Okawa and Miya Ishihara
Regulation of miRNA Expression by Low-Level Laser Therapy (LLLT) and Photodynamic Therapy (PDT)
Reprinted from Int. J. Mol. Sci. **2013**, *14(7)*, 13542–13558 889
http://www.mdpi.com/1422-0067/14/7/13542

Chapter 6. CRISPR system

Hagen Richter, Lennart Randau and André Plagens
Exploiting CRISPR/Cas: Interference Mechanisms and Applications
Reprinted from Int. J. Mol. Sci. **2013**, *14(7)*, 14518–14531 907
http://www.mdpi.com/1422-0067/14/7/14518

Chapter 7. Plant and Fungal ncRNAs

Mikhail M. Pooggin
How Can Plant DNA Viruses Evade siRNA-Directed DNA Methylation and Silencing?
Reprinted from Int. J. Mol. Sci. 2013, *14(8)*, 15233–15259 921
http://www.mdpi.com/1422-0067/14/8/15233

Virginie Gébelin, Julie Leclercq, Songnian Hu, Chaorong Tang and Pascal Montoro
Regulation of *MIR* Genes in Response to Abiotic Stress in *Hevea brasiliensis*
Reprinted from Int. J. Mol. Sci. 2013, *14(10)*, 19587–19604 949
http://www.mdpi.com/1422-0067/14/10/19587

Akihiro Matsui, Anh Hai Nguyen, Kentaro Nakaminami and Motoaki Seki
Arabidopsis Non-Coding RNA Regulation in Abiotic Stress Responses
Reprinted from Int. J. Mol. Sci. 2013, *14(11)*, 22642–22654 967
http://www.mdpi.com/1422-0067/14/11/22642

Li-Ling Lin, Chia-Chi Wu, Hsuan-Cheng Huang, Huai-Ju Chen, Hsu-Liang Hsieh and Hsueh-Fen Juan
Identification of MicroRNA 395a in 24-Epibrassinolide-Regulated Root Growth of *Arabidopsis thaliana* Using MicroRNA Arrays
Reprinted from Int. J. Mol. Sci. 2013, *14(7)*, 14270–14286 981
http://www.mdpi.com/1422-0067/14/7/14270

Yong Zhuang, Xiao-Hui Zhou and Jun Liu
Conserved miRNAs and Their Response to Salt Stress in Wild Eggplant *Solanum linnaeanum* Roots
Reprinted from Int. J. Mol. Sci. 2014, *15(1)*, 839–849 999
http://www.mdpi.com/1422-0067/15/1/839

Francisco E. Nicolás and Rosa M. Ruiz-Vázquez
Functional Diversity of RNAi-Associated sRNAs in Fungi
Reprinted from Int. J. Mol. Sci. 2013, *14(8)*, 15348–15360 1011
http://www.mdpi.com/1422-0067/14/8/15348

Chapter 8. 3' UTRs of mRNA

Eva Michalova, Borivoj Vojtesek and Roman Hrstka
Impaired Pre-mRNA Processing and Altered Architecture of 3' Untranslated Regions Contribute to the Development of Human Disorders
Reprinted from Int. J. Mol. Sci. **2013**, *14(8)*, 15681–15694 1025
http://www.mdpi.com/1422-0067/14/8/15681

Reprinted from *IJMS*. Cite as: Curtale, G.; Citarella, F. Dynamic Nature of Noncoding RNA Regulation of Adaptive Immune Response. *Int. J. Mol. Sci.* **2013**, *14*, 17347-17377.

Preface

These are exciting times for RNA molecular biologists! With the discovery of thousands of new non-coding RNA (ncRNA) transcripts in the last few years, and especially the new human genome transcripts, great opportunities and challenge are provided for determining functions in normal and disease states. This text is an outgrowth of a special issue of *IJMS* devoted to regulation by non-coding RNAs and contains both original research and review articles. In all there are 50 peer-reviewed articles presented that were submitted to the Journal within a period of 8 months. An attempt has been made to provide an up-to-date analysis of this very fast moving field and to cover regulatory roles of both microRNAs and long non-coding RNAs. Multifaceted functions of these RNAs in normal cellular processes, as well as in disease progression, are highlighted. We hope the readers will enjoy the articles and find the concepts presented challenging.

Nicholas Delihas

Guest Editor

Reprinted from *IJMS*. Cite as: Stoller, M.L.; Chang, H.C.; Fekete, D.M. Bicistronic Gene Transfer Tools for Delivery of miRNAs and Protein Coding Sequences. *Int. J. Mol. Sci.* **2013**, *14*, 18239-18255.

Article

Bicistronic Gene Transfer Tools for Delivery of miRNAs and Protein Coding Sequences

Michelle L. Stoller [1,2], Henry C. Chang [1,2] and Donna M. Fekete [1,2,*]

[1] Department of Biological Sciences, Purdue University, 915 W State St, West Lafayette, IN 47907-1392, USA; E-Mails: mlstolle@purdue.edu (M.L.S.); hcchang@purdue.edu (H.C.C.)

[2] Purdue University Center for Cancer Research, Purdue University, 201 S University Dr, West Lafayette, IN 47907-2064, USA

***** Author to whom correspondence should be addressed; E-Mail: dfekete@purdue.edu; Tel.: +1-765-496-3058; Fax: +1-765-494-0876.

Received: 22 July 2013; in revised form: 8 August 2013 / Accepted: 13 August 2013 / Published: 5 September 2013

Abstract: MicroRNAs (miRNAs) are a category of small RNAs that modulate levels of proteins via post-transcriptional inhibition. Currently, a standard strategy to overexpress miRNAs is as mature miRNA duplexes, although this method is cumbersome if multiple miRNAs need to be delivered. Many of these miRNAs are found within introns and processed through the RNA polymerase II pathway. We have designed a vector to exploit this naturally-occurring intronic pathway to deliver the three members of the sensory-specific miR-183 family from an artificial intron. In one version of the vector, the downstream exon encodes the reporter (GFP) while another version encodes a fusion protein created between the transcription factor Atoh1 and the hemaglutinin epitope, to distinguish it from endogenous Atoh1. *In vitro* analysis shows that the miRNAs contained within the artificial intron are processed and bind to their targets with specificity. The genes downstream are successfully translated into protein and identifiable through immunofluorescence. More importantly, Atoh1 is proven functional through *in vitro* assays. These results suggest that this cassette allows expression of miRNAs and proteins simultaneously, which provides the opportunity for joint delivery of specific translational repressors (miRNA) and possibly transcriptional activators (transcription factors). This ability is attractive for future gene therapy use.

Keywords: miRNAs; gene therapy; miR-183 family

---

1. Introduction

MicroRNAs (miRNAs) are small non-coding RNAs that are usually between 22 and 24 nucleotides in length. Each miRNA contains a seed region defined by nucleotides 2–7/8 that is perfectly complementary to a sequence found within its target messenger RNA (mRNA). This bond allows the miRNA to control the levels of its target proteins by downregulating the translation or stability of the target mRNA [1]. Since the discovery of microRNAs (miRNAs), research has focused on identifying conserved miRNA families and determining how these small molecules regulate a multitude of cellular processes that occur during cancer [2,3] and development [4].

In development, a subset of miRNAs has received attention because their expression patterns are relatively specific for distinct cell types or organs. One approach to explore the function of such miRNAs is to either knockdown their levels or to force their overexpression *in vivo* or *in vitro*. Intracellular injection or transfection of miRNA mimics has been successful to overexpress mature miRNAs, although the elevated level of miRNA mimics is transient because they are not stably transduced. As an alternative, exogenous miRNAs can be stably expressed using two distinct transcriptional pathways. Some vectors use the RNA polymerase III (polIII) pathway via the U6 promoter to drive expression of pre-miRNA hairpins [5,6], while others use the RNA polymerase II (polII) pathway, for example to express two pre-miRNA sequences downstream of a tet-responsive PolII promoter [7]. A major drawback of these approaches is that cells overexpressing miRNAs cannot be easily identified, making subsequent phenotypic analysis difficult. To circumvent this problem, it is common to combine the delivery of these miRNA elements with some type of reporter gene using IRES (internal ribosomal entry sites) to make a bicistronic mRNA [2,8] or to deliver miRNAs and a reporter gene using two different promoters. In the latter case, a polII- or polIII-based promoter controls the production of the miRNA and a polII-based promoter drives expression of the reporter gene [9,10]. While the use of two promoters allows production of miRNAs and a protein-coding gene, the production of the two factors is not necessarily coordinated. Such a tenuous link between the relative levels of miRNAs and any associated reporter (such as GFP) could compromise the use of the latter as an estimate of the former in functional studies. An alternative approach to overexpress miRNAs is to generate vectors that resemble the 38% of endogenous miRNA genomic loci where miRNAs are found within the introns of protein-coding genes [11]. When used in this context, both miRNAs and an exogenous gene, such as a *GFP* reporter or cell-surface marker, can be placed under the control of the same polII-dependent promoter [12].

Here we describe the development and functional testing of an intronic cassette to deliver a small family of miRNAs, the miR-183 family, that is specifically expressed in primary sensory cells in variety of vertebrate sensory systems, including vision, hearing, taste, olfaction and somatosensory. The evolutionarily conserved miR-183 family miRNAs has three members (miR-183, -96 and -182) that are transcribed as a single polycistronic pri-miRNA [13]. Although our interest is in the role these miRNAs play in the specification of mechanosensory cells of the inner ear, members of this sensory-specific miRNA family are also upregulated in several different types of cancer [14–16].

MIR96 was the first miRNA locus to be associated with a hereditary human disease when it was linked to the DFNA50 locus in two families with dominant non-syndromic progressive hearing loss [17]. Each family has a point mutation in the seed region of *MIR96*, but at different nucleotides. A

third deafness allele of DFN50 maps to a location in the pre-miR-96 transcript that likely interferes with miRNA processing [18]. Further supporting the link between deafness and mutations in miR-96, a semidominant deaf mouse mutant (diminuendo) was found with yet a third point mutation in the seed region [19]. The physiological and anatomical defects, present in either heterozygous or homozygous diminuendo mice, indicate that hair cells (HC) fail to fully mature [20].

In mouse, *Mir183*, *96* and *182* are located within an intronic region on Chromosome 6, and are transcribed as a single polycistronic pri-miRNA [21,22]. This coordinated expression is restricted to HCs as they begin to differentiate in both mice and zebrafish [23–26], suggesting that these miRNAs participate in HC development. Indeed, morpholino-mediated knockdown of each of the three miRNAs in zebrafish caused smaller inner ear sensory organ size and reduced numbers of HCs 2 days after injection [26]. Furthermore, overexpression of miR-96 and miR-182, by injection of double-stranded miRNA mimics into one-celled zebrafish, generated duplicate inner ears and produced supernumerary and ectopic inner ear HCs [26].

In total, data from humans, mice and zebrafish argue that the miR-183 family is essential for proper HC development and maintenance. As such, they should be considered as potential therapeutic agents for treating deafness due to HC loss. The vast majority (90%) of hearing loss is categorized as sensorineural, of which the most common type results from the destruction or malformation of the HCs occupying the organ of Corti, while sparing the associated supporting cells. One therapeutic approach is to deliver the HC-promoting transcription factor, Atonal1 (Atoh1), to the supporting cells of damaged ears. This has met with some success in animal models [27,28], although further studies are needed. Since it has been established that initiation and maturation of HCs require a complex regulatory network to turn off and on certain genes [29], we reasoned that the reprogramming of supporting cells into HCs might be enhanced by combining the delivery of an activating factor (Atoh1) and repressive elements (the miR-183 family). As every miR-183 family member is present during HC formation, we desired a gene transfer strategy that could efficiently and simultaneously deliver all 3 miRNAs along with a known HC-specification gene (*Atoh1*) to the same target cell population.

We produced two vectors containing the entire miR-183 family within a single artificial intron located upstream of a protein-encoding exon. The exon encoded either GFP as a reporter gene or a traceable version of Atoh1. We demonstrated that this design facilitates the coordinated expression of all three mature miRNAs and the associated protein. The data suggest that by simply exploiting one of the natural miRNA production pathways, it is possible to simultaneously deliver multiple negative and positive regulatory elements. Since many cellular processes require the joint activation and repression of downstream pathways, this delivery system provides an opportunity to achieve that dual manipulation efficiently.

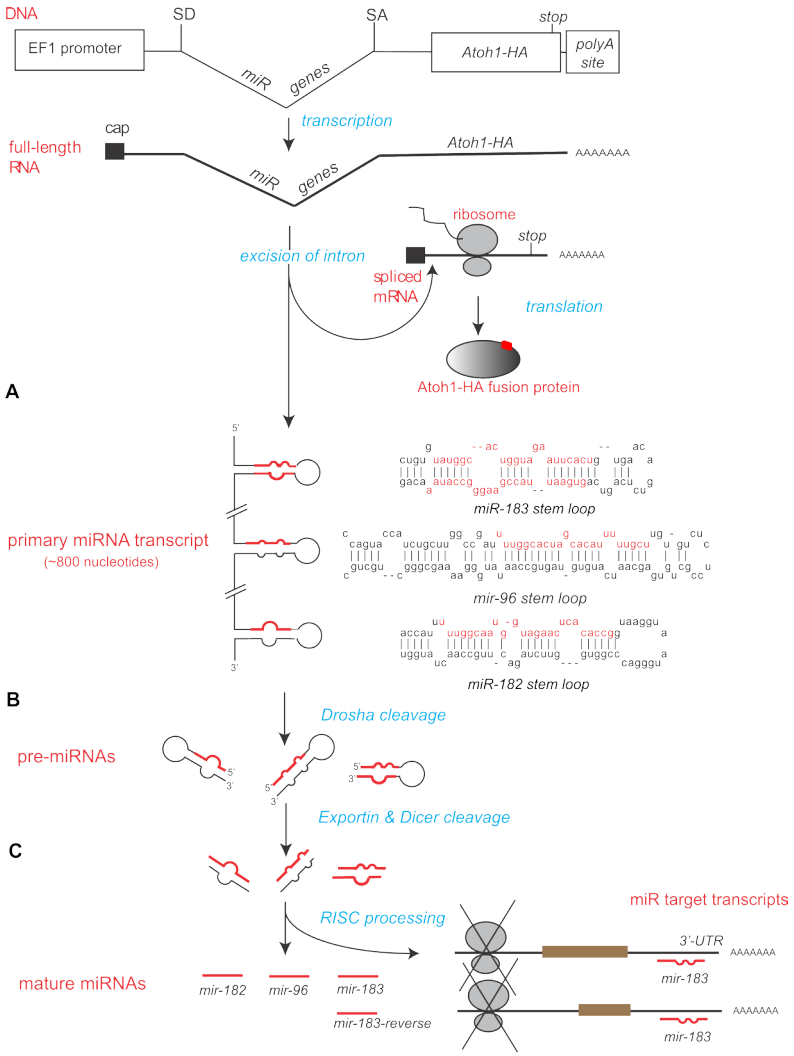
2. Results and Discussion

2.1. Construction of Bifunctional Atoh1-HA and miRNA Expression Vector

In order to coordinate the expression of the miRNAs and Atoh1 with high precision within the same cell, both elements are synthesized from the same RNA transcript. To accomplish this, an artificial intron containing the miRNAs is placed downstream of EF1α (human elongation factor 1 alpha;

pEF1X [30]) and upstream of *Atoh1* coding sequence (Figure 1). Within the mouse genome, about 3.5 kb of sequence separates *Mir182* from the nearest other family member *Mir96* [21]. To accommodate the size restrictions of certain delivery vectors planned for the future, such as the RCAS avian retrovirus and adeno-associated virus, we removed this large intervening stretch between *Mir182* and *Mir96* while retaining the natural pre-miRNA sequences for all 3 family members. Thus, all of the endogenous sequence between *Mir183* and *Mir96* (~120 bp) along with ~100 bp of sequence flanking the end of each pre-miRNA sequence was kept. Then, the pre-miR-182 sequence, with 120 bps flanking each end, was fused to the *Mir183*/*Mir96* fragment by PCR.

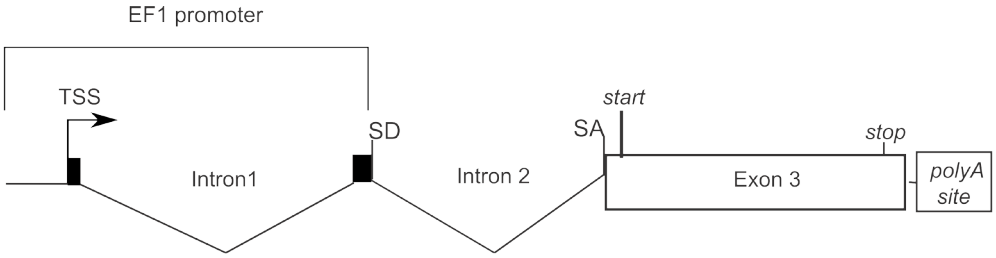
Figure 1. Bifunctional vector design and processing of transcripts. The vector consists of the EF1α promoter that will drive expression of the *miR-183* family of genes from the intron designated by the splice donor (SD) and splice acceptor (SA) site, and an exon encoding Atoh1 fused to the hemagluttinin influenza epitope (HA). Once the plasmid is transcribed into RNA, endogenous enzymes present in transfected cells should recognize the SD and SA sites to release the intron containing the primary miRNA transcript (**A**); clip it into the three distinct pre-miRNAs, export them from the nucleus (**B**); and then further process them into mature miRNAs (**C**). As the miRNAs follow their own maturation pathway, the spliced Atoh1-HA-encoding polyA+ transcript is exported from the nucleus and processed as mRNA.



The combined pri-miRNA sequences were inserted into the artificial intron sequence (intron discussed by Lin and colleagues [31,32]). This artificial intron of only ~100 bps contains a splice donor site at the 5' end of the pri-miRNAs. The 3' end flanking the pri-miRNAs houses a branch point domain, polypyrimidine tract, and splice acceptor site. The polypyrimidine tract allows spliceosome assembly, while the branch point is necessary for the cell to recognize that a splicing event should occur to excise the element between the splice donor and acceptor sites.

Downstream of the miRNA intron is the murine *Atoh1* coding region. This *Atoh1* sequence was proven bioactive by its ability to induce ectopic HCs in utero [33]. To facilitate the detection of Atoh1 expression from transfected plasmids, an influenza hemaglutinin (HA) peptide tag (YPYDVPDYA) was fused in-frame to the *Atoh1* coding sequence. Figure 2 displays the overall design of the resulting plasmid, pEF1X-sd-miR183F-sa-Atoh1-HA, hereafter referred to as p183F-Atoh1-HA. Figure 2 also provides details for the introns and exons of the other constructs and their abbreviated names that will be introduced below.

Figure 2. Content and design of overexpression vectors. Each overexpression vector discussed in the paper is listed with its formal name, abbreviated name, and contents. Black boxes represent exons. Intron 1 and exons 1 and 2 were present within the plasmid backbone prior to modification. Checkmarks indicate the presence of artificial intronic flanking sequences. Empty spaces indicate a specific component is not found within that particular vector. TSS: transcription start site.



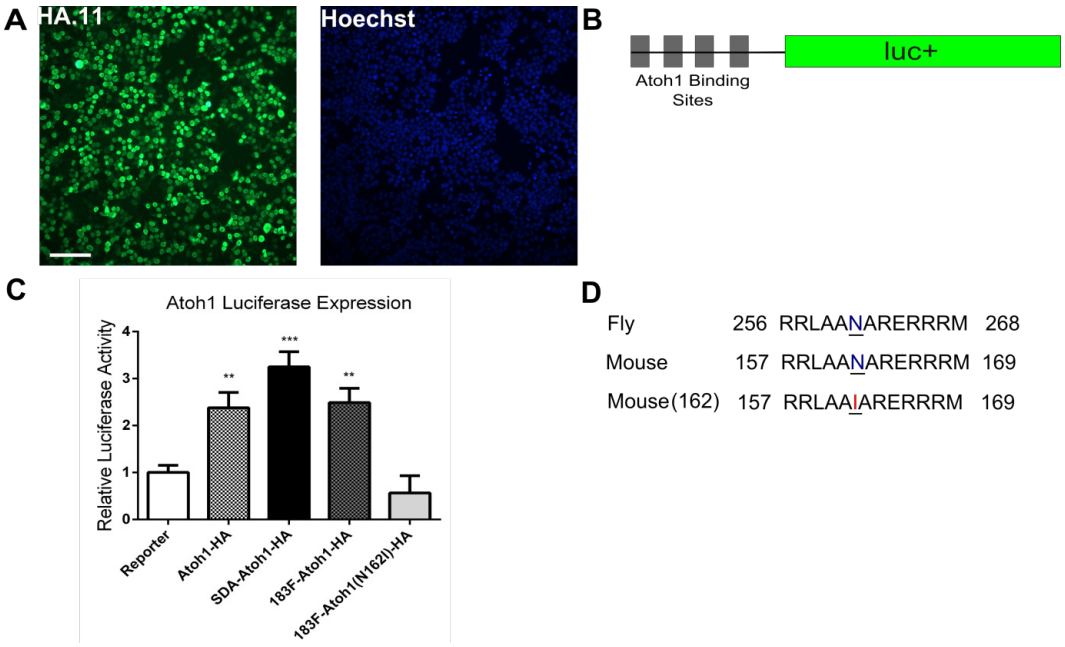
Formal Names	Abbreviated Names	SD	Intron 2	SA	Exon 3
pEF1X-Atoh1-HA	pAtoh1-HA				Atoh1-HA
pEF1X-sda-Atoh1-HA	pSDA-Atoh1-HA	√		√	Atoh1-HA
pEF1X-sd-miR183F-sa-Atoh1-HA	p183F-Atoh1-HA	√	miR-183 family	√	Atoh1-HA
pEF1X-sd-miR183F-sa-Atoh1(N162I)-HA	p183F-Atoh1(N162I)-HA	√	miR-183 family	√	Atoh1(N162I)-HA
pEF1X-sd-miR9-sa-Atoh1-HA	p9-Atoh1-HA	√	miR-9	√	Atoh1-HA
pEF1X-GFP	pGFP				GFP
pEF1X-sd-miR9-sa-GFP	p9-GFP	√	miR-9	√	GFP
pEF1X-sd-miR183F-sa-GFP	p183F-GFP	√	miR-183 family	√	GFP

2.2. Confirmation of Atoh1-HA Production and Function from a Bifunctional Cassette

To ascertain that Atoh1 is expressed from this bicistronic system, HEK293T cells transfected with p183F-Atoh1-HA were stained with anti-HA antibody. In cells 24 h after transfection, HA-positive staining was readily seen in the nuclei using immunofluorescence (Figure 3A), consistent

with the fact that Atoh1 is a transcription factor. No HA-positive staining was seen in mock-transfected cells, demonstrating that the signal in p183F-Atoh1-HA-transfected cells is specific (data not shown).

Figure 3. The Atoh1-HA fusion protein is functional and detectable by immunofluorescence. (**A**) Detection of Atoh1-HA fusion protein with HA.11 antibody in cells transfected with p183F-Atoh-HA. Scale bar = 100 microns; (**B**) Illustration of Atoh1 reporter construct; (**C**) Relative luciferase activity of cells transfected with Atoh1 reporter alone or with the indicated versions of the Atoh1-HA overexpression constructs. Luciferase activities are all referenced to cells transfected only with the reporter construct, which is set at 1.0. All constructs showed a significant increase in luminescence compared to the control except p183F-Atoh1(N162I)-HA. Each bar represents mean (±standard error) within each group. Each experiment was replicated at least three times; (**D**) Alignment of conserved Atoh1 segment between fly and mouse. Highlighted is the location of the amino acid mutated to make Atoh1 non-functional while maintaining the HA tag. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0001$.



While the immunofluorescence suggests that HA-tagged Atoh1 was expressed and properly localized, it remains possible that the addition of a peptide hinders its bioactivity. To ensure that the HA-tagged Atoh1 is functional, we tested its ability to activate the expression of a luciferase-based reporter gene (4E-box), which has a firefly (FF) luciferase coding sequence placed under the control of four Atoh1-binding sites [34]. In addition, hpRL-SV40 (Promega, Madison, WI, USA), a plasmid with Renilla luciferase driven by a constitutive promoter, was included for normalization. HEK293T cells transfected with pAtoh1-HA showed a 138% increase ($p = 0.0031$) in FF luminescence, compared to those transfected with the pEF1X empty vector. Similarly, cells transfected with pSDA-Atoh1-HA and p183F-Atoh1-HA showed significant increase in FF luciferase luminescence (Figure 3C; pSDA-Atoh1-HA, 225% increase, $p < 0.0001$; p183F-Atoh1-HA, 149% increase, $p = 0.0004$).

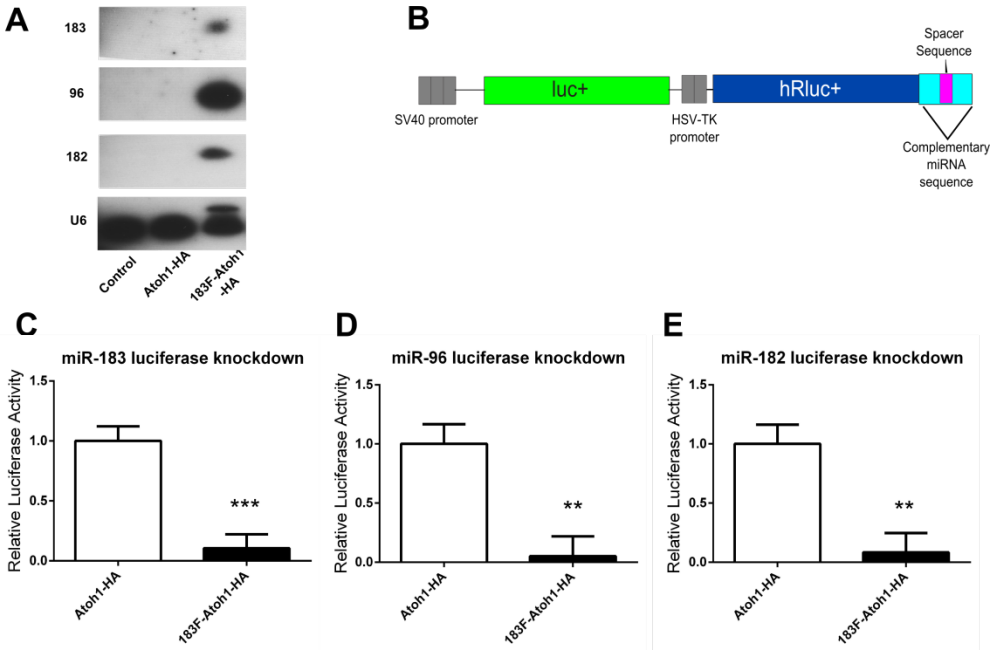
To ensure that this increase in FF luciferase expression required a functional Atoh1 protein, we took advantage of the fact that mutating a highly conserved asparagine in the homeobox domain has been shown to disrupt Atoh1 function [35]. We generated p183F-Atoh1(N162I)-HA, which expresses Atoh1-HA with the N162I substitution (this mutation is analogous to the point mutation affecting amino acid 261 in the fly) (Figure 3D). In cells transfected with p183F-Atoh1(N162I)-HA, mutant Atoh1-HA was still detectable by immunofluorescence (data not shown), although its ability to activate FF luciferase expression was diminished (69% decrease in luminescence relative to 3 vectors carrying the wild type *Atoh1* sequence; ANOVA; $p < 0.0001$). Interestingly, the N162I mutation seems to act as a dominant negative, as the luminescence in p183F-Atoh1(N162I)-HA transfected cells decreased by 43% compared to the control (Figure 3C; $p = 0.8213$). Atohl is believed to act as a heterodimer that binds to other bHLH (basic helix loop helix) transcription factors such as E47 [36]. It is likely that expression of N162I prevents the formation of functional Atoh1 heterodimers by depleting the pool of endogenous E47 or other such transcription factors. In any case, our results showed clearly that functional HA-tagged Atoh1 is expressed from these constructs.

2.3. Confirmation of miRNA Production and Function from a Bifunctional Cassette

To assess whether the miRNAs were synthesized from the artificial intron, small RNAs collected from HEK293T cells 30 h after p183F-Atoh1-HA transfection were analyzed by Northern blots. While none was detected in untransfected or pAtoh1-HA transfected cells, bands corresponding to mature miRNA of each 183 family member were seen in p183F-Atoh1-HA-transfected cells (Figure 4A). It is notable that the relative levels of the three miRNAs are distinctly different, with miR-96 most prominent. The observation that these family members are not uniformly expressed has also been reported for murine retina [22] and cochlea [24].

A dual luciferase assay system was used to confirm bioactivity of the miR-183 family miRNAs produced from the cassette. For each miRNA, a reporter construct was created beginning with psiCHECK-2 (Promega, Madison, WI, USA), into which two binding sites complementary to a mature miRNA and separated by a spacer sequence were inserted downstream of the *Renilla* luciferase gene (Figure 4B). The psiCHECK-2 vector also contains the firefly luciferase gene driven off a separate promoter, so that luminescence from the firefly protein serves as an internal transfection control. Reporters containing the miRNA binding sites for miR-182, miR-96 or miR-183 were co-transfected withp183F-Atoh1-HA into HEK293T cells. The luminescence ratio (corrected for transfection efficiency) from the experimental wells was compared to control wells transfected with the relevant miR-183 family reporter and the pAtoh1-HA plasmid lacking the miRNA intron. As shown in Figure 4C–E, each miRNA-reporter construct showed a significant knockdown in luminescence compared to its corresponding control (miR-96: 95% knockdown, $p = 0.0013$; miR-182: 92% knockdown, $p = 0.0008$; miR-183: 89% knockdown, $p < 0.0001$). Thus, all 3 miRNAs are produced from the bifunctional cassette and appear functional.

Figure 4. miRNAs are expressed from the 183F-Atoh1-HA vector and functional. (**A**) p183F-Atoh1-HA transfection leads to production of mature miR-183 family members. Untransfected HEK293T cells and cells transfected with pAtoh1-HA show no detectable expression of 183 family members; whereas miR-183, -96, and -182 are each detected in cells transfected with p183F-Atoh1-HA. U6 levels are provided as the loading control; (**B**) Illustration of miRNA-specific reporter plasmids. PsiCHECK-2 luciferase reporters contain 2 sites complementary to miR-96, -182, or -183 in the 3'UTR; (**C**–**E**) Knockdown of luciferase activity in reporters specific to each member of the miR-183 family; (**C**) Cells co-transfected with reporter containing miR-183 sites and p183F-Atoh1-HA showed a marked decrease in luciferase activity compared to wells transfected with the miR-183 reporter and pAtoh1-HA. Experiments in (**D**) and (**E**) were conducted in the same manner except with miR-96 or miR-182 complementary binding sites in the luciferase reporter. All showed significant decrease in luminescence. Each bar represents mean (±standard error) for each group. Each experiment was replicated at least three times. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0001$.



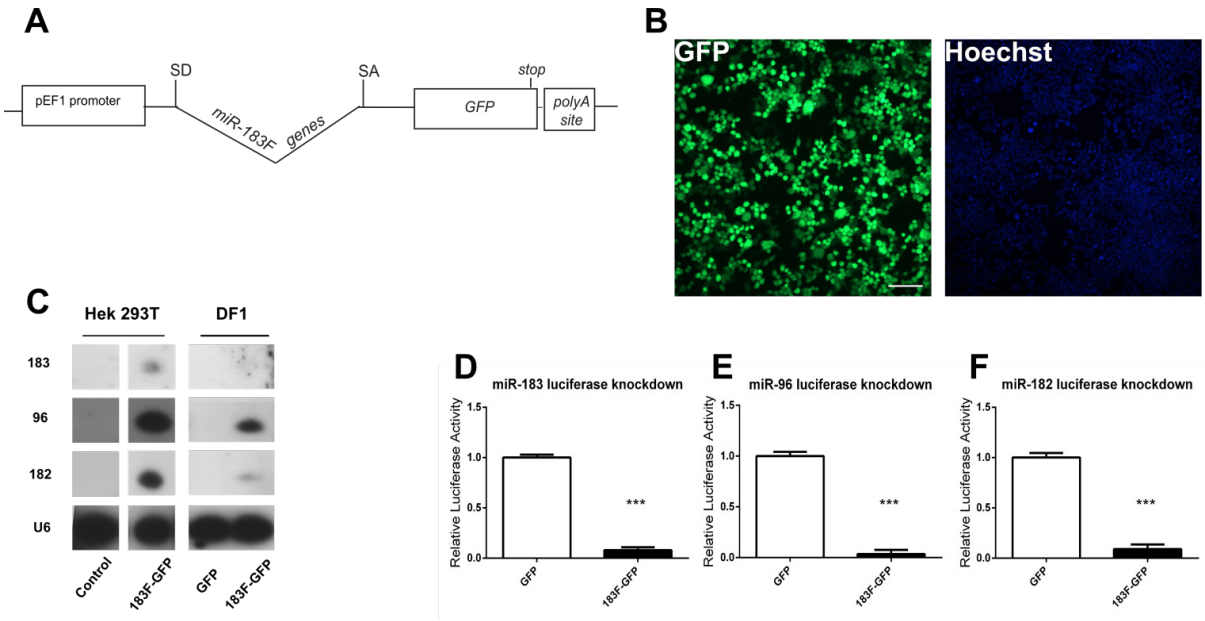
2.4. Overexpression of Functional miRNAs from GFP Expression Vectors

While our primary interest involved the dual expression of the miR-183 family and Atoh1, it was also desirable to express the miR-183 family alone to assess how much of an impact this family can have on HC development by themselves. To accomplish this, the *Atoh1-HA* coding region was replaced with *GFP*, which would allow the identification of cells expressing transfected miRNA constructs. Furthermore, as the design of p183F-GFP is the same as the p183F-Atoh1-HA (Figure 5A), phenotypic analysis using vectors with or without *Atoh1-HA* is less likely to be confounded by changes in the processing of the RNA transcripts that may affect transcript levels.

To test whether functional GFP protein is expressed from p183F-GFP, HEK293T cells were transfected and observed not only for direct GFP fluorescence but also after enhancing the signal with

anti-GFP antibodies. After 24 h, green emissions were detected from the majority of fixed, transfected cells both before (not shown) and after immunofluorescence (Figure 5B).

Figure 5. Plasmid containing *miR-183* family and the *GFP* gene produces functional miRNAs and GFP protein in HEK293T cells. (**A**) Vector design of p183F-GFP; (**B**) Visualization of GFP in cells transfected with p183F-GFP. Scale bar = 100 microns; (**C**) The miR-183 family is expressed from p183F-GFP in mammalian and avian cells. Cells transfected with p183F-GFP showed expression of mature miR-183, -96, and -182. Control (untransfected cells) and pGFP transfected cells exhibit no detectable miRNA. U6 levels serve as the loading control; (**D–F**) Luciferase activity is decreased by expression of miRNAs from p183F-GFP expressing vector; (**D**) Cells co-transfected with p183F-GFP and the psiCHECK-2 reporter containing sites complementary to miR-183 show a significant decrease in luciferase activity compared to cells co-transfected with the pGFP and reporter; (**E**) and (**F**) show results of experiments similar to (**D**) except the reporter contained different complementary binding sites: 96 for (**E**) and 182 for (**F**). Each bar represents mean (±standard error) for each group. Each experiment was replicated at least three times. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0001$.



All three mature miRNAs of the miR-183 family could be detected in HEK293T cells transfected with p183F-GFP but not in cells transfected with a GFP vector lacking the miRNA intron, as assessed by Northern blots (Figure 5C). Notably, miRNA expression levels appear to remain consistent regardless of the identity of downstream coding sequence (Figure 5C; HEK293T cells).

To ascertain whether avian cells are able to process and express mammalian miRNAs, small RNAs from DF1 cells (chicken embryo fibroblast cells) transfected with p183F-GFP or pGFP were analyzed by Northern blotting. While both p183F-GFP and pGFP transfected groups expressed GFP fluorescence (data not shown), only those transfected with p183F-GFP showed bands corresponding to miR-182, -96 and -183 (Figure 5C; DF1 cells). The relative levels of the miR-183 family members appeared lower in transfected DF-1 cells than HEK293T cells. This discrepancy could result from the

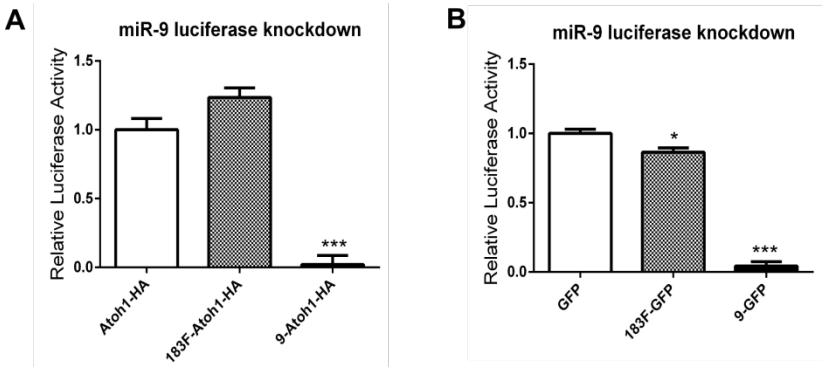
species difference of the transfected cells (chicken *vs.* human, respectively) or the difference in their respective tissue origins (embryonic day 10 fibroblasts *versus* fetal kidney, respectively). Nevertheless, our data demonstrated clearly that the miRNAs from p183F-GFP can be processed and produced in avian cells, allowing the option of using them in avian-specific vectors, like the RCAS retroviral vector [37].

Using the miRNA luciferase reporters discussed above, the function of the three miRNAs expressed from p183F-GFP was tested. Compared to HEK293T cells transfected with pGFP (which lacks the miRNA-producing intron), p183F-GFP transfection showed significant decrease in the expression of all three targets (miR-96: 97% knockdown, $p < 0.0001$; miR-182: 91% knockdown, $p < 0.0001$; miR-183: 92% knockdown, $p < 0.0001$) (Figure 5D–F). These observations demonstrated that miRNAs processed from the intron can successfully knockdown miRNA-specific targets.

2.5. MiRNAs Produced from Expression Vectors Bind with Specificity

To demonstrate the specificity of the knockdown mediated by these intronic miRNAs, we generated another luciferase-based reporter with two sites complementary to miR-9, a miRNA unrelated to the miR-183 family. Two intronic-miR-9 expression vectors with different downstream protein coding sequences (*Atoh1-HA or GFP*) were also generated to ensure this miR-9 reporter functions properly.

Figure 6. The miRNAs produced from both miRNA expression vectors bind to their targets with specificity. (**A**,**B**) Luciferase activity is decreased when vectors containing miR-9 are co-expressed with the miR-9 luciferase reporter but not when co-expressed with the miR-183 family expressing plasmids; (**A**) Luciferase activity of transfected cells containing the miR-9 luciferase reporter with pAtoh1-HA were compared to cells co-transfected with the miR-9 reporter and p183F-Atoh1-HA or p9-Atoh1-HA; (**B**) Experiments similar to those in A except control cells were co-transfected with miR-9 reporter and pGFP, while experimental cells were co-transfected with reporter and p183F-GFP or p9-GFP. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0001$. Bars represent mean (±standard error) for each group. Each experiment was replicated at least three times.



In cells transfected with miR-9 reporter, co-transfection of p9-Atoh1-HA or p9-GFP resulted in greater than 95% decrease in *Renilla* luciferase expression (98% decrease, $p < 0.0001$ for p9-Atoh1-HA; 96% decrease, $p < 0.0001$ for p9-GFP), demonstrating that these miR-9 expression vectors are functional. The comparable knockdown with both suggests that miR-9 vectors are similar

to the 183F-expressing plasmid series in being effectively processed from the artificial intron regardless of the identity of downstream coding sequences. Most importantly, co-transfection of p183F-Atoh1-HA or p183F-GFP, while capable of knocking down the expression of 183F-based reporters (see above), showed only negligible effects on the *Renilla* luciferase level from the miR-9 reporter (23% increase, $p = 0.10$ for p183F-Atoh1-HA; 13% decrease, $p = 0.02$ for p183F-GFP) (Figure 6). These data suggest that intronic miRNAs produced from 183F- and miR-9-expressing vectors regulate the expression of their target genes with high specificity.

3. Experimental Section

3.1. Bifunctional Plasmid Construction

The murine *Atoh1* coding sequence was PCR amplified from the pEF1-Atoh-IRES-GFP vector, a gift from John Brigande [33]. To facilitate cloning and protein detection, one primer introduced an EcoRI site, while the other primer added an influenza hemaglutinin (HA) tag (YPYDVPDYA) to the *C*-terminus of Atoh1 coding sequence and a *Not*I site (see Table S1). *Atoh1-HA* was cloned into pEF1X (provided by Cliff Ragsdale) [30] as an *Eco*RI–*Not*I fragment and the entire fusion was verified by sequencing (Purdue Genomics Center).

To construct an artificial miR183-containing intron, a *Sal*I–*Hind*III fragment containing a splice donor, three restriction sites *Xba*I, *BamH*I and *Xho*I, polypyrimidine tract, branch point, and a splice acceptor was generated by PCR and cloned into pME-MCS [38], generating pMCS-SDA. PCR primers were generated based on published sequences [31,32]. Then, PCR amplification was used to extract the primary miRNA DNA of all three members of the *miR-183* family from the mouse genome, and to flank the genomic DNA with *Spe*I and *Sal*I. This sequence was then inserted between the *Xba*I and *Xho*I sites found within the intron contained within pME-MCS-sda to create pME-MCS-sd-miR183F-sa. Then, Kpn1 was used to extract the artificial intron containing about 800 bp of mouse *miR-183* family genomic primary miRNA sequence from pME-MCS-sd-miR183F-sa. The *Kpn*I site was used to insert the intron with the *miR-183* family upstream of the Atoh1-HA fusion protein in pEF1X to generate pEF1X-sd-miR183F-sa-Atoh1-HA.

pEF1X-sd-miR183F-sa-GFP vector was constructed by extracting *GFP* from pAAV2.1-CMV-eGFP3-WPRE, provided by Alberto Auricchio [39], via PCR (see Table S3). Using the *Spe*I and *Not*I sites found on the 5' and 3' ends respectively, *GFP* was inserted downstream of the miRNA-containing artificial intron in pMCS-sd-miR183F-sa to create pMCS-sd-miR183F-sa-GFP. The pEF1X vector was converted to a Gateway Destination vector (Invitrogen, Carlsbad, CA, USA) by inserting cassette B from the Gateway Conversion Kit (Invitrogen, Carlsbad, CA, USA to create pEF1X-cB. A LR recombination reaction between pMCS-miR183F-GFP and pEF1X-cB generated pEF1X-sd-miR183F-sa-GFP.

pEF1X-sd-miR9-sa-Atoh1-HA and pEF1X-sd-miR9-sa-GFP were constructed in a similar manner as the aforementioned vectors except *Mir9-1* genomic sequence was inserted into the artificial intron instead of the *miR-183* family. PCR was used to extract the endogenous mouse *Mir9-1* sequence and flanking regions using primer sequences described in Shibata *et al.* [9].

For each miRNA reporter, primers were designed to contain two sequences that are fully complementary to the mature miRNA of interest (miR-183, -96, -182, or -9; see Table S4). These sites were separated by a 17 nucleotide spacer sequence. For all cases, the forward primer contains a *Xho*I site, while the reverse primer houses a *Not*I site to allow the resulting PCR fragments to be introduced downstream of the *Renilla* luciferase gene located in the psiCHECK-2 vector (Promega, Madison, WI, USA).

3.2. Mutation of Atoh1-HA Fusion Protein

To introduce the N162I substitution in Atoh-1, site-directed mutagenesis was performed with Quikchange 2XL (Agilent Technologies, Santa Clara, CA, USA) according to the manufacturer's instructions. Primers were designed to induce a point mutation to change the amino acid 162 from Asparagine to Isoleucine in the Atoh1-HA fusion protein encoded by pEF1X-sd-miR183F-sa-Atoh1-HA creating pEF1X-sd-183F-sa-Atoh1(N162I)-HA (see Table S2).

3.3. HEK293T Plasmid Transfection

HEK293T were cultured with modified DMEM supplemented with L-glutamine, antibiotics, and 10% calf serum. Using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA), cells seeded in 6-well plates were transfected with plasmids of interest. Collection time was assay dependent.

3.4. HEK293T Immunostain and Imaging

Cells transfected with pEF1X-sd-miR183F-sa-Atoh1-HA or pEF1X-sd-miR183F-sa-GFP were fixed 24 h post-transfection with 4% paraformaldehyde. The following primary antibodies (1:1000) were used: for detection of the HA tag, anti-HA.11 mouse IgG_1 monoclonal (Covance, Indianapolis, IN, USA); for detection of GFP, anti-GFP rabbit polyclonal (Molecular Probes, Eugene, OR, USA). Secondary antibodies (1:500) used were Alexa Fluor (Molecular Probes, Eugene, OR, USA) 488 anti-mouse IgG_1 and Alexa Fluor 488 anti-rabbit IgG. Immunostained cells were imaged under theE800 fluorescence microscope (Nikon, Elgin, IL, USA) with the 20× objective.

3.5. Atoh1 and MiRNA Luciferase Assays

The cells, 24 h after transfection, were lysed and luciferase activity was assessed using the dual luciferase assay kit (Promega, Madison, WI, USA) in the Luminoskan Ascent luminometer (Thermo Fisher Scientific, Waltham, MA, USA). For the Atoh1 luciferase assays, the firefly luciferase luminescence readings (at 560nm) were normalized to the *Renilla* luciferase readout (at 480 nm) to account for variation in transfection efficiency. In the case of the miRNA luciferase assays the ratio is inverted: the *Renilla* luciferase readout was normalized to the firefly luciferase readout. These ratios are expressed as relative luciferase activity. Experimental values were referenced to the control values which were arbitrarily set to one. Each treatment condition was conducted at least in duplicate. The experiments were repeated at least three times.

3.6. Northern Blots

HEK 293T/17 cells (abbreviated HEK293T cells; ATCC #CRL-11268) or UMNSAH-DF1 cells (abbreviated DF-1 cells; ATTC #CRL-12203) seeded in 35 mm plates were lysed ~30 h post-transfection and small RNAs were collected according to manufacturer's instructions using the PureLink miRNA Isolation Kit (Invitrogen, Carlsbad, CA, USA). Small RNA (300 ng) was probed for miR-183, -96, or -182 using the High Sensitive miRNA Northern Blot Assay Kit (Signosis, Santa Clara, CA, USA), a chemiluminescence system, according to manufacturer's instructions.

3.7. Statistical Analysis

All results are reported as mean ± standard error. The mean of each group is computed from measurements collected from at least three independent experiments. Statistical significance was determined by using a one-way analysis of variance with block (ANOVA), which was followed by Tukey's or Tukey-Kramer's multiple comparisons test (SAS 9.3, SAS Institute, www.sas.com). *p*-values below 0.05 were considered statistically significant.

4. Conclusions

Previous research has focused primarily on generating vectors to overexpress siRNAs: small non-coding RNAs that are a perfect complement to their targets. Researchers have shown that it is possible to express multiple siRNAs from one vector using a universal hairpin scaffold within or without an intron under the control of a PolII promoter [40–42]. Others have experimented with siRNA expression designs that allow joint delivery of siRNAs and protein coding regions under control of the same promoter without placement of the siRNA within an intron [43,44]. While these vectors do produce both small RNAs and protein, research shows that it is not the most efficient means of simultaneous delivery due to interference between miRNA/siRNA processing and protein production [45]. Thus, researchers sought to design a dual-delivery vector that would more reliably express a protein-coding gene and siRNA by inserting the siRNA within an artificial intron [31,32,45–47]. Capitalizing on the success of these vectors, others created miRNA vectors with a similar design to deliver one or two miRNAs within an endogenous intron upstream of a cell-surface marker [12]. We took that design one step further and created a dual-delivery vector to express a traceable transcription factor and multiple miRNA genes (in this case an entire miRNA family, using its endogenous sequence, contained within a single artificial intron). Luciferase assays showed that the transcription factor (Atoh1) fused to the HA tag is produced from the vector and is transcriptionally active, while immunofluorescence proved the fusion protein can be detected with HA.11 antibody. With this design, not only can the cells receiving the bifunctional vector be monitored, but artificially-expressed Atoh1 can be distinguished from the endogenous Atoh1 via its tag. Northern blots demonstrated that both human and chicken cells can process the single artificial intron containing 800 bp of endogenous sequence to produce each of the 3 mature miRNAs. These miRNAs were also shown to be functional and bind with specificity to their artificial target sequences.

In order to overexpress just the miRNA family and monitor its expression, we created a vector that replaced *Atoh1-HA* with *GFP*. We opted to keep the miRNAs in an intron upstream (rather than

downstream) of a coding exon because research has shown that expressing siRNAs within introns upstream of reporters leads to more stable *GFP* mRNA and better GFP expression [45]. Indeed, GFP expression was robust and efficient: every cell present within the transfected wells was GFP-positive. Northern blots also showed that each intronic miRNA is expressed, while luciferase assays illustrated that these miRNA repress their targets and bind with specificity.

The expression construct reported here uses a polII-based promoter to control the expression of an entire miRNA family within the context of an artificial intron and its downstream *GFP* reporter gene. Naturally co-expressed miRNA family members may function optimally if their overexpression is coordinated with one another for healthy cell function. This same design was used to express a functional transcription factor (in the place of *GFP*) and the miR-183 family, while still maintaining the ability to monitor transfected cells by creating the Atoh1-HA fusion protein. Each cassette was transferred into a Gateway-compatible (Invitrogen, Carlsbad, CA, USA) shuttle vector (unpublished) to facilitate their transfer into alternative delivery vectors such as lentivirus or adenovirus. Such vectors will allow delivery of the cassette into tissues that are difficult to access or to transfect, including into the mouse organ of Corti depleted of HCs. Thus, we believe this novel design can be manipulated for multiple overexpression uses to study a variety of different complex cellular systems and possibly for future therapeutic purposes, including hair cell regeneration.

Acknowledgments

This research was funded by NIH (R01DC002756), NIH/NIDCD (1F31DC011687), and NOHR (National Organization for Hearing Research). We would like to thank the following researchers for their reagents: John Brigande (Oregon Health Sciences University), Nissim Ben-Arie (The Hebrew University of Jerusalem), Alberto Auricchio (Telethon Institute of Genetics and Medicine), and Cliff Ragsdale (University of Chicago).

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Ying, S.Y.; Lin, S.L. Intron-mediated RNA interference and microRNA biogenesis. In *Methods of Molecular Biology, siRNA and miRNA Gene Silencing*; Sioud, M., Ed.; Humana Press: New York, NY, USA,2009; Volume 487, pp. 387–413.
2. He, L.; Thomson, J.M.; Hemann, M.T.; Hernando-Monge, E.; Mu, D.; Goodson, S.; Powers, S.; Cordon-Cardo, C.; Lowe, S.W.; Hannon, G.J.; *et al.* A microRNA polycistron as a potential human oncogene. *Nature* **2005**, *435*, 828–833.
3. Kasinski, A.L.; Slack, F.J. Epigenetics and genetics. MicroRNAs en route to the clinic: Progress in validating and targeting microRNAs for cancer therapy. *Nat. Rev. Cancer* **2011**, *11*, 849–864.
4. Lewis, M.A.; Steel, K.P. MicroRNAs in mouse development and disease. *Semin Cell Dev. Biol.* **2010**, *21*, 774–780.

5. Furukawa, N.; Sakurai, F.; Katayama, K.; Seki, N.; Kawabata, K.; Mizuguchi, H. Optimization of a microRNA expression vector for function analysis of microRNA. *J. Control Release* **2011**, *150*, 94–101.
6. Packer, A.N.; Xing, Y.; Harper, S.Q.; Jones, L.; Davidson, B.L. The bifunctional microRNA miR-9/miR-9* regulates REST and CoREST and is downregulated in Huntington's disease. *J. Neurosci.* **2008**, *28*, 14341–14346.
7. Rissland, O.S.; Hong, S.-J.; Bartel, D.P. MicroRNA destabilization enables dynamic regulation of the miR-16 family in response to cell-cycle changes. *Mol. Cell* **2011**, *43*, 993–1004.
8. Otaegi, G.; Pollock, A.; Sun, T. An optimized sponge for microRNA miR-9 affects spinal motor neuron development *in vivo*. *Front. Neurosci.* **2011**, *5*, 146.
9. Shibata, M.; Kurokawa, D.; Nakao, H.; Ohmura, T.; Aizawa, S. *MicroRNA-9* modulates Cajal-Retzius cell differentiation by suppressing Foxg1 expression in mouse medial pallium. *J. Neurosci.* **2008**, *28*, 10415–10421.
10. Zhao, C.; Sun, G.; Li, S.; Shi, Y. A feedback regulatory loop involving *microRNA-9* and nuclear receptor TLX in neural stem cell fate determination. *Nat.Struct. Mol. Biol.* **2009**, *16*, 365–371.
11. Chiang, H.R.; Schoenfeld, L.W.; Ruby, J.G.; Auyeung, V.C.; Spies, N.; Baek, D.; Johnston, W.K.; Russ, C.; Luo, S.; Babiarz, J.E.; *et al.* Mammalian microRNAs: Experimental evaluation of novel and previously annotated genes. *Gene Dev.* **2010**, *24*, 992–1009.
12. Amendola, M.; Passerini, L.; Pucci, F.; Gentner, B.; Bacchetta, R.; Naldini, L. Regulated and multiple miRNA and siRNA delivery into primary cells by a lentiviral platform. *Mol. Ther.* **2009**, *17*, 1039–1052.
13. Pierce, M.L.; Weston, M.D.; Fritzsch, B.; Gabel, H.W.; Ruvkun, G.; Soukup, G.A. *MicroRNA-183* family conservation and ciliated neurosensory organ expression. *Evo. Dev.* **2008**, *10*, 106–113.
14. Mihelich, B.L.; Khramtsova, E.A.; Arva, N.; Vaishnav, A.; Johnson, D.N.; Giangreco, A.A.; Martens-Uzunova, E.; Bagasra, O.; Kajdacsy-Balla, A.; Nonn, L. MiR-183-96-182 cluster is overexpressed in prostate tissue and regulates zinc homeostasis in prostate cells. *J. Biol. Chem.* **2011**, *286*, 44503–44511.
15. Hannafon, B.N.; Sebastiani, P.; de las Morenas, A.; Lu, J.; Rosenberg, C.L. Expression of microRNA and their gene targets are dysregulated in preinvasive breast cancer. *Breast Cancer Res.* **2011**, *13*, R24.
16. Zhu, W.; Liu, X.; He, J.; Chen, D.; Hunag, Y.; Zhang, Y.K. Overexpression of members of the microRNA-183 family is a risk factor for lung cancer: A case control study. *BMC Cancer* **2011**, *11*, 393.
17. Mencía, A.; Modamio-Høybjør, S.; Redshaw, N.; Morín, M.; Mayo-Merino, F.; Olavarrieta, L.; Aguirre, L.A.; del Castillo, I.; Steel, K.P.; Dalmay, T.; *et al.* Mutations in the seed region of human miR-96 are responsible for nonsyndromic progressive hearing loss. *Nat. Genet.* **2009**, *41*, 609–613.
18. Soldà, G.; Robusto, M.; Primignani, P.; Castorina, P.; Benzoni, E.; Cesarani, A.; Ambrosetti, U.; Asselta, R.; Duga, S. A novel mutation within the MIR96 gene causes non-syndromic inherited hearing loss in an Italian family by altering pre-miRNA processing. *Hum. Mol. Genet.* **2012**, *21*, 577–585.

19. Lewis, M.A.; Quint, E.; Glazier, A.M.; Fuchs, H.; de Angelis, M.H.; Langford, C.; van Dongen, S.; Abreu-Goodger, C.; Piipari, M.; Redshaw, N.; *et al*. An ENU-induced mutation of miR-96 associated with progressive hearing loss in mice. *Nat. Genet.* **2009**, *41*, 614–618.
20. Kuhn, S.; Johnson, S.L.; Furness, D.N.; Chen, J.; Ingham, N.; Hilton, J.M.; Steffes, G.; Lewis, M.A.; Zampini, V.; Hackney, C.M.; *et al.* MiR-96 regulates the progression of differentiation in mammalian cochlear inner and outer hair cells. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 2355–2360.
21. Xu, S.; Witmer, P.D.; Lumayag, S.; Kovacs, B.; Valle, D. MicroRNA (miRNA) transcriptome of mouse retina and identification of a sensory organ-specific miRNA cluster. *J. Biol. Chem.* **2007**, *282*, 25053–2366.
22. Lumayag, S.; Haldin, C.E.; Corbett, N.J.; Wahlin, K.J.; Cowan, C.; Turturro, S.; Larsen, P.E.; Kovacs, B.; Witmer, P.D.; Valle, D.; *et al.* Inactivation of the *microRNA-183/96/182* cluster results in syndromic retinal degeneration. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E507–E516.
23. Weston, M.D.; Pierce, M.L.; Jensen-Smith, H.C.; Fritzsch, B.; Rocha-Sanchez, S.; Beisel, K.W.; Soukup, G.A. *MicroRNA-183* family expression in hair cell development and requirement of microRNAs for hair cell maintenance and survival. *Dev. Dynam.* **2011**, *240*, 808–819.
24. Weston, M.D.; Pierce, M.L.; Rocha-Sanchez, S.; Beisel, K.W.; Soukup, G.A. MicroRNA gene expression in the mouse inner ear. *Brain Res.* **2006**, *1111*, 95–104.
25. Sacheli, R.; Nguyen, L.; Borgs, L.; Vandenbosch, R.; Bodson, M.; Lefebvre, P.; Malgrange, B. Expression patterns of miR-96, miR-182 and miR-183 in the development inner ear. *Gene Expr. Patterns* **2009**, *9*, 364–370.
26. Li, H.; Kloosterman, W.; Fekete, D.M. *MicroRNA-183* family members regulate sensorineural fates in the inner ear. *J. Neurosci.* **2010**, *30*, 3254–3263.
27. Izumikawa, M.; Minoda, R.; Kawamoto, K.; Abrashkin, K.A.; Swiderski, D.L.; Dolan, D.F.; Brough, D.E.; Raphael, Y. Auditory hair cell replacement and hearing improvement by *Atoh1* gene therapy in deaf mammals. *Nat. Med.* **2005**, *11*, 271–276.
28. Yang, S.-M.; Chen, W.; Guo, W.-W.; Jia, S.; Sun, J.-H.; Liu, H.-Z.; Young, W.-Y.; He, D.Z.Z. Regeneration of stereocilia of hair cells by forced Atoh1 expression in the adult mammalian cochlea. *PLoS One* **2012**, *7*, e46355.
29. Groves, A.K.; Zhang, K.D.; Fekete, D.M. The genetics of hair cell development and regeneration. *Annu. Rev. Neurosci.* **2013**, *36*, 361–381.
30. Agarwala, S.; Sanders, T.A.; Ragsdale, C.W. Sonic hedgehog control of size and shape in midbrain pattern formation. *Science* **2001**, *291*, 2147–2150.
31. Lin, S.L.; Chang, D.; Wu, D.Y.; Ying, S.Y. A novel RNA splicing-mediated gene silencing mechanism potential for genome evolution. *Biochem. Biophy. Res. Commun.* **2003**, *310*, 754–760.
32. Lin, S.-L.; Ying, S.-Y. New drug design for gene therapy—Taking advantage of introns. *Lett. Drug Des. Discov.* **2004**, *1*, 256–262.
33. Gubbels, S.P.; Woessner, D.W.; Mitchell, J.C.; Ricci, A.J.; Brigande, J.V. Functional auditory hair cells produced in the mammalian cochlea by in utero gene transfer. *Nature* **2008**, *455*, 537–541.
34. Krizhanovsky, V.; Soreq, L.; Kliminski, V.; Ben-arie, N. Math1 target genes are enriched with evolutionarily conserved clustered E-box binding sites. *J. Mol. Neurosci.* **2006**, *28*, 211–229.

35. Jarman, A.P.; Grell, E.H.; Ackerman, L.; Jan, L.J.; Jan, Y.N. Atonal is the proneural gene drosophila photoreceptors. *Nature* **1994**, *369*, 398–400.
36. Mulvaney, J.; Dabdoub, A. Atoh1, an essential transcription factor in neurogenesis and intestinal and inner ear development: Function, regulation, and context dependency. *JARO* **2012**, *13*, 281–293.
37. Hughes, S.H. The RCAS vector system. *Folia Biol.* **2004**, *50*, 107–119.
38. Kwan, K.M.; Fujimoto, E.; Grabher, C.; Mangum, B.D.; Hardy, M.E.; Campbell, D.S.; Parant, J.M.; Yost, H.J.; Kanki, J.P.; Chien, C.-B. The Tol2kit: A multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev. Dynam.* **2007**, *236*, 3088–3099.
39. Karali, M.; Manfredi, A.; Puppo, A.; Marrocco, E.; Gargiulo, A.; Allocca, M.; Corte, M.D.; Rossi, S.; Giunti, M.; Bacci, M.L.; *et al.* MicroRNA-restricted transgene expression in the retina. *PLoS One* **2011**, *6*, e22166.
40. Chung, K.-H. Polycistronic RNA polymerase II expression vectors for RNA interference based on BIC/miR-155. *Nucleic Acids Res.* **2006**, *34*, e53.
41. Shin, K.-J.; Wall, E.A.; Zavzavadjian, J.R.; Santat, L.A.; Liu, J.; Hwang, J.-I.; Rebres, R.; Roach, T.; Seaman, W.; Simon, M.I.; *et al.* A single lentiviral vector platform for microRNA-based conditional RNA interference and coordinated transgene expression. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13759–13764.
42. Zhu, X.; Santat, L.A.; Chang, M.S.; Liu, J.; Zavzavadjian, J.R.; Wall, E.A.; Kivork, C.; Simon, M.I.; Fraser, I.D. A versatile approach to multiple gene RNA interference using microRNA-based short hairpin RNAs. *BMC Mol. Biol.* **2007**, *8*, 98.
43. Stegmeier, F.; Hu, G.; Rickles, R.J.; Hannon, G.J.; Elledge, S.J. A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13212–13217.
44. Wang, X.; Liu, P.; Liu, H.; Yang, W.; Liu, Z.; Zhuo, Z.; Gao, Y. Delivery of interferons and siRNA targeting STAT3 using lentiviral vectors suppresses the growth of murine melanoma. *Cancer Gene Ther.* **2012**, *19*, 822–827.
45. Du, G.; Yonekubo, J.; Zeng, Y.; Osisami, M.; Frohman, M.A. Design of expression vectors for RNA interference based on miRNAs and RNA splicing. *FEBS J.* **2006**, *273*, 5421–5427.
46. Karwacz, K.; Bricogne, C.; MacDonald, D.; Arce, F.; Bennett, C.L.; Collins, M.; Escors, D. PD-L1 co-stimulation contributes to ligand-induced T cell receptor down-modulation on CD8+ T cells. *EMBO Mol. Med.* **2011**, *3*, 581–592.
47. Arce, F.; Breckpot, K.; Stephenson, H.; Karwacz, K.; Ehrenstein, M.R.; Collins, M.; Escors, D. Selective ERK activation differentiates mouse and human tolerogenic dendritic cells, expands antigen-specific regulatory T cells, and suppresses experimental inflammatory arthritis. *Arthritis Rheum.* **2011**, *63*, 84–95.

Reprinted from *IJMS*. Cite as: Mayr, F.; Heinemann, U. Mechanisms of Lin28-Mediated miRNA and mRNA Regulation—A Structural and Functional Perspective. *Int. J. Mol. Sci.* **2013**, *14*, 16532-16553.

Review

Mechanisms of Lin28-Mediated miRNA and mRNA Regulation—A Structural and Functional Perspective

Florian Mayr [1,2] and Udo Heinemann [1,2,*]

[1] Crystallography, Max-Delbrück Center for Molecular Medicine, Robert-Rössle Straße 10, Berlin 13125, Germany; E-Mail: florian.mayr@mdc-berlin.de

[2] Institute for Chemistry and Biochemistry, Freie Universität Berlin, Takustraße 6, Berlin 14195, Germany

***** Author to whom correspondence should be addressed; E-Mail: heinemann@mdc-berlin.de; Tel.: +49-30-9406-3420; Fax: +49-30-9406-2548.

Received: 31 May 2013; in revised form: 22 July 2013 / Accepted: 25 July 2013 / Published: 9 August 2013

Abstract: Lin28 is an essential RNA-binding protein that is ubiquitously expressed in embryonic stem cells. Its physiological function has been linked to the regulation of differentiation, development, and oncogenesis as well as glucose metabolism. Lin28 mediates these pleiotropic functions by inhibiting *let-7* miRNA biogenesis and by modulating the translation of target mRNAs. Both activities strongly depend on Lin28's RNA-binding domains (RBDs), an N-terminal cold-shock domain (CSD) and a C-terminal Zn-knuckle domain (ZKD). Recent biochemical and structural studies revealed the mechanisms of how Lin28 controls *let-7* biogenesis. Lin28 binds to the terminal loop of pri- and pre-*let-7* miRNA and represses their processing by Drosha and Dicer. Several biochemical and structural studies showed that the specificity of this interaction is mainly mediated by the ZKD with a conserved GGAGA or GGAGA-like motif. Further RNA crosslinking and immunoprecipitation coupled to high-throughput sequencing (CLIP-seq) studies confirmed this binding motif and uncovered a large number of new mRNA binding sites. Here we review exciting recent progress in our understanding of how Lin28 binds structurally diverse RNAs and fulfills its pleiotropic functions.

Keywords: Lin28; *let-7* miRNA; miRNA processing; RNA-binding protein; cold-shock domain; zinc-knuckle domain; TUTase; oncogene; stem cell

---

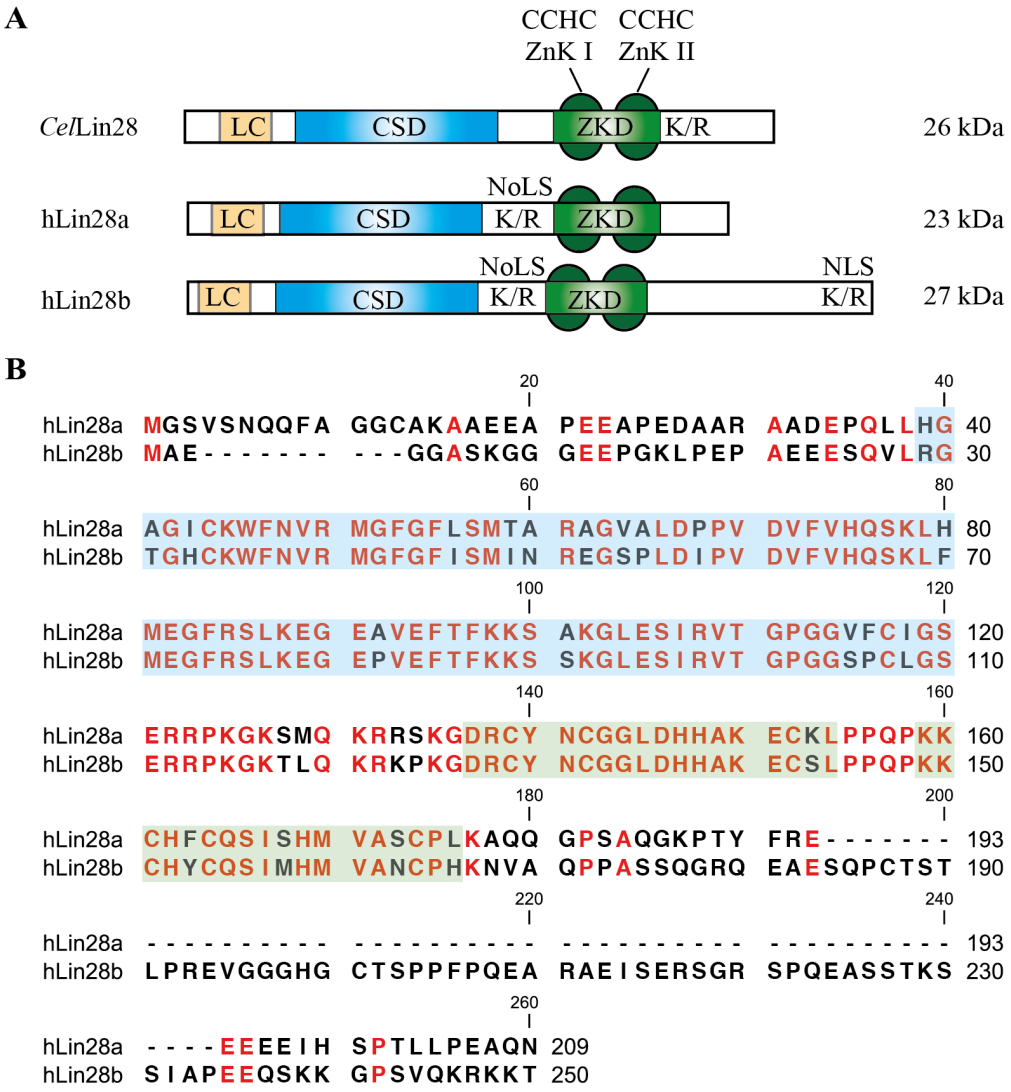
1. Introduction

Lin28 (cell lineage abnormal 28) is a conserved RNA-binding protein in higher eukaryotes that regulates several important cellular functions associated with development, glucose metabolism, differentiation and pluripotency. It was first described as a heterochronic gene in *Caenorhabditis elegans (C. elegans)*, since mutations within *lin-28* disturbed the developmental timing of the worm and accelerated differentiation of hypodermal seam cells and vulva stem cells [1,2]. Subsequent experiments revealed that Lin28 is expressed early in nematode embryonic and larval development, but its expression is down-regulated by *lin-4* and *let-7* miRNA as differentiation proceeds [2,3].

A similar expression pattern and physiological function was also shown for *Drosophila*, *Xenopus* and mammalian Lin28 [4]. The human paralogs Lin28a (routinely termed simply Lin28) and Lin28b encode for basic 23- or 28-kDa proteins that are highly expressed in embryonic stem cells (ESC) but are down-regulated upon differentiation of ESCs into embryoid bodies [5]. Reciprocally, Yu and colleagues used Lin28a, in conjunction with Oct4, Sox2 and Nanog, to reprogram adult human fibroblasts to induced pluripotent stem cells (iPSCs) [6]. A knockdown of Lin28a expression in mouse ESCs led to loss of Oct4 and Nanog expression, indicating an impaired self-renewal potential [7]. Increased Lin28a/Lin28b expression was reported in various hepatocellular and other carcinomas and was associated with poor patient prognosis [8–13]. Recently, Lin28a was linked to the regulation of developmental and metabolic processes. After ectopic overexpression of Lin28a mice developed a bigger size and delayed sexual maturation, whereas Lin28 knockout mice were smaller and died shortly after birth [14]. In addition, Lin28a overexpression was associated with increased insulin sensitivity and glucose metabolism, while a depletion of Lin28a resulted in insulin resistance and glucose intolerance [15].

On the molecular level, Lin28a and Lin28b act as both negative regulator of *let-7* miRNA biogenesis and post-transcriptional regulator of mRNA translation. Both activities strongly depend on Lin28's two RNA-binding domains (RBDs): an N-terminal cold-shock domain (CSD) and a C-terminal Zn-knuckle domain (ZKD) composed of two tandemly arranged retroviral-type CCHC Zn knuckles. The individual domain combination of both RBDs is unique in animals with the RBDs being highly conserved. The human Lin28 paralogs share an overall sequence identity of 65% (FFAS, [16]) and contain low-complexity regions at the N-terminus, a putative bipartite nucleolar localization sequence (NoLS) as well as a C-terminal nuclear localization signal (NLS) in the case of Lin28b [17] (Figure 1). Lin28a and Lin28b can localize to both cytosol and nucleus [2,4,17–21] and interact with primary (pri-) or precursor (pre-) *let-7* miRNAs thereby preventing their maturation [20,22,23]. In addition, binding of Lin28a to messenger ribonucleoprotein complexes containing translation initiation (eIF3B, eI4E) and elongation factors (EF1α, EF1α2), poly(A) binding proteins, Igf2bps and RNA helicase A was reported in various studies [18,20,24,25]. Under stress conditions, Lin28a was shown to localize to cytoplasmic stress granules and P-bodies where mRNA translation is temporally stalled [18]. Since a mutation in Lin28a's ZKD caused Lin28a to accumulate in the nucleus, it was suggested that Lin28a exits the nucleus in a complex with bound RNA and thus regulates the post-transcriptional processing of its target RNAs [18].

Figure 1. Domain organization of Lin28. (**A**) *Caenorhabditis elegans* (*Cel*) and human (h) Lin28a/Lin28b contain two RNA-binding domains (RBDs): an N-terminal cold-shock domain (CSD) and a C-terminal Zn-knuckle domain (ZKD) comprised of two retroviral type CCHC Zn knuckles (ZnK). Additionally, Lin28 harbors low-complexity sequences, Lys/Arg (K/R)-rich stretches, bipartite nuclear localization signals (NLS) or putative nucleolar localization sequences (NoLS); (**B**) Sequence alignment of hLin28a and hLin28B. Amino acids belonging to CSD or ZKD are shaded in blue or green, respectively.



2. Lin28 Blocks *let-7* Processing

The opposing expression pattern of Lin28 and *let-7* miRNA became initially apparent when studying *C. elegans* larval development [2–4,26,27]. At an early stage in larval development both Lin28 and pri-*let-7* are present, however, no levels of either pre-*let-7* or mature *let-7* can be detected, indicating a regulation at a post-transcriptional level [28]. As larval development proceeds, a heterogenic cascade involving *lin-4* miRNA and the *let-7* sisters *mir-48/84/241* lead to a relief of pri-*let-7* processing inhibition and to a subsequent down-regulation of Lin28 expression (reviewed in [29]). This inverse relationship between Lin28 and *let-7* miRNA is also present in mammalian cells,

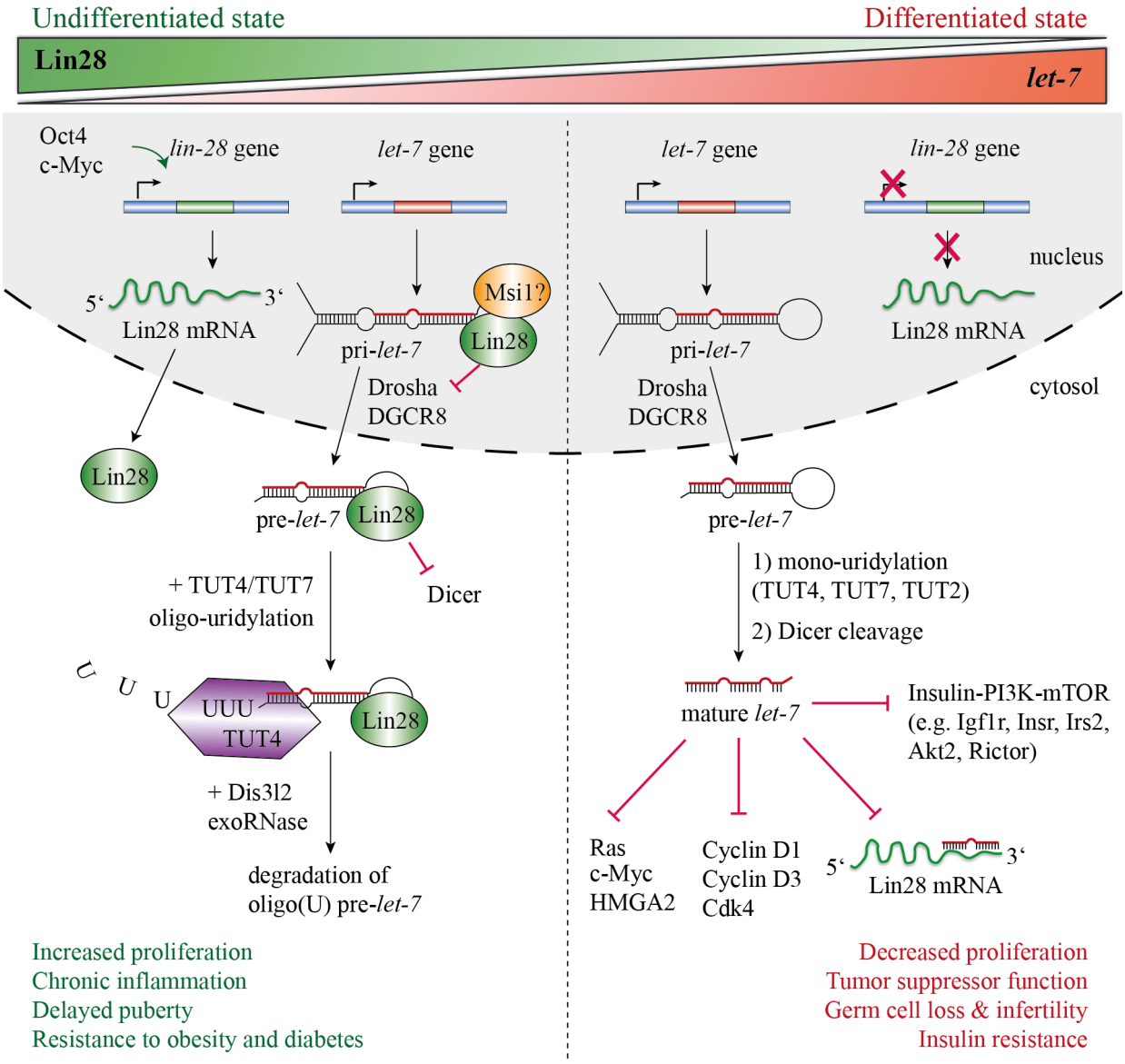
where Lin28a/b are mainly expressed in undifferentiated cells, and mature *let-7* is only detectable upon differentiation or tissue development [5,20,23]. Furthermore, levels of pri-*let-7* remain constant throughout the entire differentiation or development process suggesting a negative regulation of *let-7* biogenesis by Lin28a/b in stem or progenitor cells [30–32]. Purification of pre-*let-7* bound complexes and subsequent analysis via mass spectrometry revealed that both human Lin28 paralogs specifically associate with pri- or pre-*let-7 in vivo* [20,22,23]. Moreover, *in vitro* purified Lin28a could inhibit pri- and pre-*let-7* processing by Drosha and Dicer by binding to the double-stranded stem close to the Dicer cleavage site and the pre-element (preE, terminal loop or hairpin) [33]. Mutations within Lin28's CSD and ZKD impaired pre-*let-7* binding and inhibition of Dicer processing, suggesting a competitive relationship between Lin28 and Dicer [33–35]. Moreover, recent studies provided evidence that Lin28a/b can induce a structural change within pre-*let-7*'s preE, thereby leading to an opening of the double-stranded stem including the Dicer cleavage site [34–36].

Heo and colleagues revealed an additional inhibition mode of *let-7* miRNA processing, which irreversibly targets pre-*let-7* to a decay pathway [7,37]. They demonstrated that Lin28a/b induce oligo-uridylation of pre-*let-7*'s 3' overhang. Oligo-uridylated pre-*let-7* is resistant to Dicer cleavage given that Dicer normally recognizes a 2-nt 3' overhang in miRNAs via its PAZ domain. Thus, Dicer is unable to recognize the elongated 3' overhang and to process pre-*let-7*. Furthermore, it was reported that oligo-uridylated RNAs recruit 3'–5' exonucleases and are targeted for decay [38,39]. Indeed, oligo-uridylated pre-*let-7* was more rapidly degraded than unmodified pre-*let-7* [37]. Recently, Chang and colleagues identified the 3'–5' exonuclease Dis3l2 that catalyzes the decay of oligo-uridylated pre-*let-7* in mouse ESCs [40]. Consistent with this, a knockdown of Dis3l2 in mouse ESCs caused an accumulation of uridylated pre-*let-7*. Oligo-uridylation of pre-*let-7* is catalyzed by the non-canonical poly(A) polymerase TUT4 (terminal uridyl transferase 4/Zcchc11) and to a minor extent by TUT7 (Zcchc6) in a Lin28-dependent manner [7,41,42]. Interestingly, these enzymes catalyze mono-uridylation of pre-miRNAs with a 1-nt 3' overhang (like most pre-*let-7* family members) in the absence of Lin28, thereby enhancing Dicer-mediated processing [43]. However, in the presence of Lin28, pre-*let-7* and other miRNAs containing a GGAG motif within their preE were subjected to oligo-uridylation. Upon mutation of this motif, both Lin28 binding and oligo-uridylation were impaired, indicating that the GGAG motif is essential for these processes [7].

In *C. elegans* a similar mechanism for inhibiting pre-*let-7* processing has been reported [44]. The poly(U) polymerase PUP-2 was shown to oligo-uridylate pre-*let-7* in a Lin28-dependent fashion, thereby suppressing premature expression of mature *let-7* during larval development. In addition, subsequent RNA and chromatin immunoprecipitation assays revealed a specific interaction between Lin28 and pri-*let-7* that co-transcriptionally inhibits pri-*let-7* processing by Drosha [28]. An interaction between Lin28 and endogenous pri-*let-7* was also described for human ESCs and neuronal stem/progenitor cells [28]. Here, a highly expressed RNA-binding protein called Musashi1 (Msi1) selectively recruits Lin28a to the nucleus and synergistically blocks the cropping step of pri-*let-7* [45]. Moreover, it was suggested that Lin28b predominantly localizes to the nucleolus where it sequesters pri-*let-7*, thereby preventing Drosha processing in the nucleus [17]. Thus, Lin28a/b seem to obviate precocious expression of mature *let-7* during early development and differentiation by interfering with both the Drosha and Dicer complexes and by targeting pre-*let-7* towards degradation. Conversely,

upon differentiation of stem or progenitor cells, *let-7* ensures constant down-regulation of Lin28 by binding to the 3' UTR of Lin28 and its promoting transcription factor c-Myc [20] (Figure 2).

Figure 2. Lin28/*let-7* regulatory axis. In undifferentiated cells, Lin28 is highly expressed and blocks the biogenesis of *let-7* miRNA. By binding to the pre-element of pri- or pre-*let-7*, neither Drosha nor Dicer can process the corresponding *let-7* precursor. In addition, Lin28 recruits TUT4/TUT7 to pre-*let-7* and promotes its 3'-end oligo-uridylation. Oligo-uridylated pre-*let-7* cannot be cleaved by Dicer and thus serves as a signal for the cellular 3'–5' exoribonuclease Dis3l2. Upon differentiation, Lin28 expression is reduced, which leads to increased levels of mature *let-7*. The latter silences gene expression of proto-oncogenes (Ras, c-Myc, Hmga2), cell cycle progression factors (Cyclin D1 and D3, Cdk4), components of the insulin-PI3K-mTOR pathway and Lin28 itself, thereby establishing a positive feedback loop. Besides its role in differentiation, a Lin28/*let-7* regulatory network is apparently involved in several cellular processes such as proliferation, oncogenesis, development and physiology, as well as metabolism (recently reviewed in [42]).



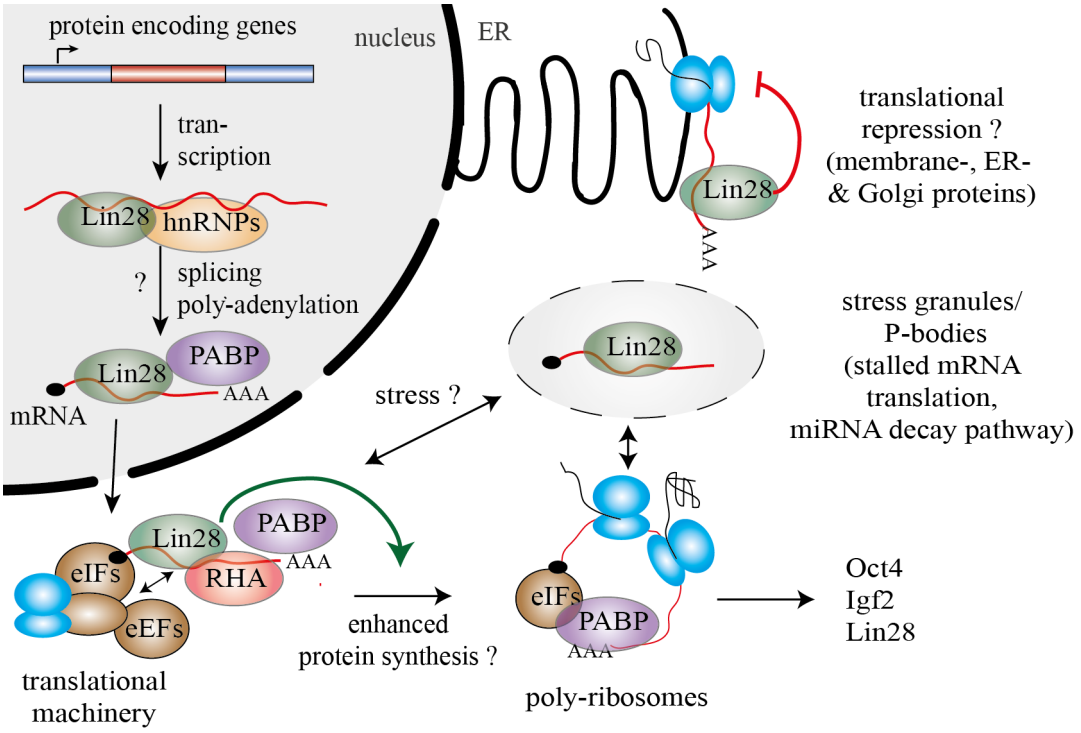
3. Lin28 Influences mRNA Translation

Besides regulation of *let-7* biogenesis, Lin28a/b can interact with various mRNAs and modulate their translation. Polesskaya and colleagues revealed that Lin28a can associate with polysomes and enhance translation of a number of mRNAs in differentiating myoblasts [25]. Among the first identified mRNA target was Igf2 (insulin-like growth factor 2), a major growth and differentiation factor in muscle tissue. Further evidence was provided that Lin28 recruits Igf2 mRNA to polysomes and enhances its translation via interactions with components of the translation initiation machinery. Subsequent studies revealed a number of additional mRNA targets of Lin28a in mouse ESCs such as H2a (histone 2a), Hmga1, Cyclin A, Cyclin B, Cdk4 and Oct4 [11,46–50]. An association of Lin28a with most of these mRNAs correlated with enhanced translation, suggesting that Lin28a maintains pluripotency by stimulating the translation of corresponding cell-cycle effectors. Further genome-wide studies revealed that Lin28a facilitates translation of genes important for growth and survival in human ESCs by recruiting RNA helicase A (RHA) to polysomes [24,51,52]. Additional mutagenesis studies revealed that the C-terminal part of Lin28a is required for RHA interactions, while mutations in the ZKD only impaired the stimulatory impact on translation, but not protein-protein interactions [46].

Very recently, a number of genome-wide Lin28 RNA crosslinking and immunoprecipitation coupled to high-throughput sequencing (HITS-Clip and PAR-CLIP) studies were conducted in human and mouse ESCs as well as somatic cells [19,49,53,54]. All of these studies have in common that only a small fraction of the identified RNA targets could be traced back to miRNAs, while the majority was mapped to thousands of mRNAs and ribosomal RNAs. For example, in mouse ESCs Lin28a was predominantly bound to mRNA transcripts (42%), mainly within the CDS and 3' UTR. Furthermore, a gene ontology analysis of target RNAs, revealed a preferential interaction of Lin28a with mRNAs that are destined for the endoplasmic reticulum. Binding of Lin28a to these mRNAs was associated with a translational repression by reducing ribosome occupancy without affecting mRNA abundance [49].

On contrary, in human HEK293 cells, binding of Lin28a and Lin28b to its mRNA targets was linked to a globally enhanced protein synthesis [19,53]. As before in mESCs, both human Lin28 paralogs predominantly bound within exonic regions of mRNAs, thereby mirroring the predominant cytosolic localization of Lin28a/b in HEK293 cells. Among the top RNA targets were mRNAs encoding for splicing factors and RNA-binding proteins, cell-cycle regulators as well as Lin28 itself. Binding of Lin28b to its own mRNA, indeed, correlated with increased levels of Lin28b protein, thereby suggesting a *let-7* independent feed-forward mechanism to maintain high levels of Lin28b in proliferative cell types [19,53,54]. Apart from their own expression Lin28a/b also seem to drive expression of important cell-cycle regulators of the ERK signaling cascade, such as Cdk1, N-Ras, Ran and ERK. This would explain the strong proliferative defects observed upon Lin28b knockdown [53]. Wilbert and colleagues further detected widespread changes in protein levels of splicing factors upon down-regulation of Lin28a and Lin28b in human ESCs. Whereas Lin28a binding to hnRNP F mRNA repressed translation, binding to TDP-43 and FUS/TLS mRNA was associated with an enhanced protein synthesis of the corresponding transcript. Consistent with Lin28's impact on alternative splicing factors, up-regulation of Lin28a in somatic HEK293 cells caused dramatic changes in alternative splicing patterns [54] (Figure 3).

Figure 3. Lin28 binds various mRNAs and modulates their translation. Both Lin28 paralogs were shown to influence mRNA processing on several levels. In the nucleus, Lin28 could regulate splicing of bound pre-mRNAs in concert with heterogeneous nuclear ribonucleoproteins (hnRNPs). In the cytosol, Lin28 was shown to interact with an RNA helicase A (RHA) thereby modulating the translation of target mRNAs via interactions with eukaryotic translation initiation factors (eIFS), elongation factors (eEFS) and poly(A)-binding proteins (PABP). Furthermore, Lin28 was found to shuttle mRNAs to poly-ribosomes and, under stress condition, to P-bodies and stress granules, thereby providing a direct link to the miRNA decay machinery. Lin28 binding to mRNAs was typically associated with a globally enhanced protein synthesis. However, in hESCs Lin28 binding repressed translation of bound mRNAs that were destined for the ER.



4. Functional Importance of Lin28-Mediated mRNA and miRNA Regulation for Stem Cell Maintenance, Cancer and Development

The functional importance of Lin28 in stem cell maintenance and reconstituting pluripotency becomes apparent when looking at the signaling pathways in which Lin28a/b are involved. Both paralogs are highly expressed in mammalian ESCs and are a central part of a conserved pluripotency network. For example, the expression of Lin28a is driven by the proto-oncogenic transcription factors Oct4, Sox2 and Nanog with Sox2 being most critical for an efficient Lin28a expression [55,56]. Once Lin28a is expressed, it antagonizes *let-7* and hence de-represses *let-7* targets such as c-Myc, Sal4, Igf2bps, Hmga2, various cyclins as well as Lin28 itself, thereby ensuring a constant expression of stemness factors and cell cycle regulator [57]. In addition, Lin28a directly or indirectly stimulates the translation of mRNAs encoding for cell-cycle regulators or growth-promoting factors such as Cyclin A/B, Oct4 and Igf2 [18,25,48,50]. Consequently, Lin28a/b up-regulate the expression of cell-cycle

regulators and growth-promoting factors via *let-7* dependent and *let-7* independent mechanisms, thereby activating and maintaining signaling pathways that are important for self-renewal and proliferation. In agreement with this, Lin28a overexpression is not essential for reprogramming human fibroblast to iPSCs but strongly accelerates reprogramming by stimulating cell proliferation [58].

The strong effect of Lin28a/b on cell progression and proliferation [59] and the frequent re-activation of Lin28a/b in multiple cancers [12] supported the role of Lin28 as a potential oncogene. Indeed, Lin28a/b overexpressing in NIH/3T3 cells led to tumor formation in nude mice and was linked to depletion of mature *let-7.* As a consequence, oncogenic *let-7* targets such as c-Myc and N-Ras were de-repressed, and, since c-Myc itself transcriptionally activates various oncogenic miRNAs as well as Lin28b, a positive feed-forward loop is established [12,60]. Iliopoulus and colleagues revealed another positive feedback loop between NF-κB, Lin28b, *let-7* and IL-6 (Interleukin 6). Transient activation of Src tyrosine kinase in immortalized breast cells led to activation of NF-κB, which binds to the Lin28b promoter and induces its expression. As a result, Lin28b represses *let-7* processing, the *let-7* target IL-6 can be produced and activate NF-κB, thereby closing the positive feedback loop [61]. Similar to its role in reprogramming somatic cells to iPSCs, an elevated expression of Lin28a/b might also be important in the formation of cancer stem cells (CSCs) [62]. This subpopulation of tumor cells is thought to be essential for the propagation of some cancer cells and might arise in a reprogramming-like mechanism [63]. Hence, Lin28a/b reactivation would contribute to the formation of metastasis thereby explaining why Lin28a/b up-regulation correlates with tumor aggressiveness and an advanced tumor stage [12,62].

Given that *let-7* family members target numerous metabolic genes, it is not surprising that Lin28a/b overexpression also has an impact on growth, developmental timing and metabolism. Using genome-wide association studies, genetic variations within the LIN28B loci was linked to changes of human height, timing of puberty and the age of menopause [64–66]. Consistent with these studies, Lin28a overexpression in transgenic mice led to similar phenotypes and was associated with increased insulin sensitivity and increased glucose uptake [14]. On the molecular level, Lin28a/b act on multiple components of the insulin-P13K-mTOR pathway, thereby explaining why administration of the mTOR inhibitor rapamycin could rescue the Lin28a-mediated metabolic phenotype [15]. Further *in vitro* studies showed that Lin28a de-represses *let-7* targets of the insulin-P13K-mTOR pathway such as Igf1r, Insr, Irs2, Akt2, Tsc1 and Rictor [15,67]. The authors could not rule out that Lin28a/b associate with these mRNAs itself and enhance their translation. Recent genome-wide Clip-seq studies indeed suggested that Lin28a/b binds to mRNAs of insulin and Igf receptors, glycolytic and mitochondrial enzymes thereby modulating their translation directly [19,52,53]. Hence, Lin28a/b seem to regulate both mRNA translation and *let-7* maturation to coordinate proliferative signaling pathways and cellular metabolism in order to maintain the self-renewal potential of stem or progenitor cells. However, given the wealth of recently identified mRNA targets of Lin28a/b, their overlap with known *let-7* targets and the interwoven signaling pathways, it remains to be determined which of the identified targets indeed contribute to the observed physiological functions.

5. Structural Basis for the RNA-Binding Specificity of Lin28

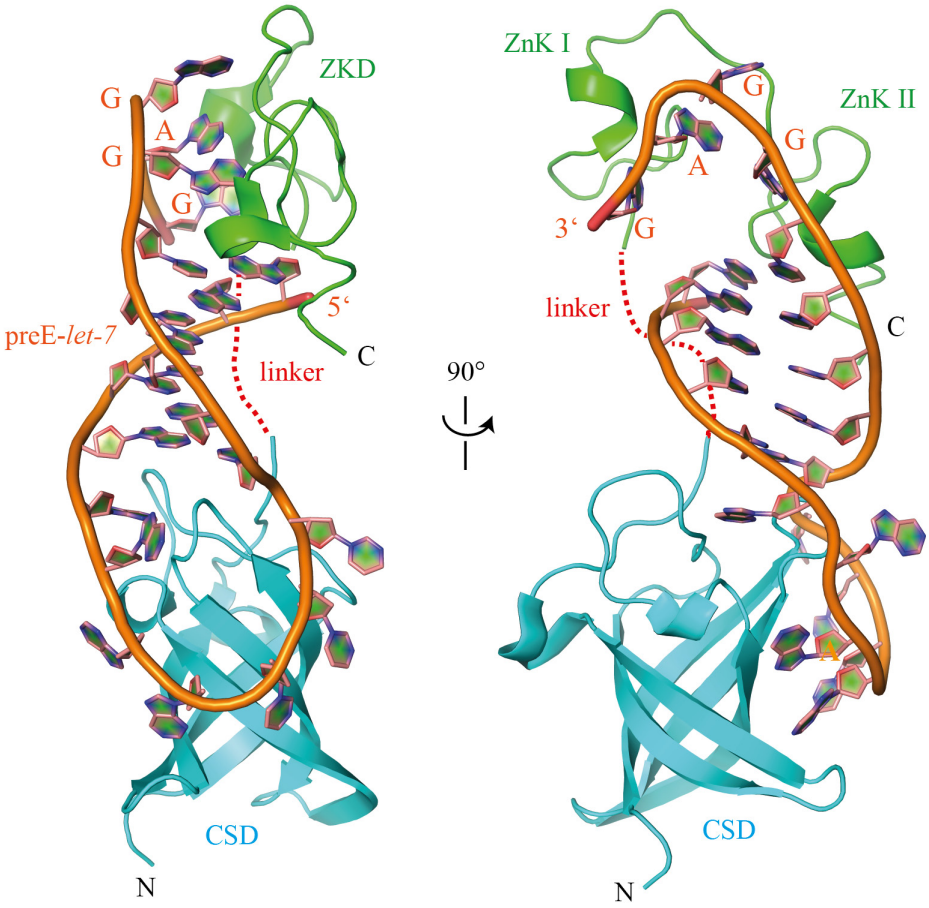
5.1. The Lin28 Zinc-Knuckle Domain Specifically Recognizes GGAG or GGAG-Like Motifs

After identifying *let-7* precursors as major targets of Lin28a and Lin28b, several groups aimed to identify the specificity of this interaction. Using electrophoretic mobility shift assays with different pre-*let-7* sequences, it became initially apparent that the terminal loop of pre-*let-7* (also called pre-element or preE) is sufficient for Lin28a binding [33]. An alignment of stem-loop precursors of *let-7* revealed a highly conserved GGAG motif within vertebrates that is critical for Lin28 binding. Mutations within this motif (GGAG→AAAG and GGAG→GUAU) released the Lin28a-mediated block of pri- or pre-*let-7* processing and impaired TUT4-mediated oligo-uridylation of pre-*let-7* [7,22]. On contrary, introduction of the GGAG motif into preE of an unrelated miRNA (pre-*miR-16-1*) allowed Lin28a binding and TUT4-mediated uridylation of this chimeric pre-miRNA [7].

Due to the close homology between Lin28's ZKD and the ZKD of HIV-1 nucleocapsid protein (HIV NC), which was known to bind GGAG- or GGUG-containing loops within the HIV Ψ-RNA recognition element [68–70], it was suggested that Lin28's ZKD mediates a specific interaction with the conserved GGAG motif (see Figure 5A). Indeed, mutations with Lin28's ZKD specifically impaired pre-*let-7* binding as well as binding of the isolated Lin28 ZKD to GGAG-containing RNAs [7,34,35,71,72]. Co-crystal structures of a minimal mouse Lin28a construct with GGAG-containing oligonucleotides derived from preE-*let-7* (Figures 4 and 5) and a NMR solution structure of human Lin28a's ZKD bound to AGGAGAU provided the final proof for the supposed interaction (Figure 5B,C) [35,72].

For Lin28:mRNA binding, so far no structural data has been obtained. However, despite their discrepancies in individual mRNA targets, most of the above mentioned genome-wide HITS-CLIP and PAR-CLIP studies identified GGAG or GGAG-like consensus motifs within Lin28a/b binding sites. For example, Wilbert *et al.* found a highly enriched GGAGA(U) consensus sequence that was enriched within loop structures [54]. Cho *et al.* detected AAGNNG, AAGNG and UGUG motifs that are often located in terminal loops of small RNA hairpins [49]. Finally, Graf *et al.* detected GGSWG (S = G or C, W = A or T) or AAGRWG (R = A or G) motifs in Lin28b binding sites. Using individual domain PAR-CLIP (iDo-PAR-CLIP) Graf and colleagues further confirmed the GGGAG sequence as the top motif within Lin28 ZKD binding sites, whereas Lin28 CSD binding sites were rather U-rich [53]. These data indicate that the GGAG motif is indeed the major determinant of Lin28 RNA binding and is recognized by the ZKD. Even a mutation of the first or second guanosine only moderately impairs the interaction, thereby mirroring the overall flexibility of both ZKD and RNA. A recent study revealed that CCHC Zn knuckles can be used to design single-stranded nucleic-acid binding proteins that specifically recognize a number of guanosines [73]. This study further demonstrated that the length of the inter-knuckle linker affects spacing between specifically bound guanosines. Hence, Lin28 ZKD probably prefers GNNG motifs over NGNG motifs as seen for HIV-1 NC ZKD (see Figure 5C). Interestingly, TUT4 and TUT7 also contain CCHC Zn knuckles that are critical for pre-*let-7* oligo-uridylation [42,74]. Compared to Lin28, the distance between these knuckles is larger (37 aa), indicating that they act independently from each other.

Figure 4. Co-crystal structure of a minimal mouse Lin28a construct with preE-*let*-7d derived RNA (PDB ID 3TRZ). The ZKD specifically binds to the conserved GGAG motif, whereas Lin28 CSD establishes extensive interactions with the less conserved terminal hairpin loop.



The second CCHC Zn knuckle undergoes a larger structural change upon RNA binding, whereby the central Zn^{2+} ion moves about 25 Å. Responsible for this large conformational shift is Pro158 within the Pro-rich linker region, since its ψ torsion angle performs a 130° rotation (Figure 5B) [72]. While HIV-1 NC ZKD specifically binds G-2 and G-4 of a GGAG tetraloop in a sequence-specific manner, each CCHC Zn knuckle of Lin28 specifically recognizes the first and fourth guanosine of the GGAG motif by sequence-specific hydrogen bonds to the bases. Hydrogen bonding is mediated by backbone carbonyl and amide groups of residues that are located within the rigid parts of the CCHC Zn knuckle. In addition to this sequence-specific interaction, both G-1 and G-4 are sandwiched in a hydrophobic pocket by one conserved Tyr and His in the first Zn knuckle and another conserved His and Met in the second Zn knuckle (Figure 5C). In the case of mLin28a:GGAG structures, G-2 is also bound in a sequence-specific manner via hydrogen bonds from backbone carbonyl groups and the N1 amino group of A-3. Even more, A-3 contributes to the formation of a strong kink within the RNA backbone, since it also contacts G-1 [35] (Figure 5D). Although such a strong bending of the RNA backbone was not observed in the hLin28a:AGGAGAU structure, the imposed structural changes within RNA and protein likely lead to a constant opening of neighboring double-stranded pre-*let*-7 stem thereby masking the Dicer cleavage site [34,35].

Figure 5. Lin28 ZKD specifically recognizes single-stranded GGAG or GGAG-like sequences. (**A**) Sequence alignment of HIV-1 NC, HIV-2 NC, hLin28a and hLin28b ZKDs. The chelating Cys and His residues of the CCHC Zn knuckles (ZnK) are shaded in red. Conserved residues are labeled from light red (100% type-conserved) to dark red (70% type-conserved); (**B**) Comparison between unbound hLin28a ZKD (green, PDB-ID 2CQF) and AGGAGAU-bound hLin28a ZKD (purple, PDB-ID 2LI8). Upon RNA binding, hLin28a ZKD undergoes a dramatic conformational shift mainly caused by a rotation of the Pro158 ψ angle; (**C**) In comparison to HIV-1 NC, the inter-knuckle linker of hLin28a ZKD harbors an additional Pro. As a consequence, the knuckles are further apart, thereby explaining why HIV-1 NC ZKD specifically binds G-2 and G-4 while hLin28a ZKD binds G-1 and G-4 of the GGAG motif in a hydrophobic pocket; (**D**) Structure of mLin28a ZKD bound to GGAG (derived from PDB-ID 3TSO). mLin28a is represented in green cartoon and the bound GGAG motif in purple (G) and pink (A). Tyr140 of the first and His162 of the second ZnK are key residues for the interaction, since they contact each other and stack with the bases, thereby establishing a kinked conformation in the RNA. All three guanosines are specifically recognized via various hydrogen bonds with backbone amide and carbonyl groups. In addition, G-1 and G-4 are bound in a hydrophobic pocket formed by His140, His162, Tyr140 and Met170.

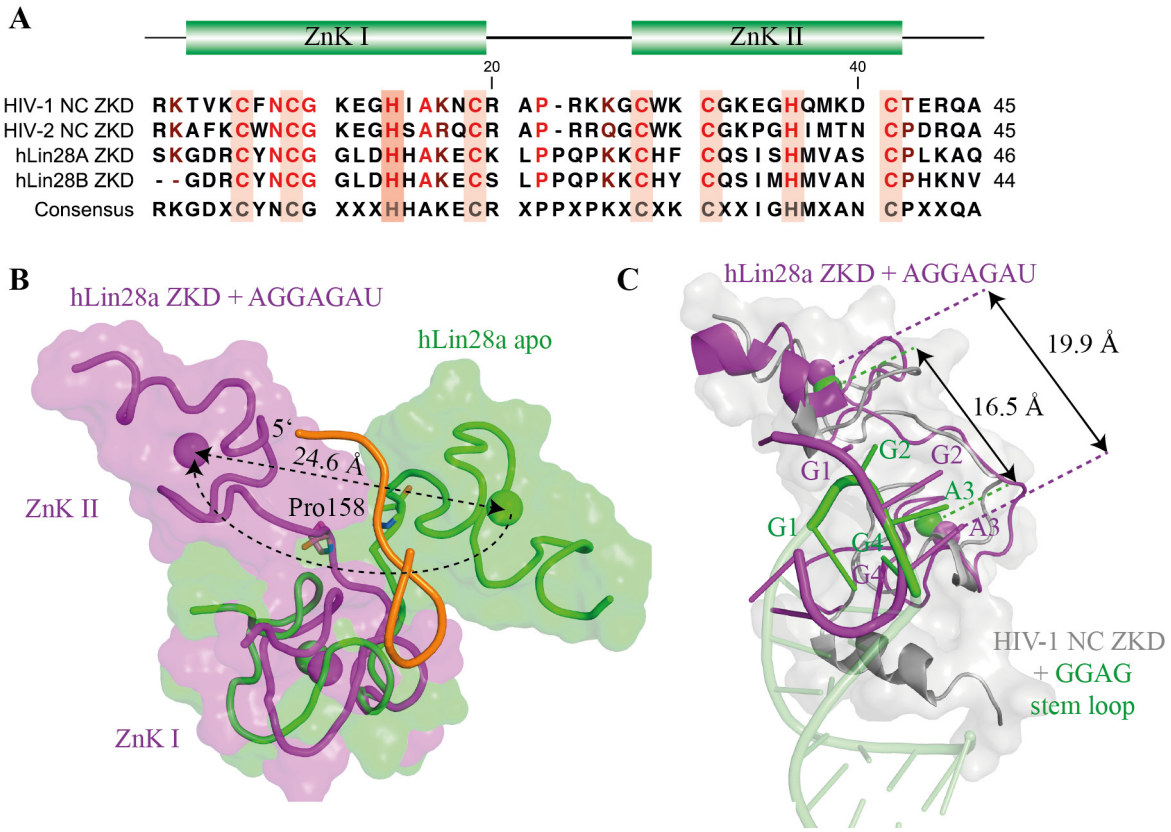
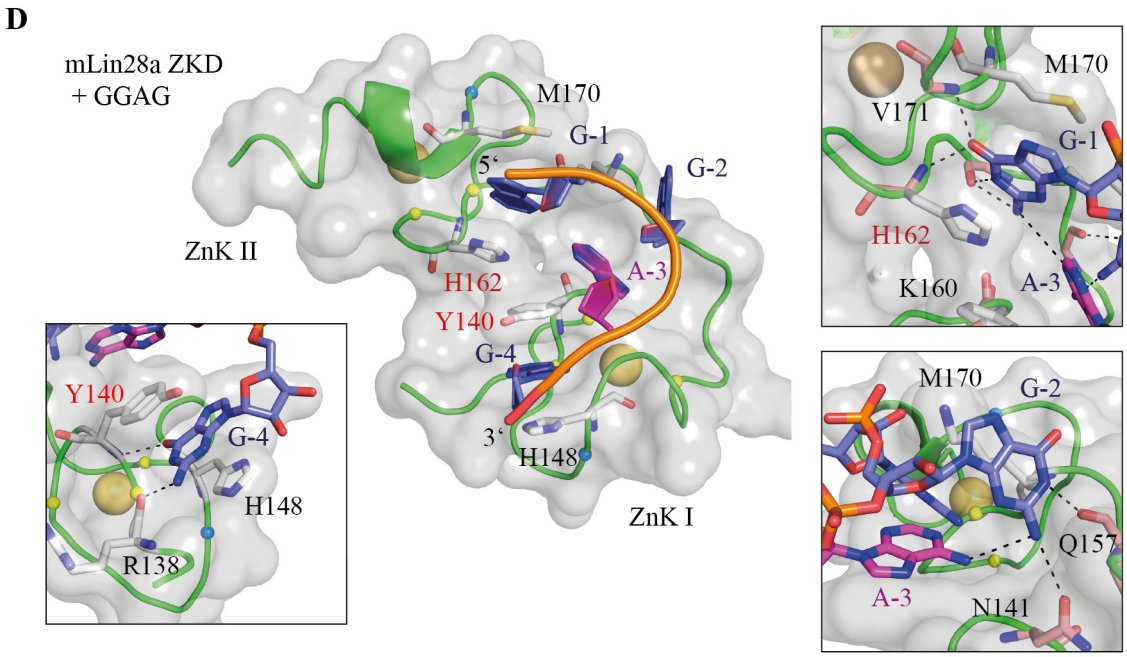


Figure 5. *Cont.*



5.2. The Lin28 CSD Has Broad Sequence Specificity and Can Induce Local Structural Changes within RNAs

Despite Lin28's specificity for GGAG-containing RNAs, the isolated Lin28 ZKD is not sufficient for binding *let-7* precursors and blocking their processing [34,35]. So what is the contribution of Lin28's CSD with respect to sequence specificity, binding affinity and inhibition of pre-*let-7* processing?

CSDs are highly conserved RBDs that are widely distributed in bacteria, animals and plants and fulfill pleiotropic functions mainly related to RNA metabolism (reviewed in [75]). Bacterial major cold-shock proteins (Csps) share between 30% and 45% sequence identity to Lin28 CSDs and are known to bind pyrimidine-rich ssDNA/ssRNA oligonucleotides with affinities in the sub-nanomolar to micromolar range [76–81]. In addition to this, they can act as RNA chaperones that destabilize local RNA secondary structures [82–84]. Crystal and NMR structures of Csps have been known since the 1990s [85–89].

A systematic binding analysis with *Xenopus tropicalis* (*Xtr*) Lin28b CSD revealed that this domain has a broad sequence specificity and shows the highest binding affinities for pyrimidine-rich RNA octamers that contain at least one guanosine at the 5' end [34]. The observation was further confirmed by genome-wide PAR-CLIP studies, in which Lin28a/b binding sites were generally uridine-rich and flanked by one or more guanosines [19,53]. Moreover, these binding sites were typically located upstream of the corresponding ZKD binding sites, indicating a defined domain orientation of Lin28s' RBDs on RNA targets [53].

Co-crystal structures of Lin28 CSDs in complex with ssDNA and preE-*let-7* derived RNA stem loops provided valuable information about Lin28's specificity and function in pre-*let-7* and mRNA binding [34,35]. Lin28 CSDs bind to single-stranded nucleic acids via a conserved nucleic acid-binding platform mainly formed of exposed aromatic residues. Unlike for Lin28 ZKD, this binding platform is already pre-formed in the apo protein and, consequently, only subtle changes are

observed upon nucleic-acid binding (Figure 6A). Binding of ssDNA and ssRNA are remarkably similar and dominated by π-stacking interactions with exposed aromatic residues (Figure 6B). Consistent with solution binding experiments and bacterial Csp:ssDNA/RNA structures [79,90], Lin28 CSD binds up to 8 nucleotides arranged in a curved single strand with defined orientation. In the case of the mLin28a:preE-*let-7* structures, an additional ninth nucleotide is visible that establishes hydrogen bonds with the first base, thereby closing the preE stem loop. Sequence-specific binding is mainly mediated at position 6, since the presence of a conserved Lys-Asp salt bridge limits the flexibility and, consequently, the size of the binding pocket and contributes to specific hydrogen bonds with the U/T base. In addition to this, at binding subsite 2 either a T, U or G is specifically recognized within a hydrophobic pocket. Despite the difference in size, the corresponding bases are recognized by similar hydrogen bonds. The lack of contacts with the CSD allows the DNA/RNA backbone to adopt slightly different conformations without disturbing hydrogen bonding.

Besides its contribution to binding affinity and specificity, Lin28 CSD can affect and reorganize secondary structures within RNA targets. The first hint in this direction came from a study that examined the effect of Lin28a binding on pre-*let-7g* secondary structure using enzymatic foot-printing [36]. Upon Lin28a binding, some regions of preE as well as a part of the double-stranded stem of pre-*let-7* became more susceptible to cleavage by single-strand-specific ribonucleases. Hence, the authors concluded that Lin28a is able to unwind the double-stranded stem of pre-*let-7*, thereby blocking the Dicer cleavage site. Second, Nam *et al.* provided evidence that Lin28a's CSD can partially melt double-stranded stem loops to generate an optimal binding interface [35]. Third, using site-directed mutagenesis in combination with a kinetic analysis of *Xtr*Lin28b-mediated remodeling of pre-*let-7g*, it was shown that Lin28's CSD first binds to pre-*let-7* and induces a structural change [34]. Consistent with earlier studies on bacterial Csps [83,91], highly conserved His (His68 in *Xtr*Lin28b-binding subsite 4) and Phe residues (Phe77-binding subsite 1,2) were crucial for the remodeling reaction. The CSD-induced remodeling might be important for proper recognition of the GGAG motif by Lin28's ZKD, since in most pre-*let-7* structures the conserved GGAG motif is involved in secondary structures and therefore not accessible for binding (Figure 7). Genome-wide PAR-CLIP and HITS-CLIP studies further supported this hypothesis, since Lin28a/b could recognize RNA binding sites that are predicted to be involved in stable secondary structures [19,49]. Such a chaperone-like function of Lin28 might be an important regulatory mechanism that allows downstream RBPs either to dissociate from or associate with RNPs and influence their processing. Most notably, a recent study provided evidence that Dis3l2 exoribonuclease degrades oligo-uridylated pre-*let-7*. This enzyme is composed of one ribonuclease II domain, two CSDs and one CSD-like S1 domain. Interestingly, both the CSDs and the S1 domain were essential for Dis3l2 binding and degradation of oligo-uridylated pre-*let-7*. Given the preference of Lin28 CSD's for U-rich binding sites, this suggests that the CSD might recognize the oligo(U) tail. In addition, it may assist in the exoribonucleolytic degradation of oligo(U)-pre-*let-7* by partially unwinding the double-stranded miRNA stem.

Figure 6. The Lin28 CSD can bind to a wide range of different RNA sequences. (**A**) Superimposition of unbound (skin color, PDB-ID 3ULJ) and heptathymidine-bound *Xtr*Lin28b CSD (green, PDB-ID 4A76). Both structures are highly conserved and reveal a pre-formed nucleic-acid binding platform with exposed aromatic residues; (**B**) Superimposition of *Xtr*Lin28b:dT_7 (green) and mLin28s:preE-*let-7f* (RNA: blue, protein: gray, PDB-ID 3TS0). Both Lin28 CSDs bind single-stranded nucleic acids predominantly via base stacking interactions in a defined orientation. The protein nucleic-acid interaction surface is similar for binding subsites 1 to 7. Binding of an additional eighth (U-8) and ninth (U-9) base in mLin28:preE-*let-7f* is triggered by the formation of a closed RNA loop; (**C**) Superimposition of bound nucleotides at binding subsite 6 derived from various bacterial and Lin28 CSDs in complex with ssDNA/ssRNA (PDB-IDs 4A76, 4A75, 3TS0, 3TS2, 3PF4, 2HAX). All structures contained T or U nucleotides at this binding pocket. A highly conserved Lys-Asp salt bridge limits the size of the pocket and establishes specific hydrogen bonds with the T/U base; (**D**) Since few interactions are formed with the sugar-phosphate backbone, the bound oligonucleotides can adopt different backbone conformations to optimize binding with Lin28 CSD. For example, at binding subsite 2, the sugar-phosphate backbone of mLin28a:preE-*let-7f* is farther displaced from the protein, thereby enabling binding of G (G-2) instead of T (T-2) without disrupting hydrogen bonds.

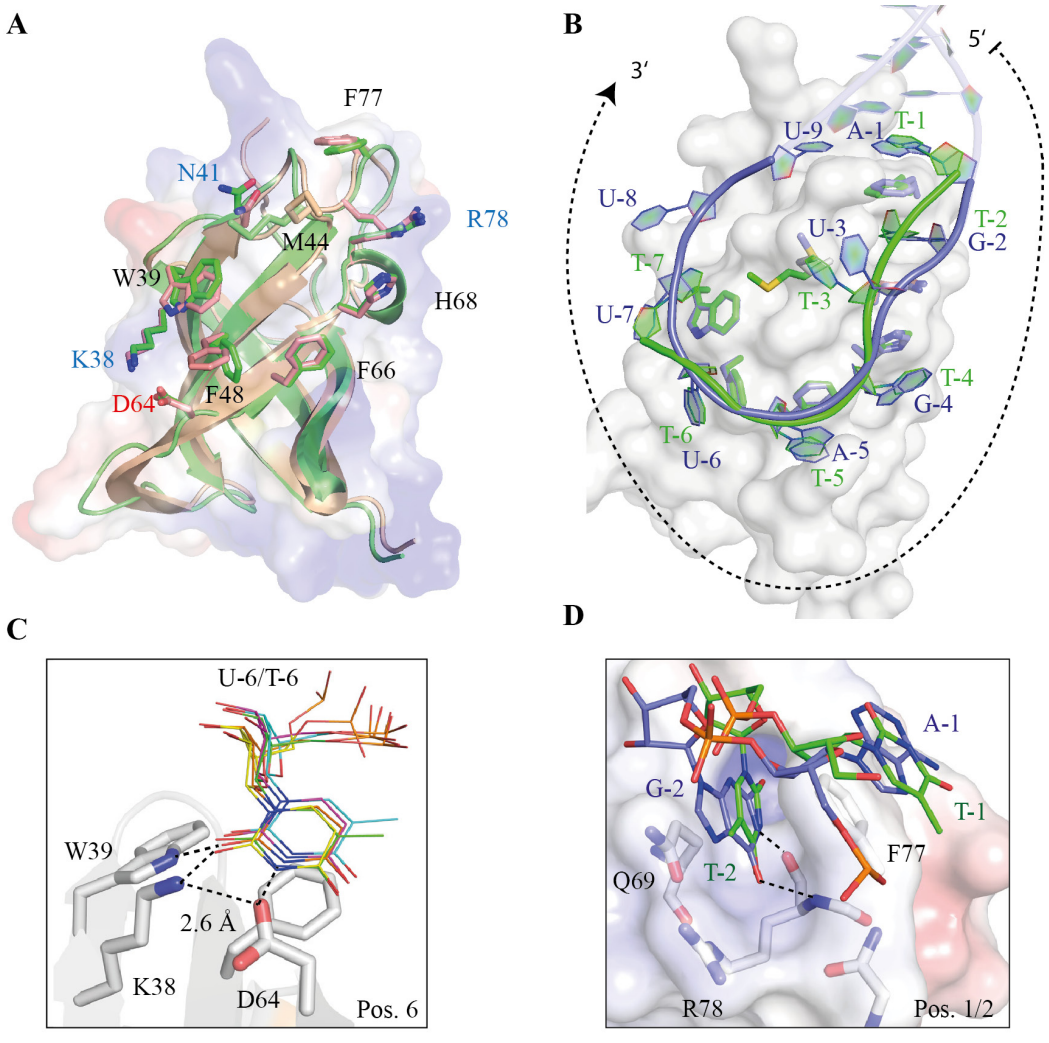
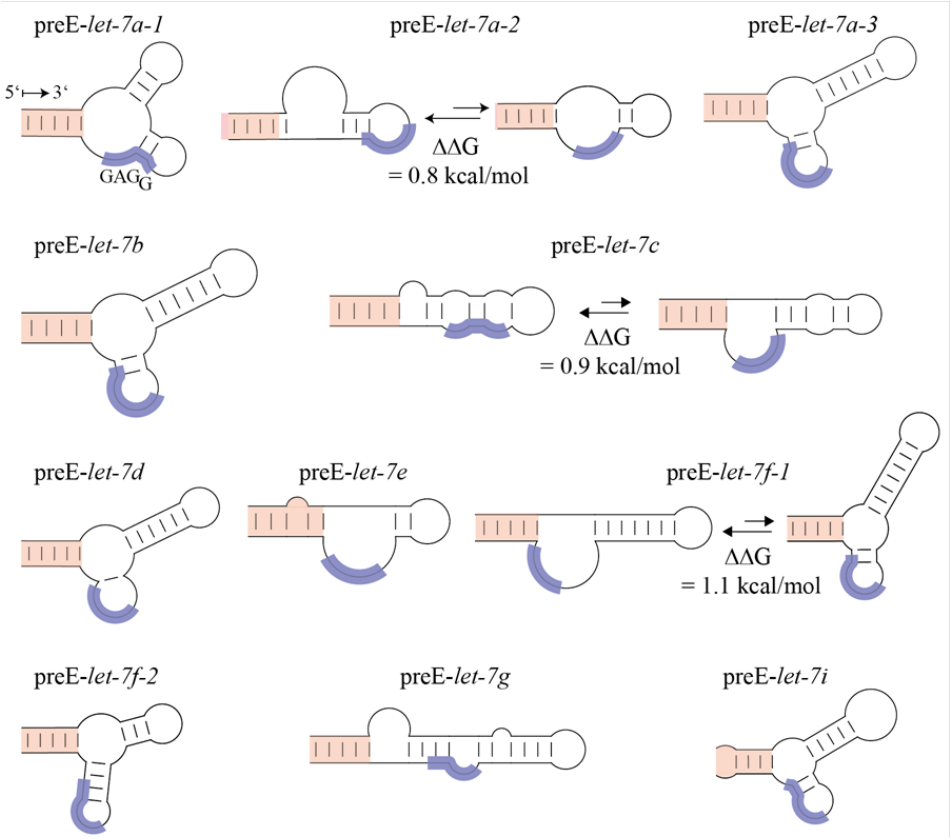


Figure 7. The pre-elements of *let-7* family members are structurally diverse. In six out of eleven human *let-7* family members, the conserved GGAG motif (blue) is inaccessible for ZKD binding in the lowest-energy folding state. Secondary structure predictions of human *let-7* family members (except miR-98 and miR-202) were calculated and visualized by CLC genomics workbench 3.65. All lowest-energy structures within a ΔΔG range of 1.5 kcal/mol are depicted. For simplicity, only 5 bp of the miRNA stem are shown (labeled in red).



6. Summary and Conclusions

Recent structural and biochemical studies along with genome-wide CLIP-seq studies have greatly improved our understanding of how Lin28 recognizes target RNAs and fulfills its pleiotropic functions related to regulation of miRNA and mRNA processing as well as mRNA translation. In the case of *let-7* biogenesis, Lin28 ZKD specifically recognizes a conserved GGAG motif within preE-*let-7* and induces a strong bending of the bound RNA backbone. Thus, the adjacent Dicer cleavage site remains constantly unwound and pre-*let-7* cannot be processed anymore. Apart from a minor preference for pyrimidine-rich sequences with one flanking guanosine, the CSD did not reveal any clear sequence specificity, but was able to remodel local RNA secondary structures. This might be important in three ways. First, initial binding of the CSD to single-stranded RNA sequences can induce a conformation in which the GGAG motif is accessible for subsequent ZKD binding. Second, the CSD might trigger structural changes within target RNAs thereby stimulating downstream processes such as pre-*let-7* oligo-uridylation. Third, the wide RNA-binding specificity of the CSD enables Lin28 to recognize all *let-7* family precursors in a defined 5'–3' orientation despite the low sequence conservation within *let-7* preEs. Consequently, Lin28 impairs *let-7* biogenesis and irreversibly targets pre-*let-7* to degradation.

On the mRNA level the combination of both domains enables Lin28a/b to bind thousands of mRNAs. Although there is still no consensus how Lin28 mRNA binding influences mRNA processing and regulates translation, the observed principles with respect to sequence specificity and RNA remodeling are also valid here. Lin28a/b recognize GGAG or GGAG-like motifs and can access these motifs even if they are embedded in predicted secondary structures. The CSD typically binds to uridine-rich regions upstream of the ZKD binding site and is responsible for Lin28's RNA chaperone-like function. The defined orientation of Lin28 on bound RNA might allow a specific recruitment of downstream factors, such as RNA helicase A. However, downstream effects of Lin28:mRNA binding were quite distinct in recent genome-wide CLIP-seq studies and comprised translational stimulation of growth-promoting and alternative splicing factors, as wells as translational repression of ER-destined mRNAs [19,49,52–54]. In addition to the localization of binding sites (coding sequence, 3'UTR), Lin28-induced structural changes within mRNAs and/or direct protein-protein interactions might affect the translation efficiency of target mRNAs.

Therefore, it will be essential to understand how Lin28a/b regulate mRNA processing and translation in more detail. Ongoing studies will need to determine which cellular factors are involved in these processes and which factors regulate the activity of Lin28a/b within different cellular compartments and cells/tissues. Furthermore, it remains to be verified which of the huge number of novel mRNA targets are indeed directly regulated by Lin28a/b and what are their impact on stem cell maintenance/differentiation, development, metabolism and cancer. Last but not least, additional structural and functional data of how Lin28 binds mRNA targets and interacts with downstream components such as RNA helicase A may help to elucidate the mechanisms behind Lin28-mediated translational enhancement. Although much effort has been undertaken to unravel the molecular mechanisms that control the Lin28/*let-7* regulatory axis, there are still a number of issues that remain to be solved. Which precise mechanisms does Lin28 use to inhibit pri-*let-7* processing by the Microprocessor in the nucleus? How does Lin28 stimulate TUT4/TUT7-mediated oligo-uridylation of pre-*let-7*? And is the observed RNA-chaperone-like function of Lin28 mandatory for fulfilling its tasks? Understanding these issues might help us to exploit Lin28's function and manipulate the involved pathways for improved tissue re-engineering and novel treatments of cancer or metabolic diseases.

Acknowledgments

The authors express their gratitude to Yvette Roske and Anja Schütz (Max-Delbrück-Center, Berlin) for their help with some of the structural studies covered in this review and helpful discussions.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Ambros, V.; Horvitz, H. Heterochronic mutants of the nematode *Caenorhabditis elegans*. *Science* **1984**, *226*, 409–416.
2. Moss, E.G.; Lee, R.C.; Ambros, V. The cold shock domain protein LIN-28 controls developmental timing in *C. elegans* and is regulated by the lin-4 RNA. *Cell* **1997**, *88*, 637–646.

3. Seggerson, K.; Tang, L.; Moss, E.G. Two genetic circuits repress the *Caenorhabditis elegans* heterochronic gene lin-28 after translation initiation. *Dev. Biol.* **2002**, *243*, 215–225.
4. Moss, E.G.; Tang, L. Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. *Dev. Biol.* **2003**, *258*, 432–442.
5. Darr, H.; Benvenisty, N. Genetic analysis of the role of the reprogramming gene LIN-28 in human embryonic stem cells. *Stem Cells* **2009**, *27*, 352–362.
6. Yu, J.; Vodyanik, M.; Smuga-Otto, K.; Antosiewicz-Bourget, J.; Frane, J.; Tian, S.; Nie, J.; Jonsdottir, G.; Ruotti, V.; Stewart, R.; *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **2007**, *318*, 1917–1920.
7. Heo, I.; Joo, C.; Kim, Y.-K.; Ha, M.; Yoon, M.-J.; Cho, J.; Yeom, K.-H.; Han, J.; Kim, V.N. TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell* **2009**, *138*, 696–708.
8. Guo, Y.; Chen, Y.; Ito, H.; Watanabe, A.; Ge, X.; Kodama, T.; Aburatani, H. Identification and characterization of lin-28 homolog B (LIN28B) in human hepatocellular carcinoma. *Gene* **2006**, *384*, 51–61.
9. King, C.E.; Cuatrecasas, M.; Castells, A.; Sepulveda, A.R.; Lee, J.-S.; Rustgi, A.K. LIN28B promotes colon cancer progression and metastasis. *Cancer Res.* **2011**, *71*, 4260–4268.
10. King, C.E.; Wang, L.; Winograd, R.; Madison, B.B.; Mongroo, P.S.; Johnstone, C.N.; Rustgi, A.K. LIN28B fosters colon cancer migration, invasion and transformation through let-7-dependent and -independent mechanisms. *Oncogene* **2011**, *30*, 4185–4193.
11. Peng, S.; Maihle, N.J.; Huang, Y. Pluripotency factors Lin28 and Oct4 identify a sub-population of stem cell-like cells in ovarian cancer. *Oncogene* **2010**, *29*, 2153–2159.
12. Viswanathan, S.; Powers, J.; Einhorn, W.; Hoshida, Y.; Ng, T.; Toffanin, S.; O'Sullivan, M.; Lu, J.; Phillips, L.; Lockhart, V.; *et al.* Lin28 promotes transformation and is associated with advanced human malignancies. *Nat. Genet* **2009**, *41*, 843–848.
13. Wang, Y.-C.; Chen, Y.-L.; Yuan, R.-H.; Pan, H.-W.; Yang, W.-C.; Hsu, H.-C.; Jeng, Y.-M. Lin-28B expression promotes transformation and invasion in human hepatocellular carcinoma. *Carcinogenesis* **2010**, *31*, 1516–1522.
14. Zhu, H.; Shah, S.; Shyh-Chang, N.; Shinoda, G.; Einhorn, W.; Viswanathan, S.; Takeuchi, A.; Grasemann, C.; Rinn, J.; Lopez, M.; *et al.* Lin28a transgenic mice manifest size and puberty phenotypes identified in human genetic association studies. *Nat. Genet* **2010**, *42*, 626–630.
15. Zhu, H.; Shyh-Chang, N.; Segre, A.V.; Shinoda, G.; Shah, S.P.; Einhorn, W.S.; Takeuchi, A.; Engreitz, J.M.; Hagan, J.P.; Kharas, M.G.; *et al.* The Lin28/let-7 axis regulates glucose metabolism. *Cell* **2011**, *147*, 81–94.
16. Jaroszewski, L.; Li, Z.; Cai, X.H.; Weber, C.; Godzik, A. FFAS server: Novel features and applications. *Nucleic Acids Res.* **2011**, *39*, W38–W44.
17. Piskounova, E.; Polytarchou, C.; Thornton, J.E.; LaPierre, R.J.; Pothoulakis, C.; Hagan, J.P.; Iliopoulos, D.; Gregory, R.I. Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. *Cell* **2011**, *147*, 1066–1079.
18. Balzer, E.; Moss, E.G. Localization of the developmental timing regulator Lin28 to mRNP complexes, P-bodies and stress granules. *RNA Biol.* **2007**, *4*, 16–25.

19. Hafner, M.; Max, K.E.; Bandaru, P.; Morozov, P.; Gerstberger, S.; Brown, M.; Molina, H.; Tuschl, T. Identification of mRNAs bound and regulated by human LIN28 proteins and molecular requirements for RNA recognition. *RNA* **2013**, *19*, 613–626.
20. Rybak, A.; Fuchs, H.; Smirnova, L.; Brandt, C.; Pohl, E.; Nitsch, R.; Wulczyn, F. A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nat. Cell Biol.* **2008**, *10*, 987–993.
21. Gaytan, F.; Sangiao-Alvarellos, S.; Manfredi-Lozano, M.; Garcia-Galiano, D.; Ruiz-Pino, F.; Romero-Ruiz, A.; Leon, S.; Morales, C.; Cordido, F.; Pinilla, L.; Tena-Sempere, M. Distinct expression patterns predict differential roles of the miRNA-binding proteins, Lin28 and Lin28b, in the mouse testis: Studies during postnatal development and in a model of hypogonadotropic hypogonadism. *Endocrinology* **2013**, *154*, 1321–1336.
22. Newman, M.A.; Thomson, J.M.; Hammond, S.M. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA* **2008**, *14*, 1539–1549.
23. Viswanathan, S.R.; Daley, G.Q.; Gregory, R.I. Selective blockade of microRNA processing by Lin28. *Science* **2008**, *320*, 97–100.
24. Jin, J.; Jing, W.; Lei, X.-X.; Feng, C.; Peng, S.; Boris-Lawrie, K.; Huang, Y. Evidence that Lin28 stimulates translation by recruiting RNA helicase A to polysomes. *Nucleic Acids Res.* **2011**, *39*, 3724–3734.
25. Polesskaya, A.; Cuvellier, S.; Naguibneva, I.; Duquet, A.; Moss, E.; Harel-Bellan, A. Lin-28 binds IGF-2 mRNA and participates in skeletal myogenesis by increasing translation efficiency. *Genes Dev.* **2007**, *21*, 1125–1138.
26. Pasquinelli, A.E.; Reinhart, B.J.; Slack, F.; Martindale, M.Q.; Kuroda, M.I.; Maller, B.; Hayward, D.C.; Ball, E.E.; Degnan, B.; Müller, P.; *et al.* Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **2000**, *408*, 86–89.
27. Reinhart, B.; Slack, F.; Basson, M.; Pasquinelli, A.; Bettinger, J.; Rougvie, A.; Horvitz, H.; Ruvkun, G. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **2000**, *403*, 901–906.
28. Van Wynsberghe, P.; Kai, Z.; Massirer, K.; Burton, V.; Yeo, G.; Pasquinelli, A. LIN-28 co-transcriptionally binds primary let-7 to regulate miRNA maturation in *Caenorhabditis elegans*. *Nat. Struct. Mol. Biol.* **2011**, *18*, 302–308.
29. Mondol, V.; Pasquinelli, A.E. Let's make it happen: The role of let-7 microRNA in development. *Curr. Top. Dev. Biol.* **2012**, *99*, 1–30.
30. Thomson, J.M.; Newman, M.; Parker, J.S.; Morin-Kensicki, E.M.; Wright, T.; Hammond, S.M. Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes Dev.* **2006**, *20*, 2202–2207.
31. Wulczyn, F.; Smirnova, L.; Rybak, A.; Brandt, C.; Kwidzinski, E.; Ninnemann, O.; Strehle, M.; Seiler, A.; Schumacher, S.; Nitsch, R. Post-transcriptional regulation of the let-7 microRNA during neural cell specification. *FASEB J.* **2007**, *21*, 415–426.
32. Yang, D.H.; Moss, E.G. Temporally regulated expression of Lin-28 in diverse tissues of the developing mouse. *Gene Expr. Patterns* **2003**, *3*, 719–726.

33. Piskounova, E.; Viswanathan, S.R.; Janas, M.; LaPierre, R.J.; Daley, G.Q.; Sliz, P.; Gregory, R.I. Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28. *J. Biol. Chem.* **2008**, *283*, 21310–21314.
34. Mayr, F.; Schutz, A.; Doge, N.; Heinemann, U. The Lin28 cold-shock domain remodels pre-let-7 microRNA. *Nucleic Acids Res.* **2012**, *40*, 7492–7506.
35. Nam, Y.; Chen, C.; Gregory, R.I.; Chou, J.J.; Sliz, P. Molecular basis for interaction of let-7 microRNAs with Lin28. *Cell* **2011**, *147*, 1080–1091.
36. Lightfoot, H.L.; Bugaut, A.; Armisen, J.; Lehrbach, N.J.; Miska, E.A.; Balasubramanian, S. A LIN28-dependent structural change in pre-let-7g directly inhibits dicer processing. *Biochemistry* **2011**, *50*, 7514–7521.
37. Heo, I.; Joo, C.; Cho, J.; Ha, M.; Han, J.; Kim, V.N. Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Mol. Cell* **2008**, *32*, 276–284.
38. Mullen, T.E.; Marzluff, W.F. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes Dev.* **2008**, *22*, 50–65.
39. Shen, B.; Goodman, H.M. Uridine addition after microRNA-directed cleavage. *Science* **2004**, *306*, 997.
40. Chang, H.M.; Triboulet, R.; Thornton, J.E.; Gregory, R.I. A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. *Nature* **2013**, *497*, 244–248.
41. Hagan, J.P.; Piskounova, E.; Gregory, R.I. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.* **2009**, *16*, 1021–1025.
42. Thornton, J.E.; Chang, H.M.; Piskounova, E.; Gregory, R.I. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *RNA* **2012**, *18*, 1875–1885.
43. Heo, I.; Ha, M.; Lim, J.; Yoon, M.J.; Park, J.E.; Kwon, S.C.; Chang, H.; Kim, V.N. Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. *Cell* **2012**, *151*, 521–532.
44. Lehrbach, N.J.; Armisen, J.; Lightfoot, H.L.; Murfitt, K.J.; Bugaut, A.; Balasubramanian, S.; Miska, E.A. LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans. *Nat. Struct. Mol. Biol.* **2009**, *16*, 1016–1020.
45. Kawahara, H.; Okada, Y.; Imai, T.; Iwanami, A.; Mischel, P.S.; Okano, H. Musashi1 cooperates in abnormal cell lineage protein 28 (Lin28)-mediated let-7 family microRNA biogenesis in early neural differentiation. *J. Biol. Chem.* **2011**, *286*, 16121–16130.
46. Lei, X.X.; Xu, J.; Ma, W.; Qiao, C.; Newman, M.A.; Hammond, S.M.; Huang, Y. Determinants of mRNA recognition and translation regulation by Lin28. *Nucleic Acids Res.* **2012**, *40*, 3574–3584.
47. Xu, B.; Huang, Y. Histone H2a mRNA interacts with Lin28 and contains a Lin28-dependent posttranscriptional regulatory element. *Nucleic Acids Res.* **2009**, *37*, 4256–4263.
48. Xu, B.; Zhang, K.; Huang, Y. Lin28 modulates cell growth and associates with a subset of cell cycle regulator mRNAs in mouse embryonic stem cells. *RNA* **2009**, *15*, 357–361.
49. Cho, J.; Chang, H.; Kwon, S.C.; Kim, B.; Kim, Y.; Choe, J.; Ha, M.; Kim, Y.K.; Kim, V.N. LIN28A is a suppressor of ER-associated translation in embryonic stem cells. *Cell* **2012**, *151*, 765–777.

50. Qiu, C.; Ma, Y.; Wang, J.; Peng, S.; Huang, Y. Lin28-mediated post-transcriptional regulation of Oct4 expression in human embryonic stem cells. *Nucleic Acids Res.* **2010**, *38*, 1240–1248.
51. Winter, J.; Jung, S.; Keller, S.; Gregory, R.I.; Diederichs, S. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat. Cell Biol*. **2009**, *11*, 228–234.
52. Peng, S.; Chen, L.-L.; Lei, X.-X.; Yang, L.; Lin, H.; Carmichael, G.G.; Huang, Y. Genome-wide studies reveal that Lin28 enhances the translation of genes important for growth and survival of human embryonic stem cells. *Stem Cells* **2011**, *29*, 496–504.
53. Graf, R.; Munschauer, M.; Mastrobuoni, G.; Mayr, F.; Heinemann, U.; Kempa, F.; Rajewski, N.; Landthaler, M. Identification of LIN28-bound mRNAs reveals features of target recognition and regulation. *RNA Biol.* **2013**, *10*, 1146–1159.
54. Wilbert, M.L.; Huelga, S.C.; Kapeli, K.; Stark, T.J.; Liang, T.Y.; Chen, S.X.; Yan, B.Y.; Nathanson, J.L.; Hutt, K.R.; Lovci, M.T.; *et al.* LIN28 binds messenger RNAs at GGAGA motifs and regulates splicing factor abundance. *Mol. Cell* **2012**, *48*, 195–206.
55. Cox, J.; Mallanna, S.; Luo, X.; Rizzino, A. Sox2 uses multiple domains to associate with proteins present in Sox2-protein complexes. *PLoS One* **2010**, *5*, e15486.
56. Marson, A.; Levine, S.S.; Cole, M.F.; Frampton, G.M.; Brambrink, T.; Johnstone, S.; Guenther, M.G.; Johnston, W.K.; Wernig, M.; Newman, J.; *et al.* Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **2008**, *134*, 521–533.
57. Melton, C.; Judson, R.L.; Blelloch, R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature* **2010**, *463*, 621–626.
58. Hanna, J.; Saha, K.; Pando, B.; van Zon, J.; Lengner, C.J.; Creyghton, M.P.; van Oudenaarden, A.; Jaenisch, R. Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **2009**, *462*, 595–601.
59. Li, N.; Zhong, X.; Lin, X.; Guo, J.; Zou, L.; Tanyi, J.L.; Shao, Z.; Liang, S.; Wang, L.P.; Hwang, W.T.; *et al.* Lin-28 homologue A (LIN28A) promotes cell cycle progression via regulation of cyclin-dependent kinase 2 (CDK2), cyclin D1 (CCND1), and cell division cycle 25 homolog A (CDC25A) expression in cancer. *J. Biol. Chem.* **2012**, *287*, 17386–17397.
60. Chang, T.C.; Zeitels, L.R.; Hwang, H.W.; Chivukula, R.R.; Wentzel, E.A.; Dews, M.; Jung, J.; Gao, P.; Dang, C.V.; Beer, M.A.; *et al*. Lin-28B transactivation is necessary for Myc-mediated let-7 repression and proliferation. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 3384–3389.
61. Iliopoulos, D.; Hirsch, H.A.; Struhl, K. An epigenetic switch involving NF-kappaB, Lin28, Let-7 MicroRNA, and IL6 links inflammation to cell transformation. *Cell* **2009**, *139*, 693–706.
62. Yang, X.; Lin, X.; Zhong, X.; Kaur, S.; Li, N.; Liang, S.; Lassus, H.; Wang, L.; Katsaros, D.; Montone, K.; *et al.* Double-negative feedback loop between reprogramming factor LIN28 and microRNA let-7 regulates aldehyde dehydrogenase 1-positive cancer stem cells. *Cancer Res.* **2010**, *70*, 9463–9472.
63. Krizhanovsky, V.; Lowe, S.W. Stem cells: The promises and perils of p53. *Nature* **2009**, *460*, 1085–1086.
64. Lettre, G.; Jackson, A.U.; Gieger, C.; Schumacher, F.R.; Berndt, S.I.; Sanna, S.; Eyheramendy, S.; Voight, B.F.; Butler, J.L.; *et al.* Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat. Genet.* **2008**, *40*, 584–591.

65. Ong, K.K.; Elks, C.E.; Li, S.; Zhao, J.H.; Luan, J.; Andersen, L.B.; Bingham, S.A.; Brage, S.; Smith, G.D.; Ekelund, U.; *et al.* Genetic variation in LIN28B is associated with the timing of puberty. *Nat. Genet.* **2009**, *41*, 729–733.
66. Perry, J.R.; Stolk, L.; Franceschini, N.; Lunetta, K.L.; Zhai, G.; McArdle, P.F.; Smith, A.V.; Aspelund, T.; Bandinelli, S.; Boerwinkle, E.; *et al.* Meta-analysis of genome-wide association data identifies two loci influencing age at menarche. *Nat. Genet.* **2009**, *41*, 648–650.
67. Frost, R.J.; Olson, E.N. Control of glucose homeostasis and insulin sensitivity by the Let-7 family of microRNAs. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 21075–21080.
68. Amarasinghe, G.K.; de Guzman, R.N.; Turner, R.B.; Chancellor, K.J.; Wu, Z.R.; Summers, M.F. NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. *J. Mol. Biol.* **2000**, *301*, 491–511.
69. De Guzman, R.N.; Wu, Z.R.; Stalling, C.C.; Pappalardo, L.; Borer, P.N.; Summers, M.F. Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science* **1998**, *279*, 384–388.
70. Pappalardo, L.; Kerwood, D.J.; Pelczer, I.; Borer, P.N. Three-dimensional folding of an RNA hairpin required for packaging HIV-1. *J. Mol. Biol.* **1998**, *282*, 801–818.
71. Desjardins, A.; Yang, A.; Bouvette, J.; Omichinski, J.G.; Legault, P. Importance of the NCp7-like domain in the recognition of pre-let-7g by the pluripotency factor Lin28. *Nucleic Acids Res.* **2012**, *40*, 1767–1777.
72. Loughlin, F.E.; Gebert, L.F.; Towbin, H.; Brunschweiger, A.; Hall, J.; Allain, F.H. Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28. *Nat. Struct. Mol. Biol.* **2012**, *19*, 84–89.
73. Guerrerio, A.L.; Berg, J.M. Design of single-stranded nucleic acid binding peptides based on nucleocapsid CCHC-box zinc-binding domains. *J. Am. Chem. Soc.* **2010**, *132*, 9638–9643.
74. Mayr, F. *Structural and Functional Analysis of Lin28-mediated Inhibition of let-7 miRNA Biogenesis*; Freie Universität Berlin: Berlin, Germany, 2013.
75. Mihailovich, M.; Militti, C.; Gabaldón, T.; Gebauer, F. Eukaryotic cold shock domain proteins: Highly versatile regulators of gene expression. *Bioessays* **2010**, *32*, 109–118.
76. Lopez, M.M.; Makhatadze, G.I. Major cold shock proteins, CspA from *Escherichia coli* and CspB from *Bacillus subtilis*, interact differently with single-stranded DNA templates. *Biochim. Biophys. Acta* **2000**, *1479*, 196–202.
77. Lopez, M.M.; Yutani, K.; Makhatadze, G.I. Interactions of the cold shock protein CspB from *Bacillus subtilis* with single-stranded DNA. Importance of the T base content and position within the template. *J. Biol. Chem.* **2001**, *276*, 15511–15518.
78. Max, K.E.A.; Zeeb, M.; Bienert, R.; Balbach, J.; Heinemann, U. T-rich DNA single strands bind to a preformed site on the bacterial cold shock protein Bs-CspB. *J. Mol. Biol.* **2006**, *360*, 702–714.
79. Max, K.E.A.; Zeeb, M.; Bienert, R.; Balbach, J.; Heinemann, U. Common mode of DNA binding to cold shock domains. Crystal structure of hexathymidine bound to the domain-swapped form of a major cold shock protein from *Bacillus caldolyticus*. *FEBS J.* **2007**, *274*, 1265–1279.
80. Morgan, H.P.; Estibeiro, P.; Wear, M.A.; Max, K.E.A.; Heinemann, U.; Cubeddu, L.; Gallagher, M.P.; Sadler, P.J.; Walkinshaw, M.D. Sequence specificity of single-stranded DNA-binding proteins: A novel DNA microarray approach. *Nucleic Acids Res.* **2007**, *35*, e75.

81. Zeeb, M.; Max, K.E.A.; Weininger, U.; Löw, C.; Sticht, H.; Balbach, J. Recognition of T-rich single-stranded DNA by the cold shock protein Bs-CspB in solution. *Nucleic Acids Res.* **2006**, *34*, 4561–4571.
82. Phadtare, S. Unwinding activity of cold shock proteins and RNA metabolism. *RNA Biol.* **2011**, *8*, 394–397.
83. Phadtare, S.; Severinov, K. Nucleic acid melting by *Escherichia coli* CspE. *Nucleic Acids Res.* **2005**, *33*, 5583–5590.
84. Phadtare, S.; Tyagi, S.; Inouye, M.; Severinov, K. Three amino acids in *Escherichia coli* CspE surface-exposed aromatic patch are critical for nucleic acid melting activity leading to transcription antitermination and cold acclimation of cells. *J. Biol. Chem.* **2002**, *277*, 46706–46711.
85. Newkirk, K.; Feng, W.; Jiang, W.; Tejero, R.; Emerson, S.D.; Inouye, M.; Montelione, G.T. Solution NMR structure of the major cold shock protein (CspA) from *Escherichia coli*: Identification of a binding epitope for DNA. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 5114–5118.
86. Schindelin, H.; Herrler, M.; Willimsky, G.; Marahiel, M.A.; Heinemann, U. Overproduction, crystallization, and preliminary X-ray diffraction studies of the major cold shock protein from *Bacillus subtilis*, CspB. *Proteins* **1992**, *14*, 120–124.
87. Schindelin, H.; Jiang, W.; Inouye, M.; Heinemann, U. Crystal structure of CspA, the major cold shock protein of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 5119–5123.
88. Schindelin, H.; Marahiel, M.A.; Heinemann, U. Universal nucleic acid-binding domain revealed by crystal structure of the *B. subtilis* major cold-shock protein. *Nature* **1993**, *364*, 164–168.
89. Schnuchel, A.; Wiltscheck, R.; Czisch, M.; Herrler, M.; Willimsky, G.; Graumann, P.; Marahiel, M.A.; Holak, T.A. Structure in solution of the major cold-shock protein from Bacillus subtilis. *Nature* **1993**, *364*, 169–171.
90. Sachs, R.; Max, K.E.; Heinemann, U.; Balbach, J. RNA single strands bind to a conserved surface of the major cold shock protein in crystals and solution. *RNA* **2012**, *18*, 65–76.
91. Phadtare, S.; Inouye, M.; Severinov, K. The mechanism of nucleic acid melting by a CspA family protein. *J. Mol. Biol.* **2004**, *337*, 147–155.

Reprinted from *IJMS*. Cite as: Doxakis, E. Principles of miRNA-Target Regulation in Metazoan Models. *Int. J. Mol. Sci.* **2013**, *14*, 16280-16302.

Review

Principles of miRNA-Target Regulation in Metazoan Models

Epaminondas Doxakis

Basic Neurosciences Division, Biomedical Research Foundation of the Academy of Athens, Soranou Efesiou 4, Athens 11527, Greece; E-Mail: edoxakis@bioacademy.gr; Tel.: +30-210-6597-479; Fax: +30-210-6597-545

Received: 3 June 2013; in revised form: 29 July 2013 / Accepted: 31 July 2013 / Published: 7 August 2013

Abstract: MicroRNAs (miRs) are key post-transcriptional regulators that silence gene expression by direct base pairing to target sites of RNAs. They have a wide variety of tissue expression patterns and are differentially expressed during development and disease. Their activity and abundance is subject to various levels of control ranging from transcription and biogenesis to miR response elements on RNAs, target cellular levels and miR turnover. This review summarizes and discusses current knowledge on the regulation of miR activity and concludes with novel non-canonical functions that have recently emerged.

Keywords: miR; miR biogenesis; miR targets; miR turnover; isomiR; ceRNA

1. Introduction

Mature microRNAs (miRs) are a class of highly conserved small non-coding RNA molecules, about 22 *nucleotides* in length, that act to inhibit protein expression by partially hybridizing to complementary sequences, mainly in the 3' UTR, of target RNA transcripts. Each miR is estimated to regulate multiple functionally-related target mRNAs, and the combinatorial action of miRs is expected to regulate the expression of hundreds of mRNAs. Currently, over 1100 and 1800 miRs have been annotated and categorized in mice and humans, respectively (miRBase 20, [1]). However, these numbers are likely to be inflated by mistakenly identified miRs [2]. In addition, the high rate of miR family turnover in mammals—with many newly emerged miR families being lost soon after their formation—indicate that many more of the truly-identified miRs are likely to have little functional significance [3]. It is now predicted that more than half of human genes are regulated by miRs [4]. miRs have a wide variety of tissue expression patterns and are differentially expressed during development [5–8]. They are deregulated in most human diseases and the profiles they generate carry more diagnostic information than those of mRNAs or proteins [9]. Moreover, the therapeutic potential of

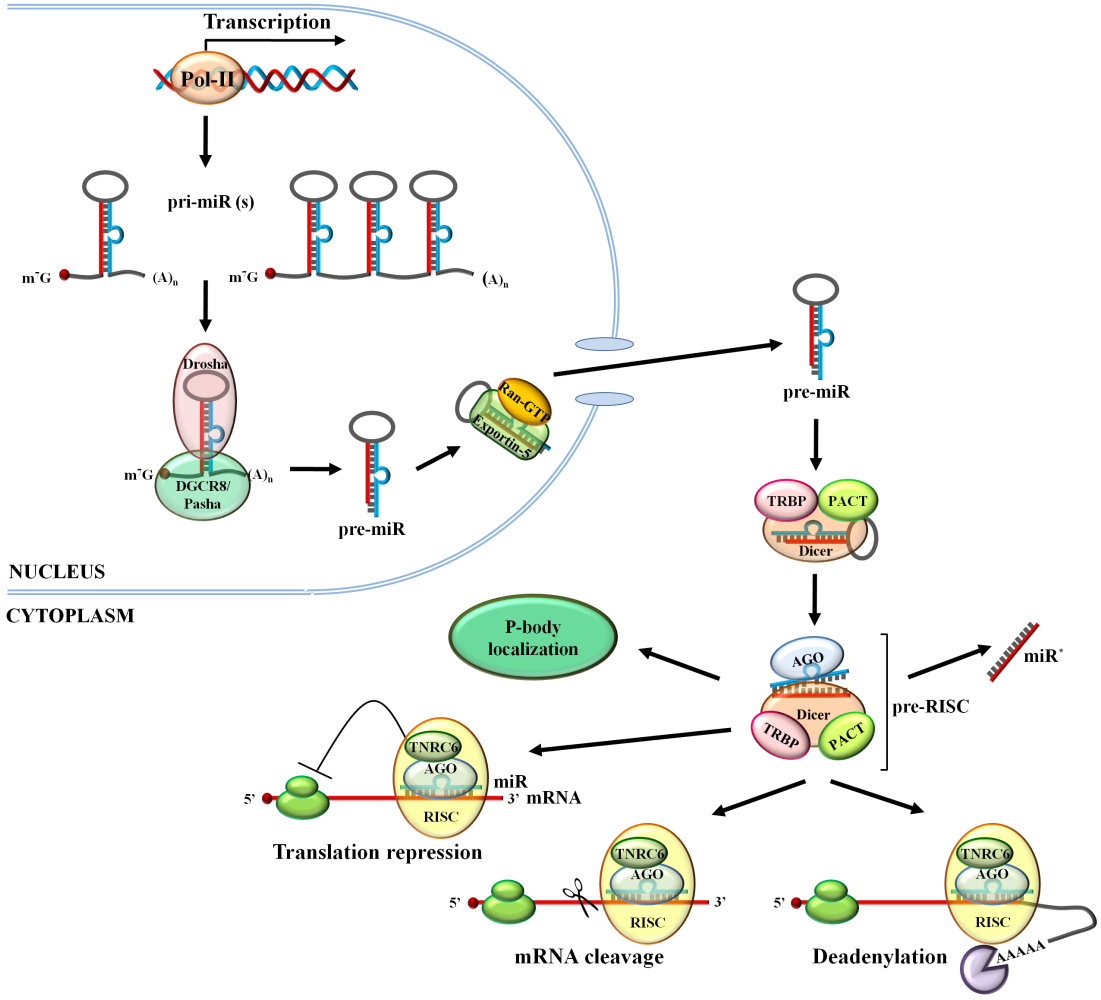
miRs, already demonstrated in numerous studies, has further heightened the importance of research that seeks to understand both their mechanism of action and their biological significance [10–13].

This paper aims at reviewing the latest information on miR biogenesis and the factors that determine the efficacy of miR-mediated repression and miR endogenous levels. It concludes with novel atypical functions that stand-out from the canonical repression activity of miRs.

2. miR Biogenesis

miRs are transcribed as part of longer primary transcripts (pri-miRs) by, mainly, RNA polymerase II (Pol II) and only few by RNA polymerase III (Pol III) [14–19]. The majority of miR genes are transcribed from introns of protein-coding genes. The remaining are transcribed as part of long non-coding RNAs that are often arranged as clusters that lead to one pri-miR being subsequently, processed into several mature miRs (Figure 1) [17,20]. Similar to other Pol II transcripts, pri-miRs possess a 5' 7-methyl-guanosine cap and a 3' poly (A) tail, the use of which is currently poorly understood. Within pri-miR long transcripts, mature miR sequences form hairpin structures that contain imperfect double-stranded stems of ~30 bp connected by a terminal loop at the top and single-stranded RNA segments at the base [21,22]. In the canonical miR biogenesis pathway, these hairpin structures are recognized in the nucleus and cleaved by a multi-protein microprocessor complex that is composed of two core components, Drosha (a RNase III ribonuclease) and DGCR8 (also known as Pasha in invertebrates which is a double-stranded RNA binding protein). Mechanistically, DGCR8, initially, recognizes the base of the miR hairpin structure and then guides Drosha to cleave the pri-miR at a distance of ~11 bp from the base generating a ~70 nucleotide (nt) hairpin RNA (named precursor miR or pre-miR) with a 2 nt 3' overhang [21–26]. This 3' overhang and the double-stranded hairpin structure of the pre-miR are subsequently recognized by exportin-5, which together with its cofactor RAN-GTP, shuttle pre-miR from the nucleus into the cytosol. The hydrolysis of GTP bound to RAN in the cytosol triggers the dissociation of the complex, allowing the pre-miR to bind Dicer, a double-stranded ribonuclease III. Dicer cleaves the pre-miR terminal loop in concert with its cofactors TRBP (also known as Loqs in *Drosophila*) and PACT. In this process, Dicer binds to the pre-miR 2 nt 3' overhang and cuts two helical turns (~22 nt) away to produce a double-stranded RNA with 3' overhangs of 2 nt at both ends. TRBP and PACT regulate Dicer's substrate recognition and RNA processing power but are not essential for Dicer's slicing activity [27–33]. After cleavage, the strand with the 5' terminus that has less stable base-pairing (the "guide strand") is transferred onto an Argonaute (Ago) protein, which is part of a poorly defined multi-protein miR-induced silencing complex (miRISC) that includes Dicer, TRBP and TNRC6 (also known as GW182) whereas the other strand (the "passenger" or "star strand") is degraded [34–37]. Additional features on the miR duplex may also play a role in the strand selection process and there are several miRs where both strands are incorporated, to varying degrees, onto Ago proteins [38–41]. Once in place, the miR nucleotide sequence serves as a guide for RNA interference (RNAi) based on the partial complementarity with the various RNA substrates, a process which is largely attained by random diffusion of miRISC into the cytosol [42]. TNRC6 proteins are essential for RNAi as they interact with poly(A)-binding protein (PABP) and the PAN2-PAN3 and CCR4-NOT deadenylase complexes to induce translation repression, deadenylation and decay of the mRNA targets [43–45].

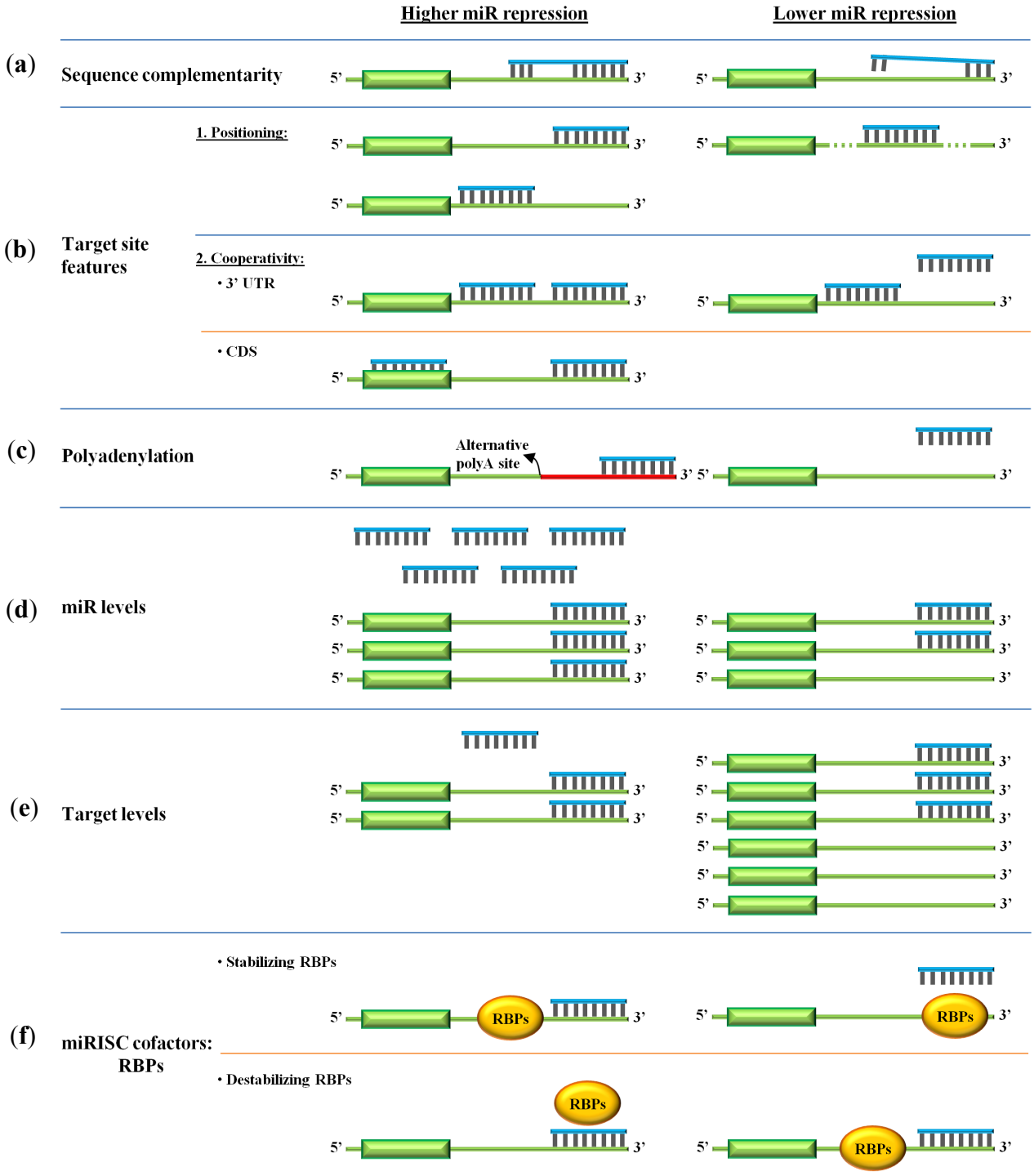
Figure 1. miR biogenesis. Monocistronic or polycistronic miRs are transcribed by RNA polymerase II into long pri-miR transcripts. These pri-miRs are, subsequently, processed by RNase III Drosha complex to ~70 nt pre-miRs that are exported out of the nucleus and into the cytosol by Exportin-5. In the cytoplasm, the RNase III Dicer complex cleaves pre-miRs to double-stranded ~22 nt miRs. One strand is then selected and loaded onto an Argonaute protein, which is part of the miRISC complex. The single-stranded miR then serves as a guide for RNA interference based on the partial complementarity with the various RNA substrates.



3. Efficacy of miR Repression

Despite a wealth of genome-wide and biochemical data on the role of miRs in the regulation of their targets, we do not yet have a clear understanding of the factors that determine which mRNAs will be targeted by miRs or by which mechanism individual mRNAs will be silenced, that is, translation repression or mRNA destabilization. Likely, this reflects on the vast repertoire of context-specific determinants that modulate miR-target interactions (Figure 2).

Figure 2. miR repression determinants. Multiple factors determine repression effectiveness of miRs. These include: (**a**) Sequence complementarity at positions 2–7 or 2–8 of the 5' end of the mature miR; (**b**) Target site features: binding site location near the edges of 3' UTR or multiple binding sites for miRs; (**c**) Alternative cleavage and polyadenylation to maintain miR binding sites; (**d**) Relatively high miR levels; (**e**) Relatively low target levels; and (**f**) Presence of stabilizing and absence of destabilizing RNA binding protein sites.



3.1. The miR Nucleotide Sequence

Large-scale transcriptomic and proteomic studies have revealed that the primary determinant for miR binding is perfect consecutive Watson-Crick base-pairing between the target RNA and the miR at positions 2–7 or 2–8 of the 5' end of the mature miR, often denoted as the "seed" region [46–49]. This signature has been reaffirmed with crystallographic studies of ribonucleoprotein Ago-miR complexes showing that the seed region is organized in a helical conformation that exposes it to base-pair with the target RNA [50–52]. More recently, a genome-wide analysis of Ago sites in murine brain revealed a variant of this target recognition pattern through a single bulged nucleotide in the middle of the 2–7 seed. These bulged sites, that likely yield overall lower repression, are evolutionarily conserved and comprise over 15% of all Ago-miR interactions, thus, expanding significantly the number of potential miR regulatory sites [53]. Despite the aforementioned basic features, a "seed" is neither necessary nor sufficient for target silencing. It has been shown that miR target sites can often tolerate G:U wobble base pairs within the seed region [54,55] and extensive base pairing at the 3' end of the miR may offset missing complementarity at the seed region [46,56]. Moreover, centered sites have also been reported showing 11–12 contiguous nt base-pairing to the central region of the miR without pairing to either end [57]. To add to this repertoire, other studies report efficient silencing from sites that do not fit to any of the above patterns, seemingly appearing random [58,59], and even sites with extensive 5' complementarity can be inactive when tested in reporter constructs [60].

3.2. Target Site Features

Considerable progress has been made to identify additional features that could predict target regulation with more precision. Grimson *et al.* have reported that local sequence context, such as AU-rich nucleotide composition near the target site, proximity to sites for co-expressed miRs, proximity to residues pairing to miR nt 13–16, and positioning away from the center of long 3' UTRs can all promote efficient miR repression of targets [61]. With respect to these findings, several different studies have reaffirmed that multiple miR sites in the same 3' UTR can potentiate the degree of translational repression. They reported that optimal downregulation is obtained when two sites are closely positioned, usually between 13 and 35 nt apart [62,63]. However, target sites spaced at substantially longer distances may still cooperate to lower the expression of proteins [64,65]. In this context, miR cooperativity is defined as the positive interaction of two or more individual miRs or one individual miR acting on multiple target sites on the same 3' UTR for target regulation. Recently, it was estimated that the miR site density of brain synaptic mRNAs is twice higher than that of the rest of cellular mRNAs, indicating that miR cooperativity may be a prevalent mechanism for physiological processes that require precise control, such as synaptic transmission [65].

An additional feature that has also emerged is that miR target sites tend to be less evolutionary preserved in the first ~15 nt downstream of the stop codon, presumably, to avoid being in the path of the translational machinery that could displace the miRISC complex [61]. However, both computational and biochemical approaches have later identified that nearly half of miR sites are located in open reading frame (ORF) sequences [66–70]. Experimental analysis indicated that the sites in coding regions and to a lesser extent 5' UTRs can confer miR repression, albeit at lower levels than

3' UTRs [71,72]. Recently, it was reported that coding region-located sites induce more rapid reduction in mRNA translation than 3' UTR-located sites in a process that does not involve mRNA degradation, however, the effect may only be transient. The authors elaborate that this type of response may be suited for the regulation of cell cycle proteins [73]. Further, there are several families of paralogous genes that contain multiple repeat sequences in their coding regions, arisen through evolutionary duplications, that are miR targets [74]. Like for miR cooperativity in 3' UTRs, it was shown that miR sites in the coding region potentiate the repression activity of miRs acting on 3' UTR [75].

To add a twist to miR regulation, it has been reported that individual miRs may also display distinct preference for binding to different regions of an mRNA. For instance, neuronal miR-124 seed sequences are preferentially located in the 3' UTR, while miR-107 seed sequences are enriched in the coding region of the mRNAs. Further, mRNA targets of neuronal miR-128 and miR-320 are less enriched for 6-mer seed sequences than miR-124 and mir-107 [76]. The reason for these differences is unknown but they, evidently, enrich the heterogeneity of miR-mediated repression.

3.3. Target Accessibility and Polyadenylation

Another important determinant of efficient silencing is target RNA folding with several reports indicating that miR sites are preferentially positioned in highly accessible/unstructured regions at the start and end of 3' UTRs [61–79]. Experimentally, target sites in the middle of 3' UTR have been found to be less efficient for RNAi regulation [80] while those positioned near the end of 3' UTRs are associated with highest repression [62].

Another contributing factor is the length of the 3' UTR. Approximately, half of human genes undergo alternative cleavage and polyadenylation (pA) to generate transcripts with variable 3' UTR lengths [53]. Given that 3' UTRs are the main targets of miRs, alternative pA is expected to modify target RNA translation. Consequently, a close connection between gene transcription and pA site choice has been demonstrated, in which highly expressed genes contain shorter 3' UTRs while transcripts that are expressed in lower levels are associated with longer 3' UTR isoforms [81]. Along this, higher gene expression is tightly linked to cell division where short 3' UTR isoforms with fewer miR sites are abundant in proliferating cells [82]. In contrast, differentiated cells possess longer 3' UTRs [81]. A noteworthy consequence of alternative splicing was observed in transformed cells where the loss of miR target sites by pA contributed to oncogene activation without any apparent mutagenesis [82].

3.4. miR and Target RNA Levels

An additional critical determinant for miR repression effectiveness is the cellular concentrations of (a) the target RNA, (b) the miR and (c) the miRISC complex. miRs that have multiple targets and are not highly expressed are expected to downregulate individual target genes to a lesser extent than those with a lower number of targets. Similarly, highly abundant target transcripts, that may act as decoys, dilute the effect of miRs under differential conditions [83–85] and this effect is more pronounced when the decoys are capable of perfect base-pairing with the miR [86]. Along these lines, it has been observed that lower levels of a miR may fail to regulate its target mRNA, however, it retains the ability to promote inhibition in conjunction with another miR, indicating that cooperative silencing requires lower concentration of miRs [65]. Going beyond, it is predicted that imbalances in the relative

concentrations of miRs and their gene targets may exaggerate or compensate for sequence mismatches between miR and target RNA pairs. miRISC stability has emerged as an additional level at which miR activity can be controlled. Specifically, LIN41, an E3 ubiquitin ligase, has been shown to target Ago2 for ubiquitination and proteasome degradation. Because Ago proteins are limiting factors for the activity of the miRISC complex, alterations in the levels of LIN41 result in global attenuation of miR-mediated repression [87].

3.5. RNA Binding Proteins

RNA binding proteins (RBPs) regulate key aspects of gene expression including pre-mRNA splicing, nuclear-cytosolic shuttling, cytosolic transport and storage, local translation and turnover. Although most RBPs have housekeeping functions, a subset of RBPs controls the expression of specific labile mRNAs by binding to U- and AU-rich elements (AREs) on either 5' UTR or 3' UTR. These include HuR, TIAR, TIA-1, AUF1, TTP and KSRP, collectively known as translation and turnover regulatory (TTR)-RBPs. They primarily modulate mRNA levels in response to external stimuli and have been shown to influence all aspects of cellular activities that include proliferation and differentiation [88]. Recently, a link between RBPs and miRs has emerged. Initially, it has been observed that destabilization mediated by a transfected miR is generally attenuated by the presence of destabilizing AU-rich motifs and augmented by stabilizing U-rich motifs, the binding sites of TTR-RBPs [89,90]. Subsequently, transcriptome-wide analysis for the best characterized ubiquitous RBP, HuR revealed that most miR sites were found in the immediate vicinity of HuR sites [91,92] (reviewed in [93]). The authors elaborated that where miR and HuR sites overlapped the transcripts were preferentially regulated by HuR, but when they were not overlapping the transcripts were regulated by miR. Interestingly, HuR transcript is itself a direct target of miRs and of itself, and at the same time, directly regulates stability and/or maturation of other miRs pointing to the vast repertoire of the different regulatory loops [91–96].

4. Availability of miRs

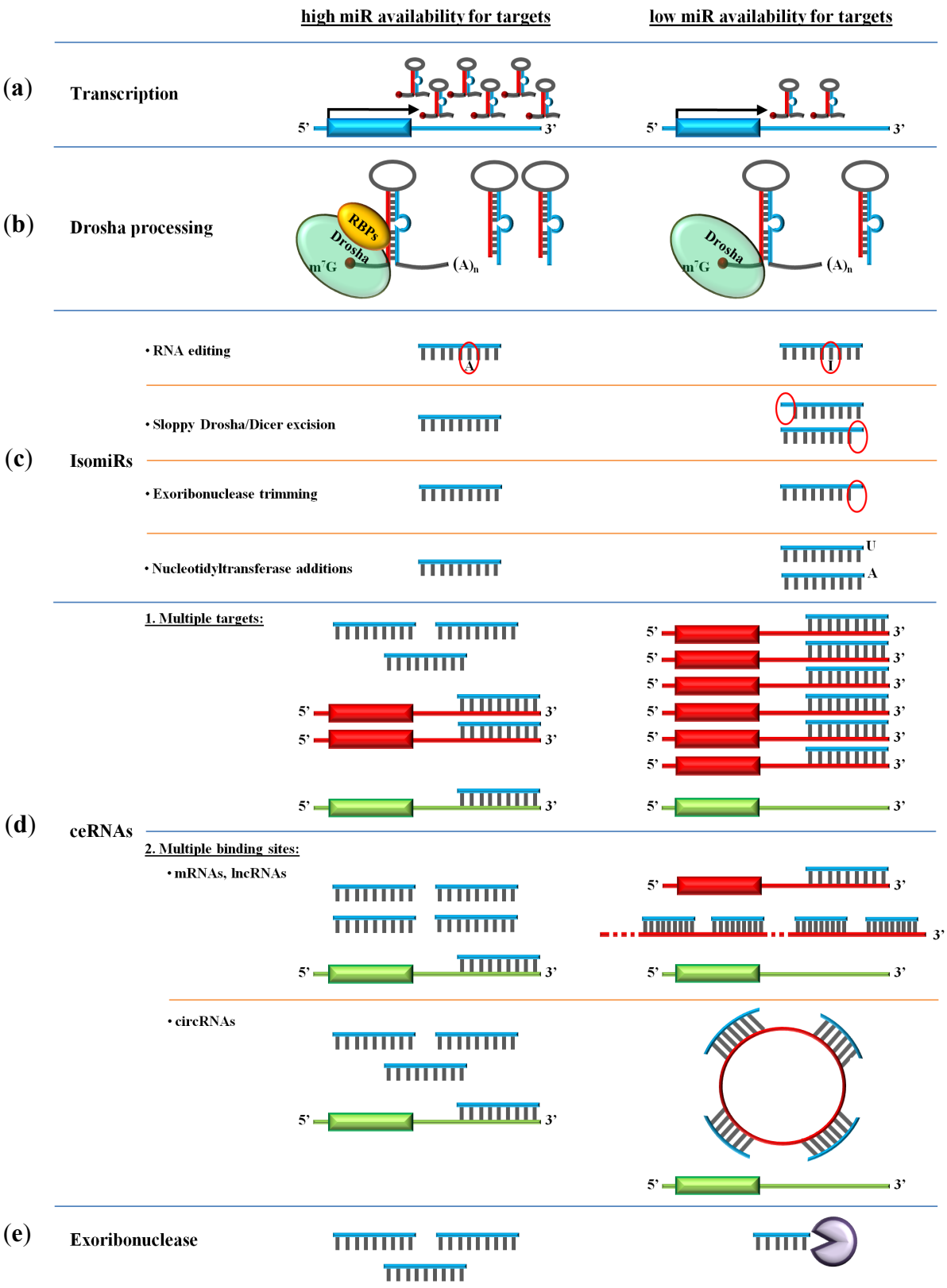
It has become increasingly evident that miR activity is determined not only by target site features but also miR levels, target abundance and the presence of multiple RNA decoys (Figure 3). It is the summation of all these inputs that ultimately shapes miR function.

4.1. Transcription

Most miRs are transcribed by RNA polymerase II (Pol II) and only few by RNA polymerase III (Pol III) [14–19,97]. Pol III-mediated transcription is usually restricted to housekeeping non-coding genes, such as tRNAs and snRNAs, that require ubiquitous expression under all conditions [98], whereas Pol II-mediated transcription permits tight control of expression during all types of regulatory conditions [99]. Nonetheless, there is evidence that the same promoter elements can be used by both polymerases in humans [100–102] and transcription factors can also regulate RNA Pol III activity to some degree [103]. Furthermore, whole genome analysis has revealed that miR promoters are, in

general, very similar to protein-coding promoters containing proportionally similar levels of CpG islands, TATA boxes, TFIIB recognition elements (BRE) and initiators (Inr) [16].

Figure 3. miR availability determinants. Multiple factors determine availability of miRs. These include: (**a**) High transcription rates; (**b**) Enhanced Drosha processing; (**c**) Lower levels of isomiRs that result from RNA editing, sloppy Drosha/Dicer cleavage, exoribonuclease trimming and nucleotidyl transferase additions; (**d**) Lower levels of miR sequestering ceRNAs; and (**e**) Lower levels of exoribonucleases.



The majority of miR genes are transcribed from introns (and to lesser extent exons) of protein-coding genes [17,20]. As a consequence, miRs, more often than not, are co-expressed with host genes [104]. Nevertheless, increasingly, there have been reports that showed that intragenic miRs could, independently, initiate transcription from own promoters [105–107]. It is now estimated that about a third of hosted miRs use their own promoters for more efficient and tailored

transcription [14,16,108]. With respect to the miRs that are located in intergenic regions, these are often arranged in clusters that lead to one pri-miR being subsequently processed into several mature miRs [104]. Using microarray profiling, Baskerville and Bartel have proposed that miRs separated by <50 kb are typically derived from a common transcript [104]. Accordingly, the latest miRBase release (Release 20) groups human and murine miRs in 153 (containing 465 miRs) and 92 (containing 366 miRs) clusters, respectively, using a default of <10 kb inter-miR distance. Clusters provide an effective mechanism to express cooperative miRs, simultaneously. Many clusters contain representatives from different miR families that together regulate specific protein networks by co-targeting downstream mRNAs [109]. This provides another layer of coordinated system-wide regulation of gene output in cells [110].

4.2. Drosha Processing

Drosha has been shown to exert selectivity over its pri-miR substrates compared to other RNAs. The mechanism by which this is achieved differs between miRs. Thus far, microarray profiling has shown that subsets of miRs contain a Smad binding RNA sequence (R-SBE) within the stem region of the pri-miR that resembles the Smad binding element in DNA. Smad proteins bind to these motifs on the miRs with one (MH1) domain while another (MH2) domain binds p68, a protein that is integral part of the microprocessor complex in the nucleus and is known to induce Drosha processing [111,112]. Similarly, DNA damage induces p53 association with p68, promoting the processing of specific miRs that subsequently exert a tumor suppressor function via repression of c-myc [113]. A different mode of regulation is demonstrated by the RNA binding proteins KHSRP and hnRNPA1 that bind to specific single- and double- stranded segments on the pre-miR hairpin, respectively, inducing microprocessor complex cleavage. Importantly, this targeted processing of the pri-miR has been shown to uncouple the uniform expression levels of clustered miRs from the maturation efficiency of individual miRs [114–116].

4.3. miR Polymorphism and isomiRs

Computational predictions have strongly suggested that miRs may have shaped the evolution of their targets based on the fact that the conservation of predicted miR target sites in mRNAs is higher than that of other conserved 3' UTR motifs [117]. Consequently, polymorphisms in miR sequences were presumed rare. Towards this, bioinformatic analysis has revealed that the density of single nucleotide polymorphisms (SNPs) in miRs is 4.5-times lower than in protein coding sequences [118] and from these polymorphisms, only 1/10 or less are located in the seed region [60,118–120]. As expected, miR SNPs in the seed region would ultimately result in the regulation of a completely different set of mRNA targets. An increasing number of epidemiological reports have now linked several of these miR SNPs to pathology and, in particular, cancer susceptibility. miR-146a-3p and miR-499-3p, for instance, have so far been associated with the largest variety of cancer pathologies affecting all organ systems (for review see [120,121]).

Recent advances in high-throughput small RNA sequencing technologies have revealed novel post-transcriptional processing mechanisms that increase mature miR sequence heterogeneity from single genomic locus in cells. It is estimated that as many as 90% of miRs are presented with some sort of modification mainly in the form of trimming and/or nucleotide addition in the 3' terminus [122].

Thus far, four mechanisms that generate functionally distinct miR isoforms, annotated as isomiRs, have been identified [123]. These are RNA editing, inexact Drosha and Dicer processing, exonuclease ribonucleotide trimming and template-independent ribonucleotide addition.

RNA editing is a chemical alteration in the primary nucleotide sequence of double-stranded RNAs. The most common RNA editing modification involves the hydrolytic deamination of adenosine-to-inosine (A-to-I) catalyzed by the adenosine deaminase acting on RNA (ADAR) enzymes [124]. Because inosine preferentially base pairs with cytidine, this conversion is equivalent to an adenosine to guanosine change. Although earlier reports identified widespread A-to-I editing in pri-miRs, more recent studies have revealed that RNA editing is rather rare for mature miRs [125,126]. A comprehensive profiling of human RNA editome revealed only 44 edited miR sites of which 11 were in the seed region [127]. This indicates that miRs exhibit low frequency of editing and that the primary biological function of miR editing in animals is the regulation of the miR maturation pathway, rather than the specificity of miR targeting [125]. Nevertheless, editing of mature miRs at seed region, such as for the most thoroughly studied mir-376, resulted in changes in the targeting profile that subsequently altered biological function in a tissue-specific manner [126] promoting carcinogenesis [128]. For another miR, mir-142, pri-miR editing resulted in impaired Drosha processing and enhanced degradation by the specific I-U nuclease Tudor-SN [129]. Recently, the adenosine deaminase ADAR1 was shown to differentiate from its deaminase activity and participate in RNAi when in heteroduplex with Dicer. Hence, when in complex with Dicer, it increased the rate of pre-miR cleave and facilitated miRISC loading of mature miRs, while in homodimer form, it mediated RNA editing [130].

Multiple isomiRs with various 5' and/or 3' ends are thought to be the result of sloppy Drosha and Dicer excision [131,132]. More recently, mammalian TRBP and its *Drosophila* ortholog Loqs have been shown to fine-tune Dicer cleavage sites for a subset of miRs generating longer miR isoforms by one nucleotide at either 5' or 3' ends [133,134]. In addition, it was shown that the hairpin loop and stem structure of the pre-miR affected Dicer-TRBP processing with different sensitivity compared with Dicer alone. The authors proposed that TRBP might induce a Dicer conformational change influencing Dicer substrate specificity and kinetics [134].

Post-Dicer processing by exoribonucleases modulates 3' shortening in miRs. Nibbler, a 3' to 5' exoribonuclease has been shown to trim Ago-bound miRs in *Drosophila*; depletion of Nibbler resulted in the loss of about a quarter of 3' isomiRs; unexpectedly, Han *et al.* also found that miRs are frequently produced by Dicer as intermediates that are longer than ~22 nt, and are subsequently trimmed to appropriate size by exoribonucleases [129,135]. It remains to be seen whether similar mechanisms exist in mammals.

Besides nucleotide excisions, post-Dicer 3' additions are widespread and conserved. These are mediated by several nucleotidyl transferases that catalyze the addition of ribonucleotides, most often adenine and uridine, to the ends of mature miR molecules [136,137]. Interestingly, these isomiRs are differentially expressed across development and different tissues. For instance, adenines are highly abundant in early *Drosophila* development, while a subset of miRs with uridines is expressed in adult tissues [138]. With respect to function, these 3' ribonucleotide additions have been shown to alter (enhance or lower) miR stability in some cases [138,139] and/or effectiveness in others [140,141]. However, the authors concluded that these effects are likely to be restricted to only a small subset of isomiRs in animals [140]. Like for differentially expressed splice mRNA variants, several isomiRs

with 3' additions have been associated with human diseases. Thus far, significant alterations have been reported in cancer, Huntington's disease and pre-eclampsia [122,142,143].

4.4. ceRNAs and miR Degrading Enzymes

Recently, a new model of post-transcriptional regulation has emerged in which RNA targets are not merely passive substrates of miR repression, but cross-talk with each other in distinct networks by competing for shared miRs. Such competing endogenous RNAs (ceRNAs) ultimately determine mature miR availability and function within cells (reviewed in [144,145]). This reverse reasoning compels a redefinition of the idea that miRs stand at the top of mRNA networks to regulate protein output, by considering that any RNA that shares same target sequences actively regulate each other and miRs through direct competition for miR binding. Initial reports that provided proof of principle to this concept have shown that exogenous overexpression of 3' UTR sequences alone titrated cellular miR abundance and inactivated miR functions by freeing target mRNAs from repression [83,146,147]. Subsequently, it was shown that tenths of protein-coding mRNAs that share multiple miR target sites with dose-sensitive phosphatase and tensin homolog (PTEN) act as decoys to modulate PTEN levels [85,148]. An implication of these studies is that any RNA with miR target sites can potentially function as ceRNA. Thus, long non-coding RNAs (lncRNAs), due to their length, may be good candidates for sequestering miRs within cells. Hence, muscle-specific lncRNA, linc-MD1, was shown to sponge out two miRs to regulate the expression of transcription factors that activate muscle-specific gene expression [149]. Similarly, the PTENP1 pseudogene that is highly homologous to PTEN regulated cellular levels of PTEN (and the reverse) by sponging out common miRs [150]. Very recently, the repertoire of ceRNAs has been expanded by the identification of a new subclass of circular RNAs (circRNAs) [151,152]. Like other ceRNAs, these circRNAs serve as miR reservoirs. Distinctly, however, circRNAs may have multiple binding sites for specific miRs and therefore, are dedicated to sequestering particular miRs. Furthermore, being circulized, they possess enhanced stability by avoiding RNA exoribonuclease enzymes that act on 3' and 5' RNA ends and hence, maintain their effects for longer. An extreme case is characterized by human circRNA, ciRS-7, that harbors 74 mismatched mir-7 seed matches of which 63 are conserved in at least one other species [152]. This circRNA acts as a mir-7 sponge; it is resistant to miR-dependent destabilization and strongly suppresses miR-7 activity [151].

Exoribonucleases and the exosome have also been implicated in miR turnover. Using microarrays, Bail *et al*. have found that most miRs are remarkably stable (half-life over 8hrs), but some, including miR-382, were short-lived and were degraded to a modest extent (1.5-fold) by XRN1, a 5' to 3' exoribonuclease, and exosome, but not by XRN2 [153]. Moreover, overexpression of polynucleotide phosphorylase hPNPase(old-35), a 3' to 5' exoribonuclease, resulted in the downregulation of specific mature miRs in human melanoma cells without affecting their pri- or pre- miR levels [154].

5. Non-Canonical miR Activities

A relatively small number of studies have demonstrated that miRs can stimulate gene expression along their assigned repressive roles. These reports indicated that miR-mediated effects via gene

promoters, extracellular receptors and 3' or 5' UTRs can be selective and controlled, ordained by either the miR sequence, associated proteins and/or cellular context.

5.1. Promoter Activation

Earlier studies have shown that exogenous application of small duplex RNAs, that are complementary to promoters, activate gene expression just like proteins and hormones, a phenomenon referred to as RNA activation (RNAa) [155,156]. Soon later, Dahiya's group discovered mir-373 target sites in the promoters of e-cadherin and cold shock domain containing protein C2 (CSDC2). miR-373 overexpression readily induced transcription of these two genes and this concurrent induction required mir-373 target sites in both promoters [157]. Subsequently, they showed that mir-205 sites are present in the promoter of interleukin (IL) tumor suppressor genes IL-24 and IL-32 and, similar to mir-373, mir-205 induced gene expression [157,158].

5.2. Target Activation

Several reports have shown that miRs can induce translation by binding to 5' or 3' UTR. In the brain, a target sequence of mir-346 was found in the 5' UTR of a splice variant of receptor-interacting protein 140 (RIP140). Gain- and loss- of-function studies established that mir-346 elevated RIP140 protein levels by facilitating association of its mRNA with the polysome fraction. Furthermore, the activity of the mir-346 did not require Ago2 indicating that other RNPs in complex with the miR or different RIP140 mRNA conformation induced by the miR mediated the effect [159]. In another study, mir-145 was shown to regulate smooth muscle cell fate and plasticity via upregulation of myocardin (*Myocd*). *Myocd* bears mir-145 sites in 3' UTR and mir-145 expression specifically upregulated luciferase expression by 150-fold; at the same time other mir-145 targets were repressed. It remains to be seen whether miR-145 interferes with binding of a destabilizing RBP to 3' UTR [160]. Along this, miR-466l, a miR discovered in mouse embryonic stem cells, upregulated IL-10 expression in TLR-triggered macrophages by antagonizing the RBP tristetraprolin (TTP)-mediated IL-10 mRNA degradation [161].

5.3. Receptors' Ligands

Members of the Toll-like receptor (TLR) family, mouse TLR7 and human TLR8, expressed on dendritic cells and B lymphocytes, physiologically recognize and bind ~20 nt viral single-stranded RNAs leading to their activation [162,163]. Because miRs are secreted in exosomes and are of similar size, it was predicted that they may also serve as TLR7/8 ligands. Indeed, Fabbri *et al*. identified that the tumor-secreted mir-21 and mir-29a were ligands for TLR7/8 and were capable of triggering a TLR-mediated prometastatic inflammatory response [164].

6. Conclusions

Over the past years, significant advances have been made into understanding how miRs interact with their RNA targets, and several key features, such as base-pair complementarity, local context factors and de/stabilization signals have been identified and finely analyzed as a result. The ultimate goal of all these studies has been to predict miR function through the identification of their targets. More recent analyses, however, demonstrated that local mRNA determinants could only explain a fraction of the miR repression activity and system level factors such as isomiRs, RBPs, and ceRNAs have been brought into attention. The very recent discovery that miRs can both regulate and be regulated by their RNA targets has presented a completely new twist into understanding the role of miRs in development and disease. It remains to be seen how miRs and RNA targets communicate using the miR nt sequence as a "language" to deliver large-scale concerted instructions in cells.

Acknowledgments

The Author wishes to thank Maria Paschou for assistance with the preparation of the figures. Work in the author's laboratory is funded by grants from the Greek General Secretariat for Research and Development (GSRT), Ministry of Education.

Conflict of Interest

The author declares no conflict of interest.

References

1. Kozomara, A.; Griffiths-Jones, S. miRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **2011**, *39*, D152–D157.
2. Chiang, H.R.; Schoenfeld, L.W.; Ruby, J.G.; Auyeung, V.C.; Spies, N.; Baek, D.; Johnston, W.K.; Russ, C.; Luo, S.; Babiarz, J.E.; *et al*. Mammalian microRNAs: Experimental evaluation of novel and previously annotated genes. *Genes Dev.* **2010**, *24*, 992–1009.
3. Meunier, J.; Lemoine, F.; Soumillon, M.; Liechti, A.; Weier, M.; Guschanski, K.; Hu, H.; Khaitovich, P.; Kaessmann, H. Birth and expression evolution of mammalian microRNA genes. *Genome Res.* **2013**, *23*, 34–45.
4. Friedman, R.C.; Farh, K.K.; Burge, C.B.; Bartel, D.P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **2009**, *19*, 92–105.
5. Beuvink, I.; Kolb, F.A.; Budach, W.; Garnier, A.; Lange, J.; Natt, F.; Dengler, U.; Hall, J.; Filipowicz, W.; Weiler, J. A novel microarray approach reveals new tissue-specific signatures of known and predicted mammalian microRNAs. *Nucleic Acids Res.* **2007**, *35*, e52.
6. Landgraf, P.; Rusu, M.; Sheridan, R.; Sewer, A.; Iovino, N.; Aravin, A.; Pfeffer, S.; Rice, A.; Kamphorst, A.O.; Landthaler, M.; *et al*. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **2007**, *129*, 1401–1414.
7. Liang, Y.; Ridzon, D.; Wong, L.; Chen, C. Characterization of microRNA expression profiles in normal human tissues. *BMC Genomics* **2007**, *8*, 166.

8. Moreau, M.P.; Bruse, S.E.; Jornsten, R.; Liu, Y.; Brzustowicz, L.M. Chronological changes in microRNA expression in the developing human brain. *PLoS One* **2013**, *8*, e60480.
9. Lu, J.; Getz, G.; Miska, E.A.; Alvarez-Saavedra, E.; Lamb, J.; Peck, D.; Sweet-Cordero, A.; Ebert, B.L.; Mak, R.H.; Ferrando, A.A.; *et al*. microRNA expression profiles classify human cancers. *Nature* **2005**, *435*, 834–838.
10. Aldrich, B.T.; Frakes, E.P.; Kasuya, J.; Hammond, D.L.; Kitamoto, T. Changes in expression of sensory organ-specific microRNAs in rat dorsal root ganglia in association with mechanical hypersensitivity induced by spinal nerve ligation. *Neuroscience* **2009**, *164*, 711–723.
11. Elmen, J.; Lindow, M.; Schutz, S.; Lawrence, M.; Petri, A.; Obad, S.; Lindholm, M.; Hedtjarn, M.; Hansen, H.F.; Berger, U.; *et al*. LNA-mediated microRNA silencing in non-human primates. *Nature* **2008**, *452*, 896–899.
12. Elmen, J.; Lindow, M.; Silahtaroglu, A.; Bak, M.; Christensen, M.; Lind-Thomsen, A.; Hedtjarn, M.; Hansen, J.B.; Hansen, H.F.; Straarup, E.M.; *et al*. Antagonism of microRNA-122 in mice by systemically administered LNA-antimiR leads to up-regulation of a large set of predicted target mRNAs in the liver. *Nucleic Acids Res.* **2008**, *36*, 1153–1162.
13. Krutzfeldt, J.; Rajewsky, N.; Braich, R.; Rajeev, K.G.; Tuschl, T.; Manoharan, M.; Stoffel, M. Silencing of microRNAs *in vivo* with "antagomirs". *Nature* **2005**, *438*, 685–689.
14. Corcoran, D.L.; Pandit, K.V.; Gordon, B.; Bhattacharjee, A.; Kaminski, N.; Benos, P.V. Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data. *PLoS One* **2009**, *4*, e5279.
15. Marson, A.; Levine, S.S.; Cole, M.F.; Frampton, G.M.; Brambrink, T.; Johnstone, S.; Guenther, M.G.; Johnston, W.K.; Wernig, M.; Newman, J.; *et al*. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **2008**, *134*, 521–533.
16. Ozsolak, F.; Poling, L.L.; Wang, Z.; Liu, H.; Liu, X.S.; Roeder, R.G.; Zhang, X.; Song, J.S.; Fisher, D.E. Chromatin structure analyses identify miRNA promoters. *Genes Dev.* **2008**, *22*, 3172–3183.
17. Saini, H.K.; Griffiths-Jones, S.; Enright, A.J. Genomic analysis of human microRNA transcripts. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 17719–17724.
18. Wang, X.; Xuan, Z.; Zhao, X.; Li, Y.; Zhang, M.Q. High-resolution human core-promoter prediction with CoreBoost_HM. *Genome Res.* **2009**, *19*, 266–275.
19. Zhou, X.; Ruan, J.; Wang, G.; Zhang, W. Characterization and identification of microRNA core promoters in four model species. *PLoS Comput. Biol.* **2007**, *3*, e37.
20. Rodriguez, A.; Griffiths-Jones, S.; Ashurst, J.L.; Bradley, A. Identification of mammalian microRNA host genes and transcription units. *Genome Res.* **2004**, *14*, 1902–1910.
21. Han, J.; Lee, Y.; Yeom, K.H.; Nam, J.W.; Heo, I.; Rhee, J.K.; Sohn, S.Y.; Cho, Y.; Zhang, B.T.; Kim, V.N. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **2006**, *125*, 887–901.
22. Zeng, Y.; Cullen, B.R. Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. *J. Biol. Chem.* **2005**, *280*, 27595–27603.
23. Basyuk, E.; Suavet, F.; Doglio, A.; Bordonne, R.; Bertrand, E. Human let-7 stem-loop precursors harbor features of RNase III cleavage products. *Nucleic Acids Res.* **2003**, *31*, 6593–6597.

24. Denli, A.M.; Tops, B.B.; Plasterk, R.H.; Ketting, R.F.; Hannon, G.J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **2004**, *432*, 231–235.
25. Gregory, R.I.; Yan, K.P.; Amuthan, G.; Chendrimada, T.; Doratotaj, B.; Cooch, N.; Shiekhattar, R. The Microprocessor complex mediates the genesis of microRNAs. *Nature* **2004**, *432*, 235–240.
26. Lee, Y.; Ahn, C.; Han, J.; Choi, H.; Kim, J.; Yim, J.; Lee, J.; Provost, P.; Radmark, O.; Kim, S.; *et al*. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **2003**, *425*, 415–419.
27. Bernstein, E.; Caudy, A.A.; Hammond, S.M.; Hannon, G.J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **2001**, *409*, 363–366.
28. Chakravarthy, S.; Sternberg, S.H.; Kellenberger, C.A.; Doudna, J.A. Substrate-specific kinetics of Dicer-catalyzed RNA processing. *J. Mol. Biol.* **2010**, *404*, 392–402.
29. Haase, A.D.; Jaskiewicz, L.; Zhang, H.; Laine, S.; Sack, R.; Gatignol, A.; Filipowicz, W. TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep.* **2005**, *6*, 961–967.
30. Hutvagner, G.; McLachlan, J.; Pasquinelli, A.E.; Balint, E.; Tuschl, T.; Zamore, P.D. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **2001**, *293*, 834–838.
31. Lee, H.Y.; Zhou, K.; Smith, A.M.; Noland, C.L.; Doudna, J.A. Differential roles of human Dicer-binding proteins TRBP and PACT in small RNA processing. *Nucleic Acids Res.* **2013**, doi:10.1093/nar/gkt361.
32. Lee, Y.; Hur, I.; Park, S.Y.; Kim, Y.K.; Suh, M.R.; Kim, V.N. The role of PACT in the RNA silencing pathway. *EMBO J.* **2006**, *25*, 522–532.
33. MacRae, I.J.; Zhou, K.; Doudna, J.A. Structural determinants of RNA recognition and cleavage by Dicer. *Nat. Struct. Mol. Biol.* **2007**, *14*, 934–940.
34. Chendrimada, T.P.; Gregory, R.I.; Kumaraswamy, E.; Norman, J.; Cooch, N.; Nishikura, K.; Shiekhattar, R. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **2005**, *436*, 740–744.
35. Khvorova, A.; Reynolds, A.; Jayasena, S.D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **2003**, *115*, 209–216.
36. Krol, J.; Sobczak, K.; Wilczynska, U.; Drath, M.; Jasinska, A.; Kaczynska, D.; Krzyzosiak, W.J. Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. *J. Biol. Chem.* **2004**, *279*, 42230–42239.
37. Schwarz, D.S.; Hutvagner, G.; Du, T.; Xu, Z.; Aronin, N.; Zamore, P.D. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **2003**, *115*, 199–208.
38. Czech, B.; Zhou, R.; Erlich, Y.; Brennecke, J.; Binari, R.; Villalta, C.; Gordon, A.; Perrimon, N.; Hannon, G.J. Hierarchical rules for Argonaute loading in *Drosophila*. *Mol. Cell* **2009**, *36*, 445–456.
39. Hu, H.Y.; Yan, Z.; Xu, Y.; Hu, H.; Menzel, C.; Zhou, Y.H.; Chen, W.; Khaitovich, P. Sequence features associated with microRNA strand selection in humans and flies. *BMC Genomics* **2009**, *10*, 413.
40. Noland, C.L.; Doudna, J.A. Multiple sensors ensure guide strand selection in human RNAi pathways. *RNA* **2013**, *19*, 639–648.

41. Okamura, K.; Liu, N.; Lai, E.C. Distinct mechanisms for microRNA strand selection by *Drosophila* Argonautes. *Mol. Cell* **2009**, *36*, 431–444.
42. Ameres, S.L.; Martinez, J.; Schroeder, R. Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* **2007**, *130*, 101–112.
43. Fabian, M.R.; Sonenberg, N. The mechanics of miRNA-mediated gene silencing: A look under the hood of miRISC. *Nat. Struct. Mol. Biol.* **2012**, *19*, 586–593.
44. Huntzinger, E.; Kuzuoglu-Ozturk, D.; Braun, J.E.; Eulalio, A.; Wohlbold, L.; Izaurralde, E. The interactions of GW182 proteins with PABP and deadenylases are required for both translational repression and degradation of miRNA targets. *Nucleic Acids Res.* **2013**, *41*, 978–994.
45. Zekri, L.; Kuzuoglu-Ozturk, D.; Izaurralde, E. GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. *EMBO J.* **2013**, *32*, 1052–1065.
46. Brennecke, J.; Stark, A.; Russell, R.B.; Cohen, S.M. Principles of microRNA-target recognition. *PLoS Biol.* **2005**, *3*, e85.
47. Krek, A.; Grun, D.; Poy, M.N.; Wolf, R.; Rosenberg, L.; Epstein, E.J.; MacMenamin, P.; da Piedade, I.; Gunsalus, K.C.; Stoffel, M.; *et al*. Combinatorial microRNA target predictions. *Nat. Genet.* **2005**, *37*, 495–500.
48. Lewis, B.P.; Burge, C.B.; Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **2005**, *120*, 15–20.
49. Lewis, B.P.; Shih, I.H.; Jones-Rhoades, M.W.; Bartel, D.P.; Burge, C.B. Prediction of mammalian microRNA targets. *Cell* **2003**, *115*, 787–798.
50. Elkayam, E.; Kuhn, C.D.; Tocilj, A.; Haase, A.D.; Greene, E.M.; Hannon, G.J.; Joshua-Tor, L. The structure of human argonaute-2 in complex with miR-20a. *Cell* **2012**, *150*, 100–110.
51. Schirle, N.T.; MacRae, I.J. The crystal structure of human Argonaute 2. *Science* **2012**, *336*, 1037–1040.
52. Wang, Y.; Sheng, G.; Juranek, S.; Tuschl, T.; Patel, D.J. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **2008**, *456*, 209–213.
53. Chi, S.W.; Hannon, G.J.; Darnell, R.B. An alternative mode of microRNA target recognition. *Nat. Struct. Mol. Biol.* **2012**, *19*, 321–327.
54. Miranda, K.C.; Huynh, T.; Tay, Y.; Ang, Y.S.; Tam, W.L.; Thomson, A.M.; Lim, B.; Rigoutsos, I. A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* **2006**, *126*, 1203–1217.
55. Vella, M.C.; Choi, E.Y.; Lin, S.Y.; Reinert, K.; Slack, F.J. The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3' UTR. *Genes Dev.* **2004**, *18*, 132–137.
56. Reinhart, B.J.; Slack, F.J.; Basson, M.; Pasquinelli, A.E.; Bettinger, J.C.; Rougvie, A.E.; Horvitz, H.R.; Ruvkun, G. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **2000**, *403*, 901–906.
57. Shin, C.; Nam, J.W.; Farh, K.K.; Chiang, H.R.; Shkumatava, A.; Bartel, D.P. Expanding the microRNA targeting code: functional sites with centered pairing. *Mol. Cell* **2010**, *38*, 789–802.
58. Lal, A.; Navarro, F.; Maher, C.A.; Maliszewski, L.E.; Yan, N.; O'Day, E.; Chowdhury, D.; Dykxhoorn, D.M.; Tsai, P.; Hofmann, O.; *et al*. miR-24 inhibits cell proliferation by targeting E2F2, MYC, and other cell-cycle genes via binding to "seedless" 3' UTR microRNA recognition elements. *Mol. Cell* **2009**, *35*, 610–625.

59. Tay, Y.; Zhang, J.; Thomson, A.M.; Lim, B.; Rigoutsos, I. microRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* **2008**, *455*, 1124–1128.
60. Didiano, D.; Hobert, O. Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nat. Struct. Mol. Biol.* **2006**, *13*, 849–851.
61. Grimson, A.; Farh, K.K.; Johnston, W.K.; Garrett-Engele, P.; Lim, L.P.; Bartel, D.P. microRNA targeting specificity in mammals: Determinants beyond seed pairing. *Mol. Cell* **2007**, *27*, 91–105.
62. Hon, L.S.; Zhang, Z. The roles of binding site arrangement and combinatorial targeting in microRNA repression of gene expression. *Genome Biol.* **2007**, *8*, R166.
63. Saetrom, P.; Heale, B.S.; Snove, O., Jr.; Aagaard, L.; Alluin, J.; Rossi, J.J. Distance constraints between microRNA target sites dictate efficacy and cooperativity. *Nucleic Acids Res.* **2007**, *35*, 2333–2342.
64. Doxakis, E. Post-transcriptional regulation of alpha-synuclein expression by mir-7 and mir-153. *J. Biol. Chem.* **2010**, *285*, 12726–12734.
65. Paschou, M.; Doxakis, E. Neurofibromin 1 is a miRNA target in neurons. *PLoS One* **2012**, *7*, e46773.
66. Chi, S.W.; Zang, J.B.; Mele, A.; Darnell, R.B. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* **2009**, *460*, 479–486.
67. Hafner, M.; Landthaler, M.; Burger, L.; Khorshid, M.; Hausser, J.; Berninger, P.; Rothballer, A.; Ascano, M., Jr.; Jungkamp, A.C.; Munschauer, M.; *et al*. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **2010**, *141*, 129–141.
68. Leung, A.K.; Vyas, S.; Rood, J.E.; Bhutkar, A.; Sharp, P.A.; Chang, P. Poly(ADP-ribose) regulates stress responses and microRNA activity in the cytoplasm. *Mol. Cell* **2011**, *42*, 489–499.
69. Rigoutsos, I. New tricks for animal microRNAS: Targeting of amino acid coding regions at conserved and nonconserved sites. *Cancer Res.* **2009**, *69*, 3245–3248.
70. Zisoulis, D.G.; Lovci, M.T.; Wilbert, M.L.; Hutt, K.R.; Liang, T.Y.; Pasquinelli, A.E.; Yeo, G.W. Comprehensive discovery of endogenous Argonaute binding sites in Caenorhabditis elegans. *Nat. Struct. Mol. Biol.* **2010**, *17*, 173–179.
71. Baek, D.; Villen, J.; Shin, C.; Camargo, F.D.; Gygi, S.P.; Bartel, D.P. The impact of microRNAs on protein output. *Nature* **2008**, *455*, 64–71.
72. Selbach, M.; Schwanhausser, B.; Thierfelder, N.; Fang, Z.; Khanin, R.; Rajewsky, N. Widespread changes in protein synthesis induced by microRNAs. *Nature* **2008**, *455*, 58–63.
73. Hausser, J.; Syed, A.P.; Bilen, B.; Zavolan, M. Analysis of CDS-located miRNA target sites suggests that they can effectively inhibit translation. *Genome Res.* **2013**, *23*, 604–615.
74. Schnall-Levin, M.; Rissland, O.S.; Johnston, W.K.; Perrimon, N.; Bartel, D.P.; Berger, B. Unusually effective microRNA targeting within repeat-rich coding regions of mammalian mRNAs. *Genome Res.* **2011**, *21*, 1395–1403.
75. Fang, Z.; Rajewsky, N. The impact of miRNA target sites in coding sequences and in 3' UTRs. *PLoS One* **2011**, *6*, e18067.
76. Wang, W.X.; Wilfred, B.R.; Xie, K.; Jennings, M.H.; Hu, Y.H.; Stromberg, A.J.; Nelson, P.T. Individual microRNAs (miRNAs) display distinct mRNA targeting "rules". *RNA Biol.* **2010**, *7*, 373–380.

77. Kertesz, M.; Iovino, N.; Unnerstall, U.; Gaul, U.; Segal, E. The role of site accessibility in microRNA target recognition. *Nat. Genet.* **2007**, *39*, 1278–1284.
78. Long, D.; Lee, R.; Williams, P.; Chan, C.Y.; Ambros, V.; Ding, Y. Potent effect of target structure on microRNA function. *Nat. Struct. Mol. Biol.* **2007**, *14*, 287–294.
79. Paschou, M.; Paraskevopoulou, M.D.; Vlachos, I.S.; Koukouraki, P.; Hatzigeorgiou, A.G.; Doxakis, E. miRNA regulons associated with synaptic function. *PLoS One* **2012**, *7*, e46189.
80. Bergauer, T.; Krueger, U.; Lader, E.; Pilk, S.; Wolter, I.; Bielke, W. Analysis of putative miRNA binding sites and mRNA 3' ends as targets for siRNA-mediated gene knockdown. *Oligonucleotides* **2009**, *19*, 41–52.
81. Ji, Z.; Lee, J.Y.; Pan, Z.; Jiang, B.; Tian, B. Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 7028–7033.
82. Sandberg, R.; Neilson, J.R.; Sarma, A.; Sharp, P.A.; Burge, C.B. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **2008**, *320*, 1643–1647.
83. Arvey, A.; Larsson, E.; Sander, C.; Leslie, C.S.; Marks, D.S. Target mRNA abundance dilutes microRNA and siRNA activity. *Mol. Syst. Biol.* **2010**, *6*, 363.
84. Karreth, F.A.; Tay, Y.; Perna, D.; Ala, U.; Tan, S.M.; Rust, A.G.; de Nicola, G.; Webster, K.A.; Weiss, D.; Perez-Mancera, P.A.; *et al*. *In vivo* identification of tumor-suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. *Cell* **2011**, *147*, 382–395.
85. Sumazin, P.; Yang, X.; Chiu, H.S.; Chung, W.J.; Iyer, A.; Llobet-Navas, D.; Rajbhandari, P.; Bansal, M.; Guarnieri, P.; Silva, J.; *et al*. An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma. *Cell* **2011**, *147*, 370–381.
86. Ameres, S.L.; Horwich, M.D.; Hung, J.H.; Xu, J.; Ghildiyal, M.; Weng, Z.; Zamore, P.D. Target RNA-directed trimming and tailing of small silencing RNAs. *Science* **2010**, *328*, 1534–1539.
87. Rybak, A.; Fuchs, H.; Hadian, K.; Smirnova, L.; Wulczyn, E.A.; Michel, G.; Nitsch, R.; Krappmann, D.; Wulczyn, F.G. The let-7 target gene mouse lin-41 is a stem cell specific E3 ubiquitin ligase for the miRNA pathway protein Ago2. *Nat. Cell Biol.* **2009**, *11*, 1411–1420.
88. Pullmann, R., Jr.; Kim, H.H.; Abdelmohsen, K.; Lal, A.; Martindale, J.L.; Yang, X.; Gorospe, M. Analysis of turnover and translation regulatory RNA-binding protein expression through binding to cognate mRNAs. *Mol. Cell. Biol.* **2007**, *27*, 6265–6278.
89. Jacobsen, A.; Wen, J.; Marks, D.S.; Krogh, A. Signatures of RNA binding proteins globally coupled to effective microRNA target sites. *Genome Res.* **2010**, *20*, 1010–1019.
90. Larsson, E.; Sander, C.; Marks, D. mRNA turnover rate limits siRNA and microRNA efficacy. *Mol. Syst. Biol.* **2010**, *6*, 433.
91. Lebedeva, S.; Jens, M.; Theil, K.; Schwanhausser, B.; Selbach, M.; Landthaler, M.; Rajewsky, N. Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. *Mol. Cell* **2011**, *43*, 340–352.
92. Mukherjee, N.; Corcoran, D.L.; Nusbaum, J.D.; Reid, D.W.; Georgiev, S.; Hafner, M.; Ascano, M., Jr.; Tuschl, T.; Ohler, U.; Keene, J.D. Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability. *Mol. Cell* **2011**, *43*, 327–339.

93. Srikantan, S.; Tominaga, K.; Gorospe, M. Functional interplay between RNA-binding protein HuR and microRNAs. *Curr. Protein Pept. Sci.* **2012**, *13*, 372–379.
94. Abdelmohsen, K.; Srikantan, S.; Kuwano, Y.; Gorospe, M. miR-519 reduces cell proliferation by lowering RNA-binding protein HuR levels. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20297–20302.
95. Guo, X.; Wu, Y.; Hartley, R.S. microRNA-125a represses cell growth by targeting HuR in breast cancer. *RNA Biol.* **2009**, *6*, 575–583.
96. Young, L.E.; Moore, A.E.; Sokol, L.; Meisner-Kober, N.; Dixon, D.A. The mRNA stability factor HuR inhibits microRNA-16 targeting of COX-2. *Mol. Cancer Res.* **2012**, *10*, 167–180.
97. Borchert, G.M.; Lanier, W.; Davidson, B.L. RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.* **2006**, *13*, 1097–1101.
98. Dieci, G.; Fiorino, G.; Castelnuovo, M.; Teichmann, M.; Pagano, A. The expanding RNA polymerase III transcriptome. *Trends Genet.* **2007**, *23*, 614–622.
99. Lee, Y.; Kim, M.; Han, J.; Yeom, K.H.; Lee, S.; Baek, S.H.; Kim, V.N. microRNA genes are transcribed by RNA polymerase II. *EMBO J.* **2004**, *23*, 4051–4060.
100. Carlson, D.P.; Ross, J. Human beta-globin promoter and coding sequences transcribed by RNA polymerase III. *Cell* **1983**, *34*, 857–864.
101. Chung, J.; Sussman, D.J.; Zeller, R.; Leder, P. The c-myc gene encodes superimposed RNA polymerase II and III promoters. *Cell* **1987**, *51*, 1001–1008.
102. Listerman, I.; Bledau, A.S.; Grishina, I.; Neugebauer, K.M. Extragenic accumulation of RNA polymerase II enhances transcription by RNA polymerase III. *PLoS Genet.* **2007**, *3*, e212.
103. Felton-Edkins, Z.A.; Kenneth, N.S.; Brown, T.R.; Daly, N.L.; Gomez-Roman, N.; Grandori, C.; Eisenman, R.N.; White, R.J. Direct regulation of RNA polymerase III transcription by RB, p53 and c-Myc. *Cell Cycle* **2003**, *2*, 181–184.
104. Baskerville, S.; Bartel, D.P. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* **2005**, *11*, 241–247.
105. Fujita, S.; Ito, T.; Mizutani, T.; Minoguchi, S.; Yamamichi, N.; Sakurai, K.; Iba, H. miR-21 gene expression triggered by AP-1 is sustained through a double-negative feedback mechanism. *J. Mol. Biol.* **2008**, *378*, 492–504.
106. He, L.; Thomson, J.M.; Hemann, M.T.; Hernando-Monge, E.; Mu, D.; Goodson, S.; Powers, S.; Cordon-Cardo, C.; Lowe, S.W.; Hannon, G.J.; *et al*. A microRNA polycistron as a potential human oncogene. *Nature* **2005**, *435*, 828–833.
107. O'Donnell, K.A.; Wentzel, E.A.; Zeller, K.I.; Dang, C.V.; Mendell, J.T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **2005**, *435*, 839–843.
108. Monteys, A.M.; Spengler, R.M.; Wan, J.; Tecedor, L.; Lennox, K.A.; Xing, Y.; Davidson, B.L. Structure and activity of putative intronic miRNA promoters. *RNA* **2010**, *16*, 495–505.
109. Yuan, X.; Liu, C.; Yang, P.; He, S.; Liao, Q.; Kang, S.; Zhao, Y. Clustered microRNAs' coordination in regulating protein-protein interaction network. *BMC Syst. Biol.* **2009**, *3*, 65.
110. Gusev, Y. Computational methods for analysis of cellular functions and pathways collectively targeted by differentially expressed microRNA. *Methods* **2008**, *44*, 61–72.
111. Davis, B.N.; Hilyard, A.C.; Lagna, G.; Hata, A. SMAD proteins control DROSHA-mediated microRNA maturation. *Nature* **2008**, *454*, 56–61.

112. Davis, B.N.; Hilyard, A.C.; Nguyen, P.H.; Lagna, G.; Hata, A. Smad proteins bind a conserved RNA sequence to promote microRNA maturation by Drosha. *Mol. Cell* **2010**, *39*, 373–384.
113. Suzuki, H.I.; Yamagata, K.; Sugimoto, K.; Iwamoto, T.; Kato, S.; Miyazono, K. Modulation of microRNA processing by p53. *Nature* **2009**, *460*, 529–533.
114. Guil, S.; Caceres, J.F. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat. Struct. Mol. Biol.* **2007**, *14*, 591–596.
115. Michlewski, G.; Caceres, J.F. Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. *Nat. Struct. Mol. Biol.* **2010**, *17*, 1011–1018.
116. Trabucchi, M.; Briata, P.; Garcia-Mayoral, M.; Haase, A.D.; Filipowicz, W.; Ramos, A.; Gherzi, R.; Rosenfeld, M.G. The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature* **2009**, *459*, 1010–1014.
117. Chen, K.; Rajewsky, N. Natural selection on human microRNA binding sites inferred from SNP data. *Nat. Genet.* **2006**, *38*, 1452–1456.
118. Muinos-Gimeno, M.; Montfort, M.; Bayes, M.; Estivill, X.; Espinosa-Parrilla, Y. Design and evaluation of a panel of single-nucleotide polymorphisms in microRNA genomic regions for association studies in human disease. *Eur. J. Hum. Genet.* **2010**, *18*, 218–226.
119. Saunders, M.A.; Liang, H.; Li, W.H. Human polymorphism at microRNAs and microRNA target sites. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3300–3305.
120. Zorc, M.; Skok, D.J.; Godnic, I.; Calin, G.A.; Horvat, S.; Jiang, Z.; Dovc, P.; Kunej, T. Catalog of microRNA seed polymorphisms in vertebrates. *PLoS One* **2012**, *7*, e30737.
121. Ryan, B.M.; Robles, A.I.; Harris, C.C. Genetic variation in microRNA networks: The implications for cancer research. *Nat. Rev. Cancer* **2010**, *10*, 389–402.
122. Marti, E.; Pantano, L.; Banez-Coronel, M.; Llorens, F.; Minones-Moyano, E.; Porta, S.; Sumoy, L.; Ferrer, I.; Estivill, X. A myriad of miRNA variants in control and Huntington's disease brain regions detected by massively parallel sequencing. *Nucleic Acids Res.* **2010**, *38*, 7219–7235.
123. Morin, R.D.; O'Connor, M.D.; Griffith, M.; Kuchenbauer, F.; Delaney, A.; Prabhu, A.L.; Zhao, Y.; McDonald, H.; Zeng, T.; Hirst, M.; *et al*. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.* **2008**, *18*, 610–621.
124. Nishikura, K. Functions and regulation of RNA editing by ADAR deaminases. *Annu. Rev. Biochem.* **2010**, *79*, 321–349.
125. De Hoon, M.J.; Taft, R.J.; Hashimoto, T.; Kanamori-Katayama, M.; Kawaji, H.; Kawano, M.; Kishima, M.; Lassmann, T.; Faulkner, G.J.; Mattick, J.S.; *et al*. Cross-mapping and the identification of editing sites in mature microRNAs in high-throughput sequencing libraries. *Genome Res.* **2010**, *20*, 257–264.
126. Kawahara, Y.; Megraw, M.; Kreider, E.; Iizasa, H.; Valente, L.; Hatzigeorgiou, A.G.; Nishikura, K. Frequency and fate of microRNA editing in human brain. *Nucleic Acids Res.* **2008**, *36*, 5270–5280.
127. Peng, Z.; Cheng, Y.; Tan, B.C.; Kang, L.; Tian, Z.; Zhu, Y.; Zhang, W.; Liang, Y.; Hu, X.; Tan, X.; *et al*. Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. *Nat. Biotechnol.* **2012**, *30*, 253–260.

128. Choudhury, Y.; Tay, F.C.; Lam, D.H.; Sandanaraj, E.; Tang, C.; Ang, B.T.; Wang, S. Attenuated adenosine-to-inosine editing of microRNA-376a* promotes invasiveness of glioblastoma cells. *J. Clin. Invest.* **2012**, *122*, 4059–4076.
129. Liu, N.; Abe, M.; Sabin, L.R.; Hendriks, G.J.; Naqvi, A.S.; Yu, Z.; Cherry, S.; Bonini, N.M. The exoribonuclease Nibbler controls 3' end processing of microRNAs in *Drosophila*. *Curr. Biol.* **2011**, *21*, 1888–1893.
130. Ota, H.; Sakurai, M.; Gupta, R.; Valente, L.; Wulff, B.E.; Ariyoshi, K.; Iizasa, H.; Davuluri, R.V.; Nishikura, K. ADAR1 Forms a complex with Dicer to spromote microRNA processing and RNA-induced gene silencing. *Cell* **2013**, *153*, 575–589.
131. Ruby, J.G.; Jan, C.; Player, C.; Axtell, M.J.; Lee, W.; Nusbaum, C.; Ge, H.; Bartel, D.P. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* **2006**, *127*, 1193–1207.
132. Ruby, J.G.; Stark, A.; Johnston, W.K.; Kellis, M.; Bartel, D.P.; Lai, E.C. Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. *Genome Res.* **2007**, *17*, 1850–1864.
133. Fukunaga, R.; Han, B.W.; Hung, J.H.; Xu, J.; Weng, Z.; Zamore, P.D. Dicer partner proteins tune the length of mature miRNAs in flies and mammals. *Cell* **2012**, *151*, 533–546.
134. Lee, H.Y.; Doudna, J.A. TRBP alters human precursor microRNA processing *in vitro*. *RNA* **2012**, *18*, 2012–2019.
135. Han, B.W.; Hung, J.H.; Weng, Z.; Zamore, P.D.; Ameres, S.L. The 3'-to-5' exoribonuclease Nibbler shapes the 3' ends of microRNAs bound to *Drosophila* Argonaute1. *Curr. Biol.* **2011**, *21*, 1878–1887.
136. Martin, G.; Keller, W. RNA-specific ribonucleotidyl transferases. *RNA* **2007**, *13*, 1834–1849.
137. Wyman, S.K.; Knouf, E.C.; Parkin, R.K.; Fritz, B.R.; Lin, D.W.; Dennis, L.M.; Krouse, M.A.; Webster, P.J.; Tewari, M. Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. *Genome Res.* **2011**, *21*, 1450–1461.
138. Fernandez-Valverde, S.L.; Taft, R.J.; Mattick, J.S. Dynamic isomiR regulation in *Drosophila* development. *RNA* **2010**, *16*, 1881–1888.
139. Katoh, T.; Sakaguchi, Y.; Miyauchi, K.; Suzuki, T.; Kashiwabara, S.; Baba, T. Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes Dev.* **2009**, *23*, 433–438.
140. Burroughs, A.M.; Ando, Y.; de Hoon, M.J.; Tomaru, Y.; Nishibu, T.; Ukekawa, R.; Funakoshi, T.; Kurokawa, T.; Suzuki, H.; Hayashizaki, Y.; *et al*. A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. *Genome Res.* **2010**, *20*, 1398–1410.
141. Jones, M.R.; Quinton, L.J.; Blahna, M.T.; Neilson, J.R.; Fu, S.; Ivanov, A.R.; Wolf, D.A.; Mizgerd, J.P. Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat. Cell Biol.* **2009**, *11*, 1157–1163.
142. Guo, L.; Yang, Q.; Lu, J.; Li, H.; Ge, Q.; Gu, W.; Bai, Y.; Lu, Z. A comprehensive survey of miRNA repertoire and 3' addition events in the placentas of patients with pre-eclampsia from high-throughput sequencing. *PLoS One* **2011**, *6*, e21072.

143. Kuchenbauer, F.; Morin, R.D.; Argiropoulos, B.; Petriv, O.I.; Griffith, M.; Heuser, M.; Yung, E.; Piper, J.; Delaney, A.; Prabhu, A.L.; *et al*. In-depth characterization of the microRNA transcriptome in a leukemia progression model. *Genome Res.* **2008**, *18*, 1787–1797.
144. Cesana, M.; Daley, G.Q. Deciphering the rules of ceRNA networks. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 7112–7113.
145. Salmena, L.; Poliseno, L.; Tay, Y.; Kats, L.; Pandolfi, P.P. A ceRNA hypothesis: The Rosetta Stone of a hidden RNA language? *Cell* **2011**, *146*, 353–358.
146. Jeyapalan, Z.; Deng, Z.; Shatseva, T.; Fang, L.; He, C.; Yang, B.B. Expression of CD44 3'-untranslated region regulates endogenous microRNA functions in tumorigenesis and angiogenesis. *Nucleic Acids Res.* **2011**, *39*, 3026–3041.
147. Lee, D.Y.; Shatseva, T.; Jeyapalan, Z.; Du, W.W.; Deng, Z.; Yang, B.B. A 3'-untranslated region (3' UTR) induces organ adhesion by regulating miR-199a* functions. *PLoS One* **2009**, *4*, e4527.
148. Tay, Y.; Kats, L.; Salmena, L.; Weiss, D.; Tan, S.M.; Ala, U.; Karreth, F.; Poliseno, L.; Provero, P.; di Cunto, F.; *et al*. Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell* **2011**, *147*, 344–357.
149. Cesana, M.; Cacchiarelli, D.; Legnini, I.; Santini, T.; Sthandier, O.; Chinappi, M.; Tramontano, A.; Bozzoni, I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **2011**, *147*, 358–369.
150. Poliseno, L.; Salmena, L.; Zhang, J.; Carver, B.; Haveman, W.J.; Pandolfi, P.P. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **2010**, *465*, 1033–1038.
151. Hansen, T.B.; Jensen, T.I.; Clausen, B.H.; Bramsen, J.B.; Finsen, B.; Damgaard, C.K.; Kjems, J. Natural RNA circles function as efficient microRNA sponges. *Nature* **2013**, *495*, 384–388.
152. Memczak, S.; Jens, M.; Elefsinioti, A.; Torti, F.; Krueger, J.; Rybak, A.; Maier, L.; Mackowiak, S.D.; Gregersen, L.H.; Munschauer, M.; *et al*. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **2013**, *495*, 333–338.
153. Bail, S.; Swerdel, M.; Liu, H.; Jiao, X.; Goff, L.A.; Hart, R.P.; Kiledjian, M. Differential regulation of microRNA stability. *RNA* **2010**, *16*, 1032–1039.
154. Das, S.K.; Sokhi, U.K.; Bhutia, S.K.; Azab, B.; Su, Z.Z.; Sarkar, D.; Fisher, P.B. Human polynucleotide phosphorylase selectively and preferentially degrades microRNA-221 in human melanoma cells. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11948–11953.
155. Janowski, B.A.; Younger, S.T.; Hardy, D.B.; Ram, R.; Huffman, K.E.; Corey, D.R. Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nat. Chem. Biol.* **2007**, *3*, 166–173.
156. Li, L.C.; Okino, S.T.; Zhao, H.; Pookot, D.; Place, R.F.; Urakami, S.; Enokida, H.; Dahiya, R. Small dsRNAs induce transcriptional activation in human cells. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 17337–17342.
157. Place, R.F.; Li, L.C.; Pookot, D.; Noonan, E.J.; Dahiya, R. microRNA-373 induces expression of genes with complementary promoter sequences. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1608–1613.
158. Majid, S.; Dar, A.A.; Saini, S.; Yamamura, S.; Hirata, H.; Tanaka, Y.; Deng, G.; Dahiya, R. microRNA-205-directed transcriptional activation of tumor suppressor genes in prostate cancer. *Cancer* **2010**, *116*, 5637–5649.

159. Tsai, N.P.; Lin, Y.L.; Wei, L.N. microRNA mir-346 targets the 5'-untranslated region of receptor-interacting protein 140 (RIP140) mRNA and up-regulates its protein expression. *Biochem. J.* **2009**, *424*, 411–418.
160. Cordes, K.R.; Sheehy, N.T.; White, M.P.; Berry, E.C.; Morton, S.U.; Muth, A.N.; Lee, T.H.; Miano, J.M.; Ivey, K.N.; Srivastava, D. miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature* **2009**, *460*, 705–710.
161. Ma, F.; Liu, X.; Li, D.; Wang, P.; Li, N.; Lu, L.; Cao, X. microRNA-466l upregulates IL-10 expression in TLR-triggered macrophages by antagonizing RNA-binding protein tristetraprolin-mediated IL-10 mRNA degradation. *J. Immunol.* **2010**, *184*, 6053–6059.
162. Heil, F.; Hemmi, H.; Hochrein, H.; Ampenberger, F.; Kirschning, C.; Akira, S.; Lipford, G.; Wagner, H.; Bauer, S. Species-specific recognition of single-stranded RNA via Toll-like receptor 7 and 8. *Science* **2004**, *303*, 1526–1529.
163. Lund, J.M.; Alexopoulou, L.; Sato, A.; Karow, M.; Adams, N.C.; Gale, N.W.; Iwasaki, A.; Flavell, R.A. Recognition of single-stranded RNA viruses by Toll-like receptor 7. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5598–5603.
164. Fabbri, M.; Paone, A.; Calore, F.; Galli, R.; Gaudio, E.; Santhanam, R.; Lovat, F.; Fadda, P.; Mao, C.; Nuovo, G.J.; *et al*. microRNAs bind to Toll-like receptors to induce prometastatic inflammatory response. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E2110–E2116.

Reprinted from *IJMS*. Cite as: Zheng, H.; Fu, R.; Wang, J.-T.; Liu, Q.; Chen, H.; Jiang, S.-W. Advances in the Techniques for the Prediction of microRNA Targets. *Int. J. Mol. Sci.* **2013**, *14*, 8179-8187.

Review

Advances in the Techniques for the Prediction of microRNA Targets

Hao Zheng [1,2,†], Rongguo Fu [3,†], Jin-Tao Wang [4], Qinyou Liu [5], Haibin Chen [6,*] and Shi-Wen Jiang [2,*]

[1] School of Electrical and Computer Engineering, Georgia Institute of Technology, 225 North Avenue NW, Atlanta, GA 30301, USA; E-Mails: hzheng7@gatech.edu

[2] Department of Biomedical Science, Mercer University School of Medicine, Savannah Campus, 4700 Waters Ave, Savannah, GA 31404, USA

[3] Department of Nephrology, Second Affiliated Hospital, School of Medicine, Xi'an Jiaotong University, Xi'an 710004, Shannxi, China; E-Mail: furongguo@mail.xjtu.edu.cn

[4] Department of Epidemiology, School of Public Health, Shanxi Medical University, 56 Xin Jian Nan Road, Taiyuan 030001, Shanxi, China; E-Mail: wjtxw@yahoo.com.cn

[5] Animal Reproduction Institute of Guangxi University, 100 Daxue Road, Nanning 530005, Guangxi, China; E-Mail: qyliu2002@126.com

[6] Department of Histology and Embryology, Shantou University Medical College, 22 Xinling Road, Shantou 515041, Guangdong, China

[†] These authors contributed equally to this work.

[*] Authors to whom correspondence should be addressed; E-Mails: jiang_s@mercer.edu (S.-W.J.); chenhb@stu.edu.cn (H.C.); Tel.: +1-912-350-0411 (S.-W.J.); Fax: +1-912-350-1269 (S.-W.J.).

Received: 7 January 2013; in revised form: 1 April 2013 / Accepted: 2 April 2013 / Published: 15 April 2013

Abstract: MicroRNAs (miRNAs) are small, non-coding, endogenous RNA molecules that play important roles in a variety of normal and diseased biological processes by post-transcriptionally regulating the expression of target genes. They can bind to target messenger RNA (mRNA) transcripts of protein-coding genes and negatively control their translation or cause mRNA degradation. miRNAs have been found to actively regulate a variety of cellular processes, including cell proliferation, death, and metabolism. Therefore, their study is crucial for the better understanding of cellular functions in eukaryotes. To better understand the mechanisms of miRNA: mRNA interaction and their cellular

functions, it is important to identify the miRNA targets accurately. In this paper, we provide a brief review for the advances in the animal miRNA target prediction methods and available resources to facilitate further study of miRNAs and their functions.

Keywords: prediction; microRNA; feature selection

1. Introduction

In addition to DNA methylation and histone modification, epigenetic mechanisms have recently been extended to microRNAs (miRNAs), which are important regulators of gene expression in many biological systems. miRNAs are small, non-coding, endogenous RNA molecules, about 19–24 nucleotides in length that can negatively control their target gene expression post-transcriptionally [1]. This is mainly achieved by recognizing and binding to the 3' untranslated region of the target messenger RNA (mRNA) sequences [2]. miRNAs have been found to actively regulate a variety of cellular processes, including cell proliferation, death, and metabolism, and therefore, their study is crucial for the better understanding of cellular functions in eukaryotes [3].

Mature miRNAs are incorporated into the RNA-induced silencing complex (RISC), where miRNAs specifically interact with target mRNAs. Approximately one thousand miRNAs have been discovered in humans and are believed to control more than half of the protein coding genes, where a single miRNAs might regulate hundreds of such genes [4]. This one-to-multiple mapping presents a hurdle in accurately identifying the miRNA targets. Furthermore, miRNAs are only partially complementary to their mRNA target sequences. Such imperfections in base matching (e.g., a mismatch or bulge) make it even more difficult to accurately predict the miRNA targets *in silico* [4].

In this paper, we provide a brief review on the advances in the miRNA target prediction methods and available resources. The readers are referred to the literature cited in this review, and the references therein for further details.

2. Methods for miRNA Target Recognition

A key step in the identification of miRNA target is the selection of features that are potentially of predictive power. Many researchers are devoted to such an effort, and quite a number of predictive features have been discovered. Such features include dinucleotide composition of flanking sequence [5,6], strong base pairing between the 3' UTR of mRNAs and the miRNA seed region [7], thermodynamic stability of binding sites [8], evolutionary conservation of binding sites (particularly the seed region) [5,9], secondary structure accessibility [10,11], and host genes expression profiles [12].

The most commonly used predictive features include characteristics in the seed regions and the phylogenetic conservation of miRNA binding sites, and almost all the existing methods take advantage of such features in the algorithm.

For example, by identifying mRNAs with strong base pairing to the 5' region of the miRNA and evaluating the number and quality of these complementary sites, Lewis *et al*. identified more than 400 regulatory target genes for the conserved vertebrate miRNAs [7]. Likewise, another popular algorithm PicTar [13–17] similarly incorporated seed constraints for the identification of miRNA

targets. The new doRiNA database offers computational miRNA target site predictions for human, mouse and worm, and these predictions constitute the most recent update of PicTar predictions [17]. It is notable that some researchers have questioned the universality of the seed assumption, demonstrating that several experimentally confirmed miRNA targets do not seem to meet the seed region criterion. So far, the seed assumption is not unanimously accepted as a method to identify all miRNA targets, and that some relevant miRNA:mRNA interactions might not exhibit the seed region property [18].

With the purpose of enhancing the specificity of prediction for functional target sites, many computational studies also incorporated the evolution conservation [9,14,19–22] or flagged conserved putative targets [8,23]. Particularly, ElMMo [22] incorporated such conservation statistics in a more general, rigorous and miRNA-dependent manner. Also, Friedman *et al*. developed a quantitative method for evaluating evolutionary conservation of binding sites and applied this to the study of vertebrate miRNA targeting With this method, they found three times as many preferentially conserved sites as detected previously, further increasing the known scope and density of conserved miRNA regulatory interactions [9].

Another commonly used feature for target recognition includes the thermodynamic stability of binding sites. It is believed that the formation of a stable miRNA:target binding *in vivo*, to some extent, must be governed by thermodynamic stability. With the rationale that this binding is a process where free energy changes occur through the formation of a miRNA:target duplex, such changes may help detect miRNA targets [24,25]. The computation of energy can vary, but most methods focus only on a particular form of energy (*i.e*., hybridization) [7,14,23,26,27]. For example, Rehmsmeier *et al*. developed a program, named RNA-hybrid, which predicts multiple potential binding sites of miRNAs in large target RNAs based on the thermodynamic stability of binding sites [8].

However, more recently, combining target accessibility and duplex stability [11,28], integrated thermodynamic features for miRNA target prediction demonstrated more effectiveness. In addition, based on the immuno-precipitation (IP) of the RISC components, AIN-1 and AIN-2, Hammell *et al*. presented that total free energy change and target accessibility yielded enrichments in miRISC-enriched transcripts [25,29]. In addition to incorporating accessibility into an energy parameter [28], methods to calculate target accessibility differ, including A/U nucleotides [5,10] and larger nucleotide window to the 5' of the binding site [29]. More specifically, for example, the Sfold method was used to fold whole 3' UTR sequences plus 300 nucleotides of adjacent coding sequence for all predicted *C. elegans* transcripts. The output of Sfold was then used to calculate the average accessibility over 25 nucleotide windows flanking each potential microRNA binding site [29].

Expression-based approaches are also becoming popular to elucidate miRNA-mRNA associations. Based on expression profiles of host genes, Radfar *et al*. introduced a new computational method InMiR, which uses a linear-Gaussian model for the prediction of targets of intronic miRNAs [12]. They separated intronic miRNAs into three classes: those that are tightly regulated with their host gene; those that are likely to be expressed from the same promoter but whose host gene is highly regulated by miRNAs; and those likely to have independent promoters. Compared to a method considering only correlation, this method recovered nearly twice as many true positives as the same fixed false positive rate [12]. Engelmann *et al*. recently also showed that entire mRNA expression profiles or large groups of them can be reconstructed only from miRNA expression, and *vice versa*.

This introduced a regression model for the prediction of canonical and non-canonical miRNA-mRNA interactions [30].

Furthermore, machine learning algorithms can also be used to intelligently search for the parameters with most predictive power of genuine miRNA binding sites. An example of a method for miRNA target prediction is TargetBoost, which uses machine learning based on a set of validated miRNA targets in lower organisms to create weighted sequence motifs that capture binding characteristics between miRNAs and their targets [31]. Combining genetic programming with boosting, TargetBoost generates a metric that represents the likelihood of a site being targeted by the miRNA.

3. Resources for miRNA Target Prediction

Various popular resources for miRNA target predictions are summarized in Table 1. Different miRNA target prediction algorithms can provide differing results, and often researchers need to cross check multiple algorithms to get an additional layer of confidence for the true positive targets. For example, Ryland *et al*. incorporated miRanda [32], microCOSM Targets [33], DIANA-MicroT [27,34] and TargetScan [9] to determine whether the variants detected in mRNA 3' UTRs occurred within miRNA binding sites [35]. To facilitate that end, starBase was developed to provide a comprehensive exploration of miRNA-target interaction maps from CLIP-Seq and Degradome-Seq data [36]. This allows for a search of commonly agreed upon targets predicted by different algorithms, including TargetScan, PicTar, PITA, miRanda and RNA22 [37]. For example, when TargetScan and PicTar are selected, the database will output target sites predicted by both TargetScan and PicTar programs. This resource greatly facilitates inter-method and inter-database consensus comparison of miRNA targets. In addition, miRTar, an integrated system for miRNA target prediction, enables biologists to easily identify biological functions and regulatory relationships between a group of known/putative miRNAs and protein coding genes. Furthermore, this database delivers perspective information on miRNA targets and their alternatively spliced transcripts [38].

Table 1. Summary of prediction techniques for miRNA target recognition.

Method	Feature	References	Availability
TargetScan(S)	Database of microRNA targets conserved in 5 vertebrates.	[7,19]	http://genes.mit.edu/tscan/ targetscanS2005.html
miRanda	Optimizes sequence complementarity based on position-specific rules and interspecies conservation.	[23,32,39]	http://www.microrna.org
RNA-hybrid	Determines the most favourable hybridization site between two sequences.	[8,40]	http://bibiserv.techfak. uni-bielefeld.de/rnahybrid
PicTar (including doRiNA)	Provides details about 3’ UTR alignments with predicted sites, and links to various public databases.	[13–17]	http://pictar.mdc-berlin.de

Table 1. *Cont.*

Method	Feature	References	Availability
TargetBoost	Learns the hidden rules of miRNA-target site hybridization based on machine learning.	[31]	http://www.interagon.com/demo
PITA	Investigates the role of target-site accessibility, as determined by base-pairing interactions within the mRNA.	[11]	http://genie.weizmann.ac.il/pubs/mir07/index.html
ElMMo	Infers miRNA targets using evolutionary conservation and pathway analysis.	[22]	http://www.mirz.unibas.ch/ElMMo2/
Singh's	Predicts and characterizes 45 miRNAs by genome-wide homology search against all the reported miRNAs.	[41]	http://www.cdfd.org.in/lmg/PDF/imb816.pdf
mirWIP	Employs structural accessibility of target sequences, the total free energy of microRNA:target hybridization, and the topology of base-pairing to the 5 seed region of the microRNA.	[29]	http://ambroslab.org
microCOSM Targets	Web resource containing computationally predicted targets for microRNAs across many species.	[33]	http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/
DIANA-microT 3.0	Individually calculate several parameters for each microRNA and combines conserved and non-conserved microRNA recognition elements into a final prediction score.	[27,34]	http://www.microrna.gr/microT
starBase	Database with intersections among targets by five predictive softwares.	[36]	http://starbase.sysu.edu.cn/clipSeqIntersection.php
InMiR	Uses a linear-Gaussian model, and provides a dataset of 1,935 predicted mRNA targets for 22 intronic miRNAs.	[12]	http://www.plosone.org
miRTar	Identifies the biological functions and regulatory relationships between a group of known/putative miRNAs and protein coding genes.	[38]	http://mirtar.mbc.nctu.edu.tw/human/

4. Next-Generation Sequencing for miRNA Target Identification

With the advances of next-generation sequencing, high-throughput, systematic identification of specific miRNAs targets in a relatively short time became realistic. Several resources using CLIP-seq data to identify miRNA targets were developed, including Piranha [42], CLIPZ [43] and starBase [36].

Piranha [42] provides a utility for peak-calling based on a zero-truncated negative binomial regression model, which is able to incorporate external information to help guide the target identification process. CLIPZ provides a database and analysis environment for experimentally determined binding sites of RNA-binding proteins [43].

5. Future Work

Although quite a number of methods and databases have been developed for the identification of miRNA targets, most methods have a false positive rate (FPR) greater than 0.3, which means that the specificity is often lower than 70%. FPR is evaluated as (1-specificity), where specificity is defined as the ratio of the number of true negatives and true negatives plus false positives. Filtering for true positive targets from the large predicted target lists is challenging and time consuming. Although conservation and functional similarities have been taken advantage of to reduce false positives, there is still much room for improvement. Since different miRNA target prediction algorithms still provide varying results, this indicates that such methods also suffer from higher rates of false negatives. As a result, highly accurate prediction algorithms with small false positive and false negative rates need to be further developed. Such algorithms are crucial to studying the exact role of miRNA in signaling pathways, as well as associations with various disease pathways.

To better perform the comparative study of different methods, it is imperative to have some "gold standard" data sets, and quantitatively evaluate different methods based on a fixed set of metrics. The establishment of a gold standard requires strong experimental evidence (reporter assay or western blot analysis) as well as consensus across independent experiments.

Acknowledgments

The authors would like to thank Charlie Bodine for his efforts to edit the manuscript.

Shi-Wen Jiang is a Distinguished Cancer Scholar supported by Georgia Cancer Coalition (GCC). This project is partially supported by Mercer University School of Medicine Seed Research Funding (S-W Jiang), The Memorial Health Hospital and Anderson Cancer Institute Pancreatic Cancer Program (S-W Jiang), and the research grant of the Natural Science Foundation of China (81070590; 81100530; R Fu).

Conflict of Interest

The authors declare no conflict of interest.

References

1. Chuang, J.; Jones, P. Epigenetics and MicroRNAs. *Pediatr. Res.* **2007**, *61*, 24R–29R.
2. Yue, D.; Liu, H.; Huang, Y. Survey of computational algorithms for microRNA target prediction. *Curr. Genomics* **2009**, *10*, 478–492.
3. Bartel, D. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297.
4. Yue, D.; Meng, J.; Lu, M.; Chen, C.; Guo, M.; Huang, Y. Understanding microRNA regulation: A computational perspective. *Signal Proc. Mag. IEEE* **2012**, *29*, 77–88.

5. Nielsen, C.; Shomron, N.; Sandberg, R.; Hornstein, E.; Kitzman, J.; Burge, C. Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA* **2007**, *13*, 1894–1910.
6. Ohler, U.; Yekta, S.; Lim, L.; Bartel, D.; Burge, C. Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification. *RNA* **2004**, *10*, 1309–1322.
7. Lewis, B.; Shih, I.; Jones-Rhoades, M.; Bartel, D.; Burge, C. Prediction of mammalian microRNA targets. *Cell* **2003**, *115*, 787–798.
8. Rehmsmeier, M.; Steffen, P.; Hochsmann, M.; Giegerich, R. Fast and effective prediction of microRNA/target duplexes. *RNA* **2004**, *10*, 1507–1517.
9. Friedman, R.; Farh, K.; Burge, C.; Bartel, D. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **2009**, *19*, 92–105.
10. Grimson, A.; Farh, K.; Johnston, W.; Garrett-Engele, P.; Lim, L.; Bartel, D. MicroRNA targeting specificity in mammals: Determinants beyond seed pairing. *Mol. Cell* **2007**, *27*, 91–105.
11. Kertesz, M.; Iovino, N.; Unnerstall, U.; Gaul, U.; Segal, E. The role of site accessibility in microrna target recognition. *Nat. Genet.* **2007**, *39*, 1278–1284.
12. Radfar, M.; Wong, W.; Morris, Q. Computational prediction of intronic microRNA targets using host gene expression reveals novel regulatory mechanisms. *PLoS One* **2011**, *6*, e19312.
13. Grun, D.; Wang, Y.; Langenberger, D.; Gunsalus, K.; Rajewsky, N. microRNA target predictions across seven drosophila species and comparison to mammalian targets. *PLoS Comput. Biol.* **2005**, *1*, e13.
14. Krek, A.; Grun, D.; Poy, M.; Wolf, R.; Rosenberg, L.; Epstein, E.; MacMenamin, P.; Piedade, I.; Gunsalus, K.; Stoffel, M.; *et al*. Combinatorial microrna target predictions. *Nat. Genet.* **2005**, *37*, 495–500.
15. Lall, S.; Grun, D.; Krek, A.; Chen, K.; Wang, Y.; Dewey, C.; Sood, P.; Colombo, T.; Bray, N.; Macmenamin, P.; *et al*. A genome-wide map of conserved microrna targets in *C. elegans*. *Curr. Biol.* **2006**, *16*, 460–471.
16. Chen, K.; Rajewsky, N. Natural selection on human microRNA binding sites inferred from SNP data. *Nat. Genet.* **2006**, *38*, 1452–1456.
17. Anders, G.; Mackowiak, S.; Jens, M.; Maaskola, J.; Kuntzagk, A.; Rajewsky, N.; Landthaler, M.; Dieterich, C. doRiNa: A database of RNA interactions in post-transcriptional regulation. *Nucleic Acids Res.* **2012**, *40*, D180–D186.
18. Liu, B.; Li, J.; Cairns, M.J. Identifying miRNAs, targets and functions. *Brief Bioinforma.* **2012**, doi:10.1093/bib/bbs075.
19. Lewis, B.; Burge, C.; Bartel, D. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **2005**, *120*, 15–20.
20. Stark, A.; Brennecke, J.; Bushati, N.; Russell, R.; Cohen, S. Animal microRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* **2005**, *123*, 1133–1146.
21. Johnson, S.M.; Grosshans, H.; Shingara, J.; Byrom, M.; Jarvis, R.; Cheng, A.; Labourier, E.; Reinert, K.L.; Brown, D.; Slack, F.J. RAS is regulated by the let-7 microRNA family. *Cell* **2005**, *120*, 635–647.
22. Gaidatzis, D.; Nimwegen, E.; Hausser, J.; Zavolan, M. Inference of miRNA targets using evolutionary conservation and pathway analysis. *BMC Bioinforma.* **2007**, *8*, 69:1–69:22.

23. John, B.; Enright, A.; Aravin, A.; Tuschl, T.; Sander, C.; Marks, D. Human microRNA targets. *PLoS Biol.* **2004**, *2*, e363.
24. Hammell, M. Computational methods to identify miRNA targets. *Semin. Cell Dev. Biol.* **2010**, *21*, 738–744.
25. Lekprasert, P.; Mayhew, M.; Ohler, U. Assessing the utility of thermodynamic features for microRNA target prediction under relaxed seed and no conservation requirements. *PLoS One* **2011**, *6*, e20622.
26. Stark, A.; Brennecke, J.; Russell, R.; Cohen, S. Identification of drosophila microRNA targets. *PLoS Biol.* **2003**, *1*, e60.
27. Maragkakis, M.; Alexiou, P.; Papadopoulos, G.L.; Reczko, M.; Dalamagas, T.; Giannopoulos, G.; Goumas, G.; Koukis, E.; Kourtis, K.; Simossis, V.A.; *et al*. Accurate microRNA target prediction correlates with protein repression levels. *BMC Bioinforma.* **2009**, *10*, 295:1–295:10.
28. Long, D.; Lee, R.; Williams, P.; Chan, C.; Ambros, V. Potent effect of target structure on microRNA function. *Nat. Struct. Mol. Biol.* **2007**, *14*, 287–294.
29. Hammell, M.; Long, D.; Zhang, L.; Lee, A.; Carmack, C. mir-WIP: MicroRNA target prediction based on microrna-containing ribonucleoprotein-enriched transcripts. *Nat. Methods* **2008**, *5*, 813–819.
30. Engelmann, J.; Spang, R. A least angle regression model for the prediction of canonical and non-canonical miRNA-mRNA interactions. *PLoS One* **2012**, *7*, e40634.
31. Saetrom, O.; Snove, O.; Saetrom, P. Weighted sequence motifs as an improved seeding step in microRNA target prediction algorithms. *Cell* **2005**, *1*, 995–1003.
32. Betel, D.; Wilson, M.; Gabow, A.; Marks, D.; Sander, C. The microrna.org resource: Targets and expression. *Nucleic Acids Res.* **2008**, *36*, D149–D153.
33. Griffiths-Jones, S.; Saini, H.; Dongen, S.; Enright, A. Mirbase: Tools for microRNA genomics. *Nucleic Acids Res.* **2008**, *36*, D154–D158.
34. Maragkakis, M.; Reczko, M.; Simossis, V.; Alexiou, P.; Papadopoulos, G. Diana-microt web server: Elucidating microRNA functions through target prediction. *Nucleic Acids Res.* **2009**, *37*, W273–W276.
35. Ryland, G.; Bearfoot, J.; Doyle, M.; Boyle, S.; Choong, D.; Rowley, S. MicroRNA genes and their target 3' untranslated regions are infrequently somatically mutated in ovarian cancers. *PLoS One* **2012**, *7*, e35805.
36. Yang, J.; Li, J.; Shao, P.; Zhou, H.; Chen, Y.; Qu, L. starBase: A database for exploring microRNAcmRNA interaction maps from Argonaute CLIP-seq and Degradome-Seq data. *Nucleic Acids Res.* **2011**, *39*, D202–D209.
37. Miranda, K.; Huynh, T.; Tay, Y.; Ang, Y.; Tam, W.; Thomson, A.; Lim, B.; Rigoutsos, I. A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* **2006**, *126*, 1203–1217.
38. Hsu, J.B.; Chiu, C.M.; Hsu, S.D.; Huang, W.Y.; Chien, C.H.; Lee, T.Y.; Huang, H.D. miRTar: An integrated system for identifying miRNA-target interactions in human. *BMC Bioinforma.* **2011**, *12*, 300:1–300:12.
39. Enright, A.; John, B.; Gaul, U.; Tuschl, T.; Sander, C.; Marks, D. MicroRNA targets in drosophila. *Genome Biol.* **2003**, *5*, R1–R14.

40. Kruger, J.; Rehmsmeier, M. RNAhybrid: MicroRNA target prediction easy, fast and flexible. *Nat. Genet.* **2005**, *37*, 495–500.
41. Singh, J.; Nagaraju, J. *In silico* prediction and characterization of microRNAs from red flour beetle (*Tribolium castaneum*). *Insect Mol. Biol.* **2008**, *17*, 427–436.
42. Uren, P.J.; Bahrami-Samani, E.; Burns, S.C.; Qiao, M.; Karginov, F.V.; Hodges, E.; Hannon, G.J.; Sanford, J.R.; Penalva, L.O.; Smith, A.D. Site identification in high-throughput RNA–protein interaction data. *Bioinformatics* **2012**, *1*, 3013–3020.
43. Khorshid, M.; Rodak, C.; Zavolan, M.; CLIPZ: A database and analysis environment for experimentally determined binding sites of RNA-binding proteins. *Nucleic Acids Res.* **2011**, *39*, D245–D252.

Reprinted from *IJMS*. Cite as: Gomes, A.Q.; Nolasco, S.; Soares, H. Non-Coding RNAs: Multi-Tasking Molecules in the Cell *Int. J. Mol. Sci.* **2013**, *14*, 16010-16039.

Review

Non-Coding RNAs: Multi-Tasking Molecules in the Cell

Anita Quintal Gomes [1,2], Sofia Nolasco [1,3,4] and Helena Soares [1,3,5,*]

[1] Health Technology College of Lisbon—Polytechnic Institute of Lisbon, 1990-096 Lisbon, Portugal; E-Mails: anita.gomes@estesl.ipl.pt (A.Q.G.); sofianolasco@fmv.utl.pt (S.N.)
[2] Institute of Molecular Medicine, Faculty of Medicine, University of Lisbon, 1649-028 Lisbon, Portugal
[3] Gulbenkian Science Institute, 2780-256 Oeiras, Portugal
[4] Interdisciplinary Centre of Research in Animal Health (CIISA), Faculty of Veterinary Medicine, 1300-666 Lisbon, Portugal
[5] Center for Chemistry and Biochemistry, Department of Chemistry and Biochemistry, Faculty of Sciences, University of Lisbon, 1749-016 Lisbon, Portugal

***** Author to whom correspondence should be addressed; E-Mail: mhsoares@fc.ul.pt; Tel.: +351-217-500-853; Fax: +351-217-500-088.

Received: 7 June 2013; in revised form: 15 July 2013 / Accepted: 19 July 2013 / Published: 31 July 2013

Abstract: In the last years it has become increasingly clear that the mammalian transcriptome is highly complex and includes a large number of small non-coding RNAs (sncRNAs) and long noncoding RNAs (lncRNAs). Here we review the biogenesis pathways of the three classes of sncRNAs, namely short interfering RNAs (siRNAs), microRNAs (miRNAs) and PIWI-interacting RNAs (piRNAs). These ncRNAs have been extensively studied and are involved in pathways leading to specific gene silencing and the protection of genomes against virus and transposons, for example. Also, lncRNAs have emerged as pivotal molecules for the transcriptional and post-transcriptional regulation of gene expression which is supported by their tissue-specific expression patterns, subcellular distribution, and developmental regulation. Therefore, we also focus our attention on their role in differentiation and development. SncRNAs and lncRNAs play critical roles in defining DNA methylation patterns, as well as chromatin remodeling thus having a substantial effect in epigenetics. The identification of some overlaps in their biogenesis pathways and functional roles raises the hypothesis that these molecules play concerted functions *in vivo*, creating complex regulatory networks where cooperation with regulatory proteins is necessary. We also highlighted the implications of biogenesis and gene expression deregulation of sncRNAs and lncRNAs in human diseases like cancer.

Keywords: sncRNAs; lncRNAs; miRNAs; siRNAs; piRNAs; gene expression regulation; epigenetic regulation

1. Introduction

1.1. The Incredible RNA Molecules

RNA has been known since the late 1800s, but its importance in cell functioning has long been in the shadow of DNA and proteins. In the 1950s, with the establishment of the molecular structure of DNA, it was proposed that RNA would be an intermediate molecule in the information flux between DNA and proteins. Later, this was experimentally demonstrated revealing that during gene expression, DNA is copied in a molecule of messenger RNA (mRNA) that is then translated into proteins with the help of other RNA molecules like transfer RNA (tRNA) and ribosomal RNAs (rRNAs). The idea that RNAs are much more than molecules involved in storage/transfer of information emerged with the discovery of ribozymes, RNA molecules that have, like proteins, active roles as catalysts of chemical reactions in cells. The two ribozymes identified first have RNAs as substrates and were the *Tetrahymena* intron of the 26S rRNA that is a self-sufficient catalytic unit capable of autoexcision and autocyclization [1], and the ribonucleoprotein, RNase P, an enzyme containing an RNA subunit essential for the catalysis required for the synthesis of tRNAs [2]. These discoveries clearly encouraged a variety of studies to search for potential new roles of RNA molecules *in vivo*, and led to the re-evaluation of RNAs as crucial molecules in the evolution of life. In view of the ability of RNAs to catalyze biological reactions, it is conceivable that the first organisms could rely only on RNA molecules and that only later an evolution of a more complex system based on proteins was established. This hypothesis gave support to the model of a primordial "RNA World" (for review [3,4]).

Progressively, the participation of RNAs in other critical molecular processes in eukaryotic cells was revealed, as in the case of DNA replication (RNA primers allow DNA polymerases to start the process), protein translation and RNA transcript maturation. For example, several ribosome functions required for protein synthesis were shown to be, at least in part, RNA-mediated, including peptidyl transferase activity [5], decoding functions [6], and the tRNA acceptor site interaction with 23S rRNA [7]. On the other hand, many small non-coding RNA molecules were isolated and characterized as being associated with proteins originating from ribonucleoprotein complexes (RNP), later identified as the components of the splicesome, including U1, U2, U4, U5 and U6 small nuclear RNA (snRNA) [8].

Furthermore, the information content of tRNA, rRNA and mRNA molecules can be biochemically altered after transcription by different molecular mechanisms that are generally designated by RNA editing [9]. These include sequence changes such as nucleoside modifications from C to U and A to I deaminations, as well as non-templated nucleotide additions and insertions. In general, RNA editing mechanisms are based on protein or protein-RNA complexes responsible for the RNA editing reaction and require a "guide RNA" molecule, which, through base-pairing with the target RNA molecule, determines the editing site. By this mechanism an mRNA sequence may be post-transcriptionally altered and consequently the amino acid sequence of the protein will then differ from that predicted by the genomic DNA sequence. Moreover, post-transcriptional processing and modifications of rRNAs

are important for the production of efficient and accurate ribosomes which is directed by two large guide families of small nucleolar RNAs (snoRNA) [10].

In the mid-1980s Blackburn and Greider, demonstrated the existence of an enzymatic activity within cell extracts that added tandem hexanucleotides to chromosome ends and led to the discovery of telomerase [11]. Today it is well established that telomerase is a specialized reverse transcriptase that uses an internal RNA template sequence that is responsible for the synthesis of telomeric repeats [12]. More recently, Quiao and Cech [13] have described that the non-template RNA part of telomerase works together with the protein reverse-transcriptase motifs to facilitate catalysis, using a mechanism resembling that of pure ribozymes [13]. According to the "RNA first" model it was speculated that telomerase arose by the association of an ancient ribozyme with the reverse-transcriptase subunit. In light of this hypothesis the telomerase RNA may be a molecular fossil and telomerase a missing link in the evolution from RNA enzymes to protein enzymes [13]. This type of close functional collaboration is also observable in snoRNPs.

At this point the growing descriptions of the importance of RNA molecules for cell function started to push them to the limelight, but the complexity of their roles and the wide variety of molecular mechanisms where RNA molecules are critical players was still far from clear. In recent years, the use of genome wide approaches and the large output of genome sequencing technologies have revealed that the mammalian transcriptome is much more complex than previously thought since it includes a large number of small non-coding RNAs (sncRNAs) and long noncoding RNAs (lncRNAs) [14,15]. In most cases, these molecules present complex and precise patterns of expression during differentiation and development, tissue specificity, and some have been related to different pathophysiological states [16]. For example, it became clear that snoRNA guide families are widely diverse, which seems to be related to variant snoRNA structures and multiple cellular RNA targets, and consequently to cellular functions beyond ribosome biogenesis [10,17]. Indeed, snoRNAs have been recently implicated in alternative splicing and in cell transformation, tumorigenesis, and metastasis (for review, see [18]) showing that we are far from having a complete picture of their roles *in vivo*. Importantly, the observations that exogenously introduced double stranded RNA (dsRNA) molecules and plasmids expressing short hair-pin RNA (shRNA) specifically base-pairing with target mRNA molecules were able to trigger mRNA degradation (RNA interference -RNAi) [19,20] revealed, for the first time, that specific silencing pathways based on sncRNAs operate in eukaryotic cells. Moreover, these observations led to the development of the powerful RNA interference (RNAi) technique that has been extensively used in the study of gene function.

The aim of this review is to give a summarized overview of the biogenesis pathways of distinct classes of sncRNAs, including miRNA, piRNA, and siRNA, as well as lncRNAs, focusing on the miRNA and lncRNAs gene regulatory roles in distinct cellular functions and developmental regulatory programs. We will highlight the implications of the deregulation of miRNA and lncRNAs biogenesis pathways further illustrating the role of these molecules in the establishment of human diseases such as cancer. Finally, we will bring to discussion the fact that the pathways where distinct family members of sncRNAs and lncRNA funtion are probably interconnected, establishing a complex network of interactions and actions required for rapid and fine-tuned gene expression regulation at multiple levels.

1.2. The Small Non-Coding RNAs

Three classes of sncRNAs, namely short interfering RNAs (siRNAs), microRNAs (miRNAs) and PIWI-interacting RNAs (piRNAs), have been extensively studied in the last decade and have been associated with pathways that lead to silencing of specific genes and to the protection of the cell/genome against viruses, mobile repetitive DNA sequences, retro-elements and transposons [16].

1.2.1. siRNAs and miRNAs

The siRNAs and miRNAs (~20–30 nucleotides long) originate from double-stranded RNA (dsRNA) precursors that are introduced into, or produced endogenously by gene transcription of both sense and anti-sense DNA strands and of pseudogenes and inverted repeats. These molecules are critical in pathways involved in mRNA degradation, translational repression, or both, therefore regulating gene expression.

In the case of siRNAs, they are small RNA duplex molecules produced by the action of Dicer, a ribonuclease III (RNaseIII) enzyme that creates RNA duplexes with 2-nt overhangs at their 3' ends and phosphate groups at their 5' ends [21].

The miRNAs are mostly transcribed by RNA polymerase II as primary-miRNA (pri-miRNA) molecule precursors that possess a characteristic stem loop structure and are subsequently subjected to processing mechanisms [22]. In animals, the first step occurs in the nucleus where the RNaseIII Drosha acts over pri-mRNAs generating a pre-miRNA, a small RNA duplex of ~65–70 nucleotides containing the hair pin. This action can be facilitated by RNA processing proteins such as hnRNP A1 [23]. The pre-miRNAs are then exported to the cytoplasm by a nuclear transport receptor complex, exportin-5–RanGTP [24] where they are processed by Dicer into ~22-nt mature miRNAs (miRNA–miRNA* duplexes, where miRNA is the antisense, or guide/mature strand, and miRNA* is the sense, or passenger strand).

An alternative nuclear pathway for miRNA biogenesis was described in invertebrates [25] where the pre-miRNA is processed via splicing/spliceosome, instead of Drosha. Accordingly, spliced lariats linearized by the lariat debranching enzyme accept monophosphates and 3' hydroxyls, the same moieties found in pre-miRNAs, that were designated by-miRNAs/introns, “mirtrons” (for review [24]). These mirtrons are subsequently exported to the cytoplasm and processed by a Dicer protein.

The next step, for both siRNA and miRNA production, is the subsequent association with members of the Argonaute protein family that have diverged into specialized clades (or subfamilies), each recognizing different sncRNA types and conferring the specific features of the various silencing pathways operating in cells [26]. Argonaute loading occurs in the RNA-induced silencing complex (RISC)-loading complex, a ternary complex that consists of an Argonaute protein, Dicer and a dsRNA-binding protein (known as TRBP in humans). During loading, the non-guide strand is cleaved by an Argonaute protein [22].

The selection of the different Argonaute proteins seems to be based on the small interfering RNA duplex structure. For example, siRNAs that are perfect duplexes in terms of base pairing are loaded into Argonaute 2 (Ago2), whereas duplexes presenting mismatches, as in the case of miRNAs, are generally driven to Argonaute 1 (Ago1) [27,28]. When the complementarity between the miRNA bound to Ago1 and the target RNA is high, this causes miRNA tailing and 3'- to 5'-trimming. The

discrimination between Ago1 and Ago2 seems to depend on the action of Hen1 an enzyme that adds the 2'-*O*-methyl group at the 3' ends of small RNAs bound to Ago2, but not those bound to Ago1 [29]. This methyl group is known to block tailing and trimming of the miRNA. The maturation and function of certain miRNAs can be also associated to enzymatic post-transcriptional modifications, like mono-uridylation [30]. These modifications will increase the variety of miRNAs and their precursor pools allowing more complex schemes of regulation in different backgrounds.

In the small RNA duplex of the siRNA the guide strand seems to be the one whose 5' end is less tightly paired to its complement [31]. In both siRNAs and miRNAs the guide strands drives the RISCs to the target mRNAs that contain complementary sequences thereby causing their degradation or translation inhibition (for review [16,32]). Recently, it has been shown that the target choice can also depend on accessory factors that interact with Dicer. For example, the *Drosophila* Loqs-PB Dicer-partner cleaves pre-miR-307a, generating a longer miRNA isoform with a distinct seed sequence and target specificity [33]. The mammalian TRBP homologue also acts together with Dicer to cleave pre-miR-132 generating a longer miRNA and consequently targets different mRNA molecules [34].

In fission yeast a specialized nuclear complex, known as the RNA-induced transcriptional silencing complex (RITS), mediates transcriptional gene silencing by inducing heterochromatin formation [35]. The RITS complex consists of Chp1 (H3K9me binding protein), Ago1, a poorly characterized protein Tas3, and siRNAs derived from centromeric repeat sequences [36]. These studies also showed the existence of a tight coupling of both siRNA and H3K9 methylation that appears to be important for the recruitment of RITS for heterochromatin assembly [37]. Therefore, there seems to exist a complex interplay between the RNAi pathway and the chromatin modifying machinery [37].

1.2.2. The piRNAs

The piRNAs are the least characterized class of sncRNAs and, contrary to the siRNAs and miRNAs that are widely expressed in different tissues and cell types, the piRNAs have been essentially detected in the germline cells of mammals, fish and *Drosophila melanogaster* [38,39] where they are important for germ line development and to suppress transposon activity. Mutations that disrupt the piRNA biogenesis pathway in mouse and fish cause germline-specific cell death and sterility, and are also associated with increased transposon expression [40].

The piRNAs (~24–31 nucleotides) got their name from the fact that they only associate to the PIWI subfamily of the Argonaute protein family (Piwi proteins). These sncRNAs usually have a uridine at the 5' end, hold a 5' monophosphate, and present a 2'-*O*-methyl (2'-*O*-Me) modification on the nucleotide at the 3' end (for review, see [32]).

Although not much is known concerning the intervening factors involved in piRNAs biogenesis pathways and transcription regulation, it is now well documented that they diverge from siRNAs and miRNAs by being generated by RNaseIII-independent pathways that do not involve dsRNA precursors. These sncRNAs are generated from long single-stranded precursors [41,42] that are preferentially cleaved at U residues and loaded onto Piwi proteins.

In *Drosophila*, as in mammals, the majority of piRNAs are transcribed from discrete genomic *loci* that are clustered in large pericentromeric or subtelomeric domains, generally spanning from 50–100 kb, and that comprise mainly various transposable DNA elements and their remnants [41].

Other piRNAs are derived from 3' UTRs of protein coding genes and dispersed euchromatin copies of transposable elements [41,43]. Most of these clusters are active specifically in germ cells, while only a single major cluster (*flamenco*) impels transposon silencing in the soma. Interestingly, if new transposons are introduced into piRNA clusters, and if they are heritable by the progeny, novel piRNAs will be produced that can lead to the control of the new transposons indicating that the mechanisms that drive adaptation to transposon invasion might be mediated by the piRNA pathway [44].

The critical role of piRNAs on transposon silencing was demonstrated by loss-of function mutations in *Drosophila* piRNAs and genes coding for the proteins involved in their biogenesis. In the germline these mutations cause a retro-transposition up-regulation causing the loss of germ cells and a variety of defects due to alterations in microtubule cytoskeleton polarization, with consequences to the polarized localization of specific proteins and mRNAs required for normal oogenesis [45]. However, it was found that the derepression of transposons activates the Chk2 DNA damage checkpoint [44] suggesting that the described phenotypes are probably an indirect consequence of transposon overexpression and DNA damage signaling (for review [44]).

Other studies also reveal that besides being involved in keeping genome integrity, a subset of piRNA genes have been implicated in the assembly of the telomere protection complex [46].

Detailed analysis of the small RNAs associated with the Piwi sub-family (PIWI, Aubergine and Argonaute 3) [41,47] in the *Drosophila* female germline showed that these sncRNAs have in their structure and sequences, signatures that give clues about their biogenesis. The most abundant piRNAs are mainly generated from the antisense strand of retro-transposons and these preferentially associate with Piwi and Aubergine proteins [41,47]. Those present in the single major somatic cluster are mainly originated from the sense strand and are associated with Argonaute 3 (Ago3).

The piRNAs from the germ cells seem to be generated by a self-amplifying loop designated by ping-pong cycle. Specifically, PIWI and/or Aubergine form complexes with antisense piRNAs that direct the slicing of sense strand transposon transcripts [41,47]. The sliced sense strands are then bound by Ago3, and this complex directs the slicing of antisense transposon transcripts [41]. A similar mechanism seems to operate in other animal genomes [42,48]. The piRNAs derived from genomic regions depleted of transposons, seem to be generated by a different pathway not completely understood called “primary processing” that operates in somatic cells, and may have a role in the regulation of target mRNAs (for review [32]).

Recent studies have shown that, in addition to their role in germ line transposon regulation and genome stability, piRNAs have a broader function in heterochromatin formation and developmental gene regulation. The analysis of a high-throughput small RNA sequencing data in *Drosophila*, mouse and rhesus macaque samples demonstrated that piRNAs are widespread and are abundant in other tissues as much as in the germline [49]. In fact, their involvement in the regulation of gene expression was demonstrated in *Drosophila*, where the degradation of a subset of maternal RNAs, *i.e.*, embryonic posterior morphogen Nanos (Nos), at the maternal-to-zygotic transition, was shown to require the zygotic expression of a piRNA cluster [50]. When this expression is inhibited, the Nos mRNA is stabilized which was accompanied by a reduced deadenylation and translational derepression, resulting in head development defects. Because the piRNAs involved in this regulation are produced from transposable elements, the authors suggested the existence of a direct developmental function for transposable elements in the regulation of gene expression through piRNAs [50].

The importance of the piRNAs pathway in the nervous system and in epigenetic regulation has also been gaining support. In the hippocampus, the inhibition of piRNAs causes a decrease of the dendrite spine area suggesting that these sncRNAs are required for spine morphogenesis [51]. More recently, in Aplysia sensory neurons, a Piwi/piRNA complex was described to facilitate the methylation of a conserved CpG island in the promoter of the transcriptional repressor of memory, CREB2, in a serotonin-dependent manner [52]. Consequently, this Piwi/piRNA complex is at the cross-roads between a transient external stimuli and alterations in the gene-expression of neurons involved in long term memory storage. Sienski *et al.* [53] have also shown that in *Drosophila* ovarian somatic cells piRNAs mediate the silencing of hundreds of transposon copies at the transcriptional level by establishing heterochromatic methylation of H3K9 on transposons and their genomic surroundings. The involvement of the piRNA pathway in *de novo* methylation of the differentially methylated region of the imprinted mouse *Rasgrf1 locus* [54] shows that the role of this pathway in methylation is also extendable to mammalian genomes.

There is also growing evidence that piRNA-pathway dependent mechanisms may have been critical during evolution, in the establishment of developmental robustness. In fact, the piRNA-pathway seems to be required for preventing phenotypic variation despite genotypic variation and environmental influences (canalization) [55]. The Hsp90 protein was previously described as a capacitator [56] being able to prevent phenotypic variation by suppressing the mutagenic activity of transposons [57]. Interestingly, it was shown in *Drosophila* that a protein complex composed of Hsp90, Piwi and Hop, is involved in canalization, probably through phosphorylation regulation of the Piwi protein by Hsp90 and Hop [55]. Therefore, it is possible that the Piwi-piRNA pathway will mediate canalization by both suppressing the generation of new genotypes and epigenetically silencing the expression of existing genetic variants [55].

The piRNAs, contrary to miRNAs, are less conserved through the eukaryotic lineage. This difference has been explained by the possible co-evolvement of miRNA with their RNA targets which have created sequence divergence constraints. There are increasing examples that piRNAs play roles in somatic cells regulating protein encoding genes. It is possible that piRNAs are more likely to be involved in epigenetic regulation rather than post-transcriptional regulation [58]. These puzzling facts suggest that our knowledge of the mechanistic relationships between piRNAs and the regulatory mechanisms based on regulatory proteins is far from being understood. On the other hand, the initial evidence that piRNAs may be involved in epigenetic regulation in tumorigenesis [59,60] requires additional attention.

The role of piRNAs in protecting genomes against parasitic nucleic acids seems to have developed early in evolution since ciliates present a mechanism that resembles that of piRNAs. Ciliates are single celled organisms that present a polyploid macronucleus that guarantees the vegetative growth of cells (the somatic nucleus) and the diploid micronucleus that is only active during sexual conjugation and constitutes the germline [61]. After conjugation the zygotic macronucleus differentiates from the micronucleus by undergoing an extensive developmentally programmed genome reorganization [62]. This reorganization involves chromosome fragmentation and elimination of germline limited sequences (internal eliminated sequences (IES), transposons and other repeated sequences) according to the pre-existing rearrangements of the maternal somatic genome. This seems to rely on a global comparison of the germline and somatic genomes and a genomic subtraction between meiosis-specific,

germline scnRNAs (small RNAs that resemble piRNAs) and longer non-coding transcripts from the somatic genome (for review, see [63]). This mechanism that parallels the patterns of heterochromatin formation in other eukaryotes allows the maintenance of an epigenetic memory of rearrangement patterns across sexual generations and establishes, in an ancestral unicellular organism, a relationship between piRNAs and development. The ancestrality of the piRNAs and the fact that they have been placed in developmental frameworks being protagonists in the establishment of developmental robustness strongly supports the view that they have been critical factors in eukaryotic evolution.

2. Long Noncoding RNAs

It is now clear that the mammalian genome produces a large transcriptome of long noncoding RNA (lncRNA, defined as RNA >100 nucleotides in length). The number of gene members integrating this class of ncRNAs is still under debate and ranges from 10,000 to >200,000 [64].

The lncRNAs can be transcribed from intergenic regions, promoter regions or be interleaved, overlapping or antisense to annotated protein-coding genes [44]. There is also growing evidence that lncRNAs molecules might be produced by transcriptional active pseudogenes [65]. Although the majority of lncRNAs are transcribed from the nuclear genome, recently it was found that some can be generated from mitochondrial genomes [66]. Like coding genes, lncRNAs undergo post-transcriptional processing, including 5'capping, alternative splicing, RNA editing, and polyadenylation [67,68].

The referred transcriptional origins have been used to establish classification classes for lncRNA, as for example promoter-associated long RNAs (lpaRNAs) [68], natural antisense transcripts (NATs) or opposite-strand transcripts [69], large intervening noncoding RNA (lincRNA) [70], and enhancer associated RNAs (eRNA) [71,72]. However, other criteria should probably be used since frequently one lncRNA molecule can be associated with more than one class.

Mammalian genomes encode a large number of natural antisense transcripts (NATs) [64,73]. For instance, the FANTOM-3 mouse transcriptome sequencing consortium identified natural antisense transcripts for more than 70% of the transcription units, the majority of which represent non-protein-coding RNAs [73].

NATs have been defined as endogenous RNA molecules at least partially complementary to transcripts of known function [74]. NATs can be transcribed from the opposite strand at the same genomic locus of their sense counterparts and will present perfect sequence complementarity being designated by *cis*-NATs. On the other hand those transcribed from different genomic loci may have imperfect sequence complementarity and are named *trans*-NATs [75]. Sense and antisense RNA pairs can present different relative orientations and variable overlapping regions. For example, they can overlap by their 5' regions (5' to 5'), by their 3' regions (3' to 3'), or fully-overlap (one gene included within the region of the other) [76]. Antisense RNAs have a tendency to have lost introns and typically show lower abundance compared with sense transcripts [77].

Studies performed in various organisms have suggested that NATs can participate in a broad range of regulatory events that will be discussed later.

3. The Emerging Roles of lncRNAS and miRNAs

3.1. LncRNAs: Implications in Different Levels of Gene Expression Regulation and Differentiation

LncRNAs have emerged as pivotal molecules for the regulation of gene expression [76]. These transcripts are biologically relevant as supported by their cell-specific expression pattern [78], subcellular distribution [79], developmental regulation and possible association with human diseases.

LncRNAs encompass a wide variety of functions which include almost all levels of gene expression regulation, ranging from epigenetic to translational regulation, including transcriptional and post-transcriptional control. The main functions of lncRNAs are summarized below.

3.1.1. Epigenetic Regulation

lncRNAs modulate chromatin through the specific recruitment of histone and chromatin modifying complexes on one hand and by the recruitment of transcription factors on the other hand. X chromosome inactivation (XCI) is the classic example of the former type of regulation and is caused by the lncRNA "Xist" which physically associates with the Polycomb repressive complex 2 (PRC2) recruiting it to the X chromosome ultimately leading to its inactivation [80]. More precisely, it is a 1.6-kb ncRNA (RepA) within Xist that targets PRC2. Depletion of RepA abolishes full-length Xist induction and trimethylation on lysine 27 of histone H3 of the X (thus abolishing X inactivation). In addition it was demonstrated that PRC2 deficiency compromises Xist up-regulation [80]. A similar process to XCI is genomic imprinting, an epigenetic event in which genes are expressed from the allele of only one parent. One of the first lncRNAs to be identified was H19, which is reciprocally imprinted with insulin-like growth factor 2 (Igf2). Even though this lncRNA is highly expressed, its deletion has no phenotype and, in fact, recently it has been proposed to function as a microRNA precursor [81].

Other lncRNAs (*i.e.*, Air, Kcnq1ot1, HOTAIR) can control chromatin states in *cis* and/or in *trans*, thereby regulating gene expression through the association with chromatin-modifying complexes [82,83]. Specifically, HOTAIR is a *trans*-acting lncRNA that serves as a scaffold for two histone modification complexes: it binds both to polycomb repressive complex 2 (PRC2) and to LSD1 (in complex with CoREST/REST). This coordinates targeting of PCR2 and LSD1 to chromatin for coupled histone H3 at lysine 27 methylation and lysine 4 demethylation leading to subsequent gene silencing [84]. Also, in the plant *Arabidopsis* it was demonstrated that environmental conditions, such as cold, are able to induce the transcription of related NATs (*i.e.*, COOLAIR) that are involved in the silencing of a flower repressor *locus* designated by flowering locus c (*FLC*) [85]. More recently it was discovered that a lncRNA, named COLDAIR, that differs from COOLAIR by the fact that it is transcribed in the sense direction relative to *FLC* mRNA transcription, interacts directly with PRC2 and targets it to *FLC*, establishing an epigenetic memory [86]. Interestingly, winter cold triggers the methylation of H3 at *FLC* and it was shown that COLDAIR is induced by cold, demonstrating that lncRNas participate in the integration of signals from the environment to cell signaling pathways.

Other *trans*-acting lncRNAs have different functions some of which remain incompletely defined. For example, the p21-associated ncRNA DNA damage-activated (PANDA) lncRNA is induced upon DNA damage in a p53-dependent manner and it interacts with the transcription factor NF-YA to limit expression of pro-apoptotic genes [87]. Mistral is another example of an lncRNA that acts on the

recruitment of the transcription factor MLL1 thereby activating Hoxa6 and Hoxa7 expression and subsequent stem cell differentiation [88].

Another group of lncRNAs that play a role in mammalian genomes are the long intergenic non-coding RNAs (lincRNAs) that range in size from ~300 nucleotides to several thousands and that, in humans, have been estimated to be around 3300, although a more correct number may be closer to 4500 [89]. This group of transcripts is heterogeneous but show significant evolutionary conservation relative to neutral sequences [70], which support the idea that they have important functions. In fact, it has been described that some groups of lincRNAs present expression patterns that correlate with those observed for protein-coding genes involved in cellular processes as diverse as cell-cycle regulation, innate immunity responses, and stem cell pluripotency [70,90]. In agreement, a reference catalog of 8195 human lincRNAs based on integratingRNA-seq data from 24 tissues and cell types showed that lincRNAs are expressed in a more tissue-specific manner than protein-coding genes [91]. By using co-immunoprecipitation and RNAi approaches it was also demonstrated that lincRNAs are associated with chromatin-modifying complexes to specific genomic loci to regulate gene expression [89]. The capacity to bind chromatin-modifying proteins or transcription factors, as exemplified, in combination with the abundance of lncRNAs suggests that lncRNAs may be part of a broad epigenetic regulatory network (reviewed in [92,93]).

3.1.2. Transcriptional Regulation

The discovery and characterization of several ncRNAs that are able to associate with promoters (promoter associated RNAs—paRNA) is also changing the traditional view of how genes encoding proteins are regulated at the transcriptional level. Promoter associated RNAs paRNAs are transcribed approximately from the start of or within the promoter, and include long, short and tiny RNA molecules (for review [49]). The long paRNAs were found at a single-gene level and were also associated with the modification of DNA methylation and demethylation patterns [94], inhibition of transposition expression in *Saccharomyces cerevisiae* [95] and gene expression in humans [96].

Interestingly, long (antisense) paRNAs have the potential to form double stranded molecules that can be processed into endo-siRNAs, and that, due to their sequence complementarity to that of a promoter, are able to induce transcriptional gene silencing [97–99] or activation [100–102] in a similar way to short paRNAs [49]. This picture is far from being complete since an increasing amount of experimental data are supporting the idea that enhancers can be transcribed and the resulting enhancer-non coding transcripts (eRNAs) may, in some cases, have functional roles, rather than represent mere transcriptional noise (for review see [103,104]).

On the other hand, lncRNAs can modulate the function of transcription factors by acting as co-regulators, modulators of transcription factors activity or by regulating the association and activity of co-regulators, among others. The ncRNA Evf-2, for example, functions as a co-activator for the homeobox transcription factor Dlx2, which plays important roles in forebrain development and neurogenesis [105]. Local ncRNAs can also recruit transcriptional factors and co-activating molecules to regulate adjacent protein-coding gene expression. The RNA binding protein TLS, binds to and inhibits the CREB binding protein (CCND1) and p300 histone acetyltransferase activities on a repressed gene target, cyclin D1. The recruitment of TLS to the promoter of cyclin D1 is directed by

single stranded, low copy number lncRNA transcripts tethered to 5' regulatory regions of CCND1 in response to DNA damage signals [106].

Finally, lncRNAs also regulate the basal transcription machinery by targeting transcription factors required for the RNAP II transcription of all genes [107]. These general factors include components of the initiation complex that assemble on promoters or are involved in transcription elongation. An example of lncRNA-mediated regulation of basal transcription is the formation of a stable RNA-DNA triplex within the major promoter of the dihydrofolate reductase (DHFR) by an lncRNA that is transcribed from an upstream minor promoter of the DHFR gene. This complex prevents the binding of the transcriptional co-factor TFIIB [96].

3.1.3. Post-Transcriptional Regulation

LncRNAs can act on splicing, on mRNA stability and translation. It has been shown that lnc antisense RNA may bind to the sense RNA, masking the splice sites and thereby changing the balances between splice variants. Thyroid hormone receptor alpha gene (*TRα*) is an example where the antisense transcript *RevErbAα* influences splicing of *TRα1* and *TRα2* mRNAs [108]. Recently, it was discovered that a new class of sno-lncRNAs, whose ends correspond to positions of intronic snoRNA, are able to interact with the splicing factor Fox2 and alter splicing patterns [109]. The authors also showed that some of these sno-lncRNAs map to a genomic region that is deleted in the patients presenting Prader-Willi syndrome, strongly suggesting an association of these sno-lncRNAs with the disease. LncRNAs can also recruit proteins to mRNA to promote its degradation or stabilization. There's evidence for lncRNA binding to sequences present in the 3' UTR of specific mRNAs, thus creating a recognition site for Staufen, a protein that binds double-stranded mRNA and induces its decay [110]. By contrast, the lncRNA TINCR (terminal differentiation-induced ncRNA) also interacts with Staufen 1 but the complex between TINCR-STAU1 seems to mediate stabilization of mRNAs encoding differentiation factors such as Keratin 80 [111]. TINCR-mRNA interaction occurs through a motif of 25 nt that is abundantly present in target interacting mRNAs [111]. Another example is that of the mRNA of BACE1, a β-secretase responsible for β-amyloid production, that is stabilized and protected from RNase cleavage by base pairing of its antisense (BACE1-AS) [112]. Therefore, different lncRNAs are able to differentially regulate factors involved in mRNA stability regulation. Translational regulation is yet another proposed function for lncRNAs. Such is the case of the antisense for *PU.1* mRNA. *PU.1* mRNA translation is inhibited by a noncoding antisense transcript, which is a polyadenylated RNA with a lower concentration but a half-life longer than the sense *PU.1* transcript [113]. On the other hand the lncRNA Uchl1, shuttles from the nucleus to the cytoplasm under the control of the mTOR pathway and is involved in the translation up-regulation of the ubiquitin carboxy terminal hydrolase L1 (UCHL1) mRNA by promoting its association with polysomes [114]. Interestingly, the UCHL1 is a specific neuronal protein involved in rampamycin neuroprotective function and more generally in cellular stress response, that has been associated with neurodegenerative diseases. The various referred examples clearly show that lncRNAs present a vast repertoire of strategies to post-transcriptionally regulate protein encoding genes and different molecules are able to differentially modulate a specific regulatory molecule or pathway.

3.1.4. Modulation of mRNA Nuclear Trafficking and Control of Nuclear Compartmentalization

NRON is a non-coding repressor of nuclear factor of the activated T cells (NFAT), which interacts with multiple proteins including members of the importin-beta superfamily and likely functions as a specific regulator of NFAT nuclear trafficking [115].

The lncRNA nuclear-enriched autosomal transcript 1 (NEAT1), and abundant 4 kb ncRNA, is retained in nuclei *foci* that are coincident with "paraspeckles" [79]. It has been demonstrated that it contributes to the formation of these dynamic structures of the interchromatin space that are implicated in mRNA retention [79].

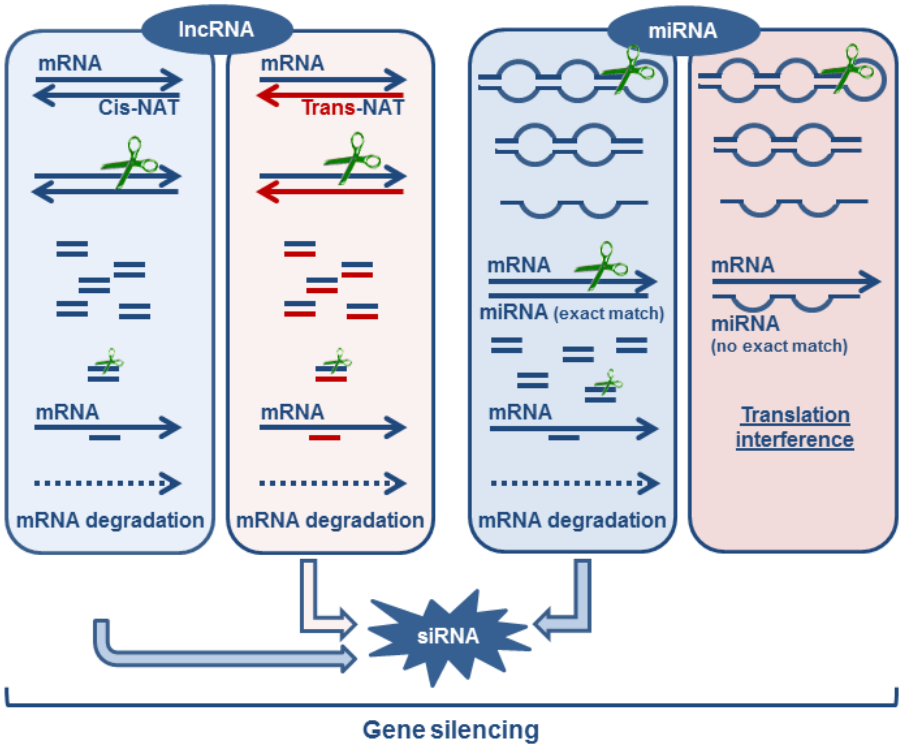
3.1.5. Formation of Endogenous siRNA

It has been described that NATs can originate siRNAs that will be involved in mRNA down-regulation. This mechanism requires the formation of a sense:anti-sense pair of transcripts that are then processed into siRNAs. This pair can be originated directly from the same *loci* (*cis*-NATs) or from different *loci* (*trans*-NATs). Interestingly, it was observed that certain *trans*-NATs are produced from pseudogene transcription. For example, in rice, a small number of pseudogenes are transcribed and processed into siRNAs, after pairing with the coding gene or a paralogous pseudogene transcript [116]. A similar observation was reported in mammals where pseudogene transcripts can be processed into small interfering RNAs (siRNA) with the ability to repress gene expression in mouse oocytes [117,118]. Therefore, NATs play their gene expression regulatory role through a mechanism equivalent to that of miRNAs and siRNAs (see Figure 1). Until recently, pseudogenes were envisaged only as copies of protein-coding genes that have lost the ability to produce functional proteins therefore constituting junk DNA in genomes [119]. Pseudogenes can be created by diverse processes, including: (1) spontaneous mutations, preventing transcription of the gene, or translation of the protein [120]; (2) duplication, in which pseudogenes are originated via tandem duplication or uneven crossing-over leading to the loss of promoters or enhancers or the appearance of crippling mutations such as frame shifts or premature stop codons [119]; and (3) retro-transposition, the mRNA transcript being reverse-transcribed and integrated into the genome at a new location originating retro-transposed or processed pseudogenes [121,122]. Therefore, their origin directly makes them prone to participate in post-transcriptional regulatory mechanisms promoted by lncRNAs. These observations started to change the vision that pseudogenes are mere junk in the genomes of organisms, and suggested that they can play important biological roles. In agreement with this hypothesis is the fact that the transcription of NATs is generally regulated in a tissue-specific manner and varying sense/antisense ratios are found [123].

It is evident that the critical roles of lncRNAs at different levels of gene expression regulation will largely contribute to establish differential profiles of gene expression required for development [124–127]. This is supported by the observation that lncRNAs such as Xist [128], TUG1 [129], PINC [130], and HOTAIR [131] have important roles in development. Moreover, Dinger *et al.* [90] using a microarray to examine the expression profiles of mouse embryonic stem cells differentiating as embryoid bodies over a 16 day time course have identified 945 ncRNAs, of which 174 were differentially expressed, many correlating with pluripotency or specific differentiation events [90]. Accordingly, it was also observed that the expression of some lincRNAs is increased in induced pluripotent stem cells (iPSCs)

in comparison to those found in stem cells. This suggests that their activation may promote the emergence of iPSCs. It was also demonstrated that one of this lincRNA (lincRNARoR) modulates the reprogramming process leading to pluripotent stem cells [132].

Figure 1. Gene silencing: mRNA post-transcriptional regulation by lncRNA and miRNA. lncRNAs can be transcribed as natural antisense transcripts, from the same loci (*cis*-NAT, the same gene is transcribed in both directions) or from a different loci (*trans*-NAT, for example from a pseudogene). These NATs transcripts can pair with the coding transcripts, originating dsRNA molecules that will activate the siRNA machinery leading to mRNA degradation. miRNAs are also complementary of coding mRNAs and can pair with a perfect match leading to the activation of the siRNA machinery or they can pair with gaps leading to translation interference.



Spermatogenesis is a very complex developmental process that requires precise microtubule cytoskeleton remodeling, creating complex microtubule structures such as the manchette and the flagellum of the sperm [133]. During this process it was observed that the gene encoding TBCA, a protein that interacts with β-tubulin and is involved in the folding and dimerization of new tubulin heterodimers (the building blocks of microtubules) is regulated by a *Tbca* pseudogene that is transcribed in both directions [134]. The *Tbca* pseudogene is down-regulated leading to the increase of the *Tbca* mRNA, during testis maturation suggesting that this *Tbca* lncRNA is required for the undifferentiated state of spermatids. Similarly, the gene encoding the nitric oxide synthase protein (NOS2A) is transcribed into a noncoding RNA containing a region of significant antisense homology with the NOS2A mRNA. As in the case of *Tbca* lncRNA, the expression patterns of the anti-NOS2A RNA and the NOS2A mRNA exhibit opposite changes in undifferentiated human embryonic stem cells (hESCs) and in hESCs induced to differentiate into neurogenic precursors [74]. In conclusion,

lncRNAs are clearly required to regulate programs of differentiation during development and seem to be generally associated with the undifferentiated states, repressing critical target genes whose expression is crucial for the cells to reach their fate.

3.2. miRNAs as Critical Regulators of Target Degradation and Translation

miRNAs act as sequence-specificity guides for the RNAi machinery to mediate repression of target gene expression. First identified as regulators of larval development in nematodes [19], miRNAs are now known to serve key roles in the regulation of almost every important cellular process in all multicellular eukaryotes. These include cell development, proliferation, differentiation, apoptosis and oncogenic transformation [135]. The genome of human cells encodes over 1000 miRNA species that regulate 60% of all protein-coding genes [32]. Most mRNA targets contain multiple miRNA binding sites, and each miRNA can regulate multiple genes. Therefore, the deregulation of miRNA levels might perturb the expression of many genes, thereby playing a key role in the occurrence of diseases (see below).

It is still unclear whether miRNAs act mainly at the mRNA translational or transcriptional levels. The miRNA repression, at the level of transcriptional inhibition, can occur as a consequence of mRNA decay, direct mRNA cleavage or through miRNA-mediated chromatin reorganization. Decay of targeted mRNA occurs without direct cleavage at the binding site. Unlike in translational inhibition where only a slight protein decrease can be obtained, protein level reductions greater than 33% indicate that mRNA decay is the major component of miRNA-driven silencing [136]. miRNA-mediated mRNA decay can occur via deadenylation, decapping or 5' to 3' degradation of the mRNA [137]. Dicer1, Ago1 and Ago2 were shown to be required for the rapid decay of mRNA containing AU-rich elements (AREs) in the 3' UTR of tumor necrosis factor-alpha suggesting that miRNA targeting of ARE is essential to mediate mRNA degradation [137]. It was also shown that upon GW182 interaction with AGO1, there is recruitment of deadenylases and decapping enzymes, leading to mRNA degradation [138]. The mRNA cleavage, another miRNA transcription repressive mechanism that is rare in animals, but frequent in plants, normally occurs when there is full complementarity between the miRNA and its mRNA target [139]. miRNAs also have the capacity to reorganize chromatin by increasing methylation of the targeted mRNA promoters thereby inhibiting their expression [140].

Finally, the repressed mRNAs, Ago proteins and miRNAs are frequently accumulated in processing bodies (P-bodies), which are cytoplasmic structures enriched in the mRNA degradation machinery but where the translational machinery is normally absent [141].

The second major mechanism of miRNAs activity includes repression of translation initiation and/or elongation, premature termination and nascent polypeptide degradation. Inhibition of translation initiation can occur at the level of cap-40S association or via 40S-AUG-60S association. Endogenous let-7 micro-ribonucleoproteins (miRNPs) or the tethering of Ago proteins to reporter mRNAs in human cells inhibit m(7)G-cap-dependent translation initiation, suggesting that miRNPs interfere with the recognition of the cap [142]. The cap-binding protein eukaryotic initiation factor 4E has in fact been proposed as a molecular target of miRNA function [143]. Ago2 represses the initiation of mRNA translation by directly binding to the m(7)G-cap of mRNA targets, thus likely precluding the recruitment of eIF4E [144]. Another Ago1 was shown to interact with GW182, this interaction being essential for miRNA-mediated inhibition of translation [145]. It was also shown that miRNA-repressed

mRNAs contain 40S but not 60S components suggesting that miRNAs repress translation initiation by preventing the 60S subunit from joining to miRNA-targeted mRNAs [146]. It has also been reported that some miRNAs can inhibit translation initiation by inducing the formation of dense miRNPs (pseudo-polysomes) [147]. The fact that various studies showed repressed mRNA targets to be associated with polyribosomes seems to indicate that miRNAs can also repress translation at the elongation step [148–150]. Silencing by miRNAs can also occur before completion of the nascent polypeptide chain causing a decrease in translational read through at a stop codon, with ribosomes on repressed mRNAs dissociating more rapidly after a block of initiation of translation, than those of control mRNAs [149]. These observations pinpoint a role for miRNAs in ribosome drop-off-mediated repression.

Intriguingly, there is also evidence for transcriptional [100] and translational [151] activation by miRNAs. The miRNA-373 was shown to induce expression of genes with complementary promoter sequences [100]. miRNA-10a can bind to the 5' UTR of ribosomal protein mRNAs and enhances their translation [152]. Further, a growing series of studies has demonstrated that miRNAs and their associated complexes (microRNPs) elicit alternate functions that enable stimulation of gene expression in addition to their assigned repressive roles [151,153].

While the global importance of miRNAs is clearly illustrated by the developmental failure of Dicer-deficient embryonic stem cells (*in vitro*) and embryos (*in vivo*) [154], unique spatial and temporal expression patterns in distinct hematopoietic and neuronal lineages are clearly suggestive of multiple roles for miRNA in hematopoiesis, immune responses and neurological differentiation. The specific profiling of hPSCs by microarray and sequencing methods has allowed the identification of miRNAS that have potential roles in differentiation and development (reviewed in [155]). Several miRNA families, including the human (hsa)-miR-302, hsa-miR-106, hsa-miR-372, hsa-miR-17, hsa-miR-520, hsa-miR-195 and hsa-miR-200 families [155] were up-regulated specifically in hPSCs compared to mature differentiated cell types. Interestingly, the “seed” sequences (short sequence at nucleotides 2-8 on the 5' end of the miRNA that binds to the 3' UTRs of their target mRNAs) for most of these miRNAs are closely related, suggesting that these miRNA families may share mRNA targets. Thus, their regulatory functions might help maintain the unique characteristics of PSCs. Contrary to the miRNA families the hsa let-7 family [155] is expressed at significantly lower levels in hPSCs than in differentiated cells. The miRNA-dependent post-transcriptional gene regulation is also crucial for neural and immune cell development. Early evidence for miRNA function in the nervous system development came partly from knockout mutations of the miRNA processing genes present in the miRNA pathway. Pioneering studies of nervous system development using maternal-zygotic mutants of zebrafish dicer revealed gross morphological defects specifically in early brain patterning and morphogenesis [156]. Detailed studies of later stages in neural development have begun to suggest a more extensive contribution of miRNAs in the formation of synaptic connections, circuit maturation, and the activity-driven plasticity of these connections. For example, the mRNA processing enzyme DGCR8 mutant mice exhibited abnormalities in synaptic connectivity due to a reduction in the number and size of dendritic spines, reduced synaptic complexity, impaired synaptic transmission, and altered short-term plasticity [157].

In the immune system, miRNAs mediate the regulation of T cell development and function, as confirmed by the observation of defective thymic and peripheric T cell subsets in Dicer deficient mice [158,159].

Individual miRNAs play different roles at distinct developmental stages. For example, miR-125b and miR-132 regulate dendritic spine development. More specifically, miR-125b and miR-132 (as well as several other miRNA) are associated with fragile X mental retardation protein (FMRP) in mouse brain. The miR-125b overexpression results in longer, thinner processes of hippocampal neurons. FMRP knockdown is shown to ameliorate the effect of overexpressed miR-125b and miR-132 on spine morphology. It has been proposed that miR-125b negatively regulates its target, NR2A, along with FMRP and AGO1 [160].

Focusing on T cells, miRNA expression patterns vary among stages of development and T cell subsets, which indicate that these molecules may contribute to the identity of the cell subsets or their functional state [161]. Consistent with this, recent reports have demonstrated that various miRNAs, namely miR-101, miR-150, miR-155, miR181a, miR-29a, miR-146a and miR-326, are expressed in particular T cell subsets and regulate several aspects of their differentiation and function [162–164] .

Like lncRNAs, miRNAs are required to regulate differentiation programs during development. However, they are associated with both undifferentiated and differentiated states repressing target genes involved in maintaining those programs.

4. ncRNAs Active Players in Cancer and Other Human Diseases

The deregulation of gene expression networks, responsible for normal cellular identity, growth and differentiation leads to cancer. The large majority of genome-wide association studies (GWAS) identify cancer risk loci outside of protein-coding regions. Of 301 single-nucleotide polymorphisms (SNPs) currently linked to cancer, only 12 (3.3%) change the protein amino-acid sequence. Most are located in the introns of protein-coding genes (40%) or intergenic regions (44%), raising the question of the function of these noncoding loci and their role in cancer development [165]. These facts, associated with the observations that miRNA and lncRNAs are involved in programs of differentiation and development soon raise the hypothesis that alterations in their profiles of expression could be correlated with cancer development. In the last years, numerous evidences have confirmed this hypothesis since miRNas and long ncRNAs, that present tissue-specific expression, were found to be deregulated in distinct types of cancers. For example, data coming from microarray expression from a wide range of distinct cancers showed that alterations in miRNAs are almost always present in the analyzed tumors [166]. More specifically, overexpression of miR-155 was reported in hematopoietic cancers, breast, lung and colon cancer [167], whereas miR-21 was found to be overexpressed in glioblastoma and to have antiapoptotic properties [168,169]. Also, transgenic mice overexpressing miR-17-92 developed lymphoproliferative disorders [170] and retroviral overexpression of the cluster accelerated lymphoma formation. The miR-17-92 cluster was also found to be overexpressed in lung, colon and gastric cancer [171]. Like miRNAs, lncRNAs have also been associated with cancer development. For example, the lncRNA MALAT1 is up-regulated in several cancer types and its overexpression has been linked to an increase in cell proliferation and migration in lung and colorectal cancer cells [165]. These phenotypes are probably related to the role of MALAT1 in controlling alternative splicing of pre-mRNAs [172]. However, this relationship is probably too simplistic since a more recent study indicates that MALAT1 may also have a role in the regulation of gene expression, different from alternative splicing, in lung metastasis [173].

Many studies have also shown that miRNA and LncRNAs themselves can function as tumor suppressor genes or oncogenes, [174–176]. Several studies found that the tumor suppressor p53 transcriptionally regulates the three gene members of the miR-34 family. On the other hand, the miR-34 activation resembles p53 activity, including induction of cell-cycle arrest and promotion of apoptosis, and loss of miR-34 can impair p53-mediated apoptosis [177]. However, the interaction between p53 and miR-34 is much more complex since mice possessing the combined loss of all three miR-34 members are viable and fertile, do not display morphological defects and are not prone to spontaneous tumor formation [178].

Similarly to miRNAs, some lincRNAs are transcriptional targets of p53 like the lincRNA-p21 that plays a role as a transcriptional repressor in the p53 pathway by triggering apoptosis. The lincRNA-p21 binds to hnRNP-K that allows for the correct localization of hnRNP-K, probably by influencing their target preference, and therefore the transcriptional repression of p53-regulated genes [179]. The precise mechanism by which lincRNA-p21 contributes to repression at specific loci remains to be defined.

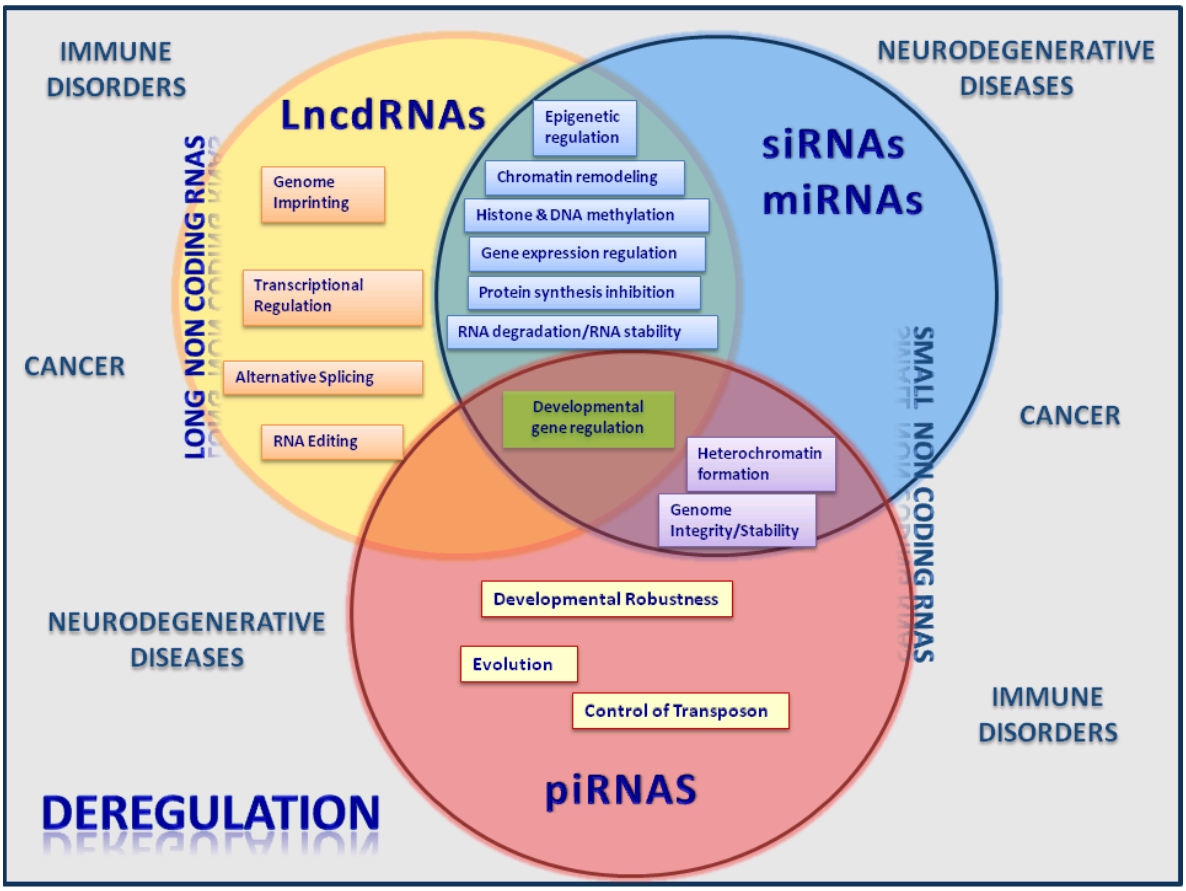
Although most of the mechanisms that implicate lncRNA in cancer biology are uncovered, the growing available data show that they are probably linked for example to chromatin remodeling. For example lncRNAs that are known to be involved in the recruitment of epigenetic modifiers to specific loci such as ANRIL, XIST, HOTAIR and KCNQ1OT1 were observed to have modified expression in a variety of cancers [176]. Also the lncRNA named TERRA, which binds telomerase, inhibiting its activity *in vitro* [180] is downregulated in many cancer cells which may be related to the longevity of cancer cells.

The broad functional classes of genes and regulatory pathways that involve ncRNA participation clearly justifies that the deregulation of their biogenesis and roles could are not restricted to cancer development (Figure 2). Perturbations in the biogenesis and actions of ncRNAs have also been associated with diverse neurodegenerative diseases such as Huntington's disease [181], Alzheimer [182] and Parkinson [183]. Moreover, recent studies have shown that miRNAs have unique expression profiles in cells of the innate and adaptive immune systems, suggesting that these molecules are important regulators of immune cell functions (reviewed in [184]). In fact, the role of miRNAs have been linked to autoimmune disorders (e.g., systemic lupus erythematosus, rheumatoid arthritis, multiple sclerosis, inflammatory bowel disease, psoriasis) and inflammatory pathologies of distinct organ (e.g., atherosclerosis, osteoarthritis, atopic eczema) and/or systemic locations like allergy. Chromatin remodeling by lncRNA is not exclusively related with cancer but is also linked to other diseases like facioscapulohumeral muscular dystrophy (FSHD) [185], lethal lung developmental disorder [186] and the HELLP syndrome, a pregnancy-associated disease [187].

The presented examples directly implicate long ncRNA and miRNAs in cancer biology and other human diseases and indicate that a complex interplay between their biogenesis pathways, their regulatory mechanisms and their targets should be seriously taken into consideration not only in cancer research but in other human pathologies and also in the definition of future strategies of diagnostics and therapeutics.

Figure 2. Diagram of functional relationships among lncRNAs, siRNAs, miRNas and piRNAs. This "venn diagram" depicts the specific function of each RNA molecule (inside each circle) as well as the shared functions (overlapping areas). Some of the disorders

caused by deregulation in the expression patterns of these RNA molecules are indicated outside the circles.



5. Concluding Remarks

In the last years we have witnessed an unprecedented discovery of numerous functions of non-coding RNAs in eukaryotic cells ranging from gene expression regulation to genome imprinting roles that were previously attributed to proteins.

This means that proteins are likely to cooperate with ncRNAs to control gene expression at different levels of regulation. They cooperate in the regulation of the transcription of genes encoding proteins, to process and maturate their transcripts and finally to regulate their mRNA stability and translation. Moreover, upstream of these regulatory steps, cooperation will also be required for altering DNA methylation profiles and the remodeling of chromatin contributing to epigenetic regulation. This means that complex networks between proteins and RNAs have been established during the course of evolution. We can envisage and speculate that due to their biochemical nature and biogenesis, ncRNAs will contribute to speed-up, make more flexible, transform and ultimately make more accurate the regulatory pathways conducted by regulatory proteins, pushing gene expression regulation to a new level. It is predictable that this complex regulatory web will have several hubs that will be composed of ncRNAs and proteins or alternatively only proteins or ncRNAs, which will also allow a rapid and better integration of different environmental signals. Although, the field of ncRNAs has been growing fast we are still far from understanding the complexity and the mechanisms underlying the establishment of the regulatory networks between RNAs and proteins.

From the evolutionary point of view it seems that the "invention" of proteins like telomerase was a critical step in the establishment of accurate spatial and temporal regulatory processes which probably allowed the evolution of eukaryotic complexity and later on, the appearance of multicellularity. In the view of an "RNA World hypothesis" it is tempting to speculate that the first RNA activities were to maintain the viability and integrity of "cell precursors" defending them from destructive invader molecules; these ancestral defense functions of ncRNAs that are still present and operate in "modern cells" seem to havebeen extended to mechanisms of gene regulation.

It should also be pointed out that the close analysis of different classes of ncRNAs (sncRNA and lncRNA), and the fact that we can detect biogenesis (see Figure 1) and functional overlaps between them (see Figure 2), strongly supports the idea that they could also have close interactions, not only at the level of their biogenesis pathways, but also at the functional level. This has been probably missed to a certain extent by the fact that they have been essentially separately studied.

From what has been compiled, the deregulation of biogenesis and functional roles of ncRNAs were, as expected at the crossroads of different human pathologies ranging from cancer to neurodegenerative and immune diseases. Finally the continued understanding of the molecular mechanisms and signaling pathways where ncRNAs participate should offer new insights to define new diagnostic strategies and open new avenues for therapies.

Acknowledgments

We are deeply in debt to João Gonçalves (Mount Sinai Hospital, Samuel Lunefeld Institute, Canada) for critical reviewing the manuscript. This work was supported by Fundação para a Ciência e a Tecnologia (FCT), PEst-OE/QUI/UI0612/2013.

Conflict of Interest

The authors declare no conflict of interests.

References

1. Cech, T.R.; Zaug, A.J.; Grabowski, P.J. *In vitro* splicing of the ribosomal RNA precursor of Tetrahymena: Involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell* **1981**, *27*, 487–496.
2. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **1983**, *35*, 849–857.
3. Dworkin, J.P.; Lazcano, A.; Miller, S.L. The roads to and from the RNA world. *J. Theor. Boil.* **2003**, *222*, 127–134.
4. Lehman, N. RNA in evolution. *Wiley Interdiscip. Rev.* **2010**, *1*, 202–213.
5. Noller, H.F.; Hoffarth, V.; Zimniak, L. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **1992**, *256*, 1416–1419.
6. Von Ahsen, U.; Noller, H.F. Identification of bases in 16S rRNA essential for tRNA binding at the 30S ribosomal P site. *Science* **1995**, *267*, 234–237.

7. Moazed, D.; Noller, H.F. Sites of interaction of the CCA end of peptidyl-tRNA with 23S rRNA. *Proc. Natl. Acad. Sci.USA* **1991**, *88*, 3725–3728.
8. Wassarman, D.A.; Steitz, J.A. Interactions of small nuclear RNA's with precursor messenger RNA during *in vitro* splicing. *Science* **1992**, *257*, 1918–1925.
9. Brennicke, A.; Marchfelder, A.; Binder, S. RNA editing. *FEMS Microbial. Rev* **1999**, *23*, 297–316.
10. Bachellerie, J.P.; Cavaille, J.; Huttenhofer, A. The expanding snoRNA world. *Biochimie* **2002**, *84*, 775–790.
11. Greider, C.W.; Blackburn, E.H. Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* **1985**, *43*, 405–413.
12. Greider, C.W.; Blackburn, E.H. A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis. *Nature* **1989**, *337*, 331–337.
13. Qiao, F.; Cech, T.R. Triple-helix structure in telomerase RNA contributes to catalysis. *Nat. Struct. Mol. Boil.* **2008**, *15*, 634–640.
14. Birney, E.; Stamatoyannopoulos, J.A.; Dutta, A.; Guigo, R.; Gingeras, T.R.; Margulies, E.H.; Weng, Z.; Snyder, M.; Dermitzakis, E.T.; Thurman, R.E.; *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **2007**, *447*, 799–816.
15. Mattick, J.S.; Makunin, I.V. Non-coding RNA. *Hum. Mol. Genet.* **2006**, *15*, R17–R29.
16. Moazed, D. Small RNAs in transcriptional gene silencing and genome defence. *Nature* **2009**, *457*, 413–420.
17. Eliceiri, G.L. Small nucleolar RNAs. *Cell. Mol. Life Sci.* **1999**, *56*, 22–31.
18. Mannoor, K.; Liao, J.; Jiang, F. Small nucleolar RNAs in cancer. *Biochim. Biophys. Acta* **2012**, *1826*, 121–128.
19. Lee, R.C.; Feinbaum, R.L.; Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **1993**, *75*, 843–854.
20. Reinhart, B.J.; Slack, F.J.; Basson, M.; Pasquinelli, A.E.; Bettinger, J.C.; Rougvie, A.E.; Horvitz, H.R.; Ruvkun, G. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **2000**, *403*, 901–906.
21. Meister, G.; Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* **2004**, *431*, 343–349.
22. Jinek, M.; Doudna, J.A. A three-dimensional view of the molecular machinery of RNA interference. *Nature* **2009**, *457*, 405–412.
23. Michlewski, G.; Guil, S.; Semple, C.A.; Cáceres, J.F. Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Mol. Cell* **2008**, *7*, 383–393.
24. Zeng, Y.; Cullen, B.R. Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic Acids Res.* **2004**, *32*, 4776–4785.
25. Berezikov, E.; Chung, W.J.; Willis, J.; Cuppen, E.; Lai, E.C. Mammalian mirtron genes. *Mol. Cell* **2007**, 28, 328–336.
26. Farazi, T.A.; Juranek, S.A.; Tuschl, T. The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development* **2008**, *135*, 1201–1214.

27. Forstemann, K.; Horwich, M.D.; Wee, L.; Tomari, Y.; Zamore, P.D. Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. *Cell* **2007**, *130*, 287–297.
28. Tomari, Y.; Du, T.; Zamore, P.D. Sorting of *Drosophila* small silencing RNAs. *Cell* **2007**, *130*, 299–308.
29. Ameres, S.L.; Horwich, M.D.; Hung, J.H.; Xu, J.; Ghildiyal, M.; Weng, Z.; Zamore, P.D. Target RNA-directed trimming and tailing of small silencing RNAs. *Science* **2010**, *328*, 1534–1539.
30. Heo, I.; Ha, M.; Lim, J.; Yoon, M.J.; Park, J.E.; Kwon, S.C.; Chang, H.; Kim, V.N. Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. *Cell* **2012**, *151*, 521–532.
31. Khvorova, A.; Reynolds, A.; Jayasena, S.D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **2003**, *115*, 209–216.
32. Siomi, H.; Siomi, M.C. On the road to reading the RNA-interference code. *Nature* **2009**, *457*, 396–404.
33. Fukunaga, R.; Han, B.W.; Hung, J.H.; Xu, J.; Weng, Z.; Zamore, P.D. Dicer partner proteins tune the length of mature miRNAs in flies and mammals. *Cell* **2012**, *151*, 533–546.
34. Lee, H.Y.; Doudna, J.A. TRBP alters human precursor microRNA processing *in vitro*. *RNA* **2012**, *18*, 2012–2019.
35. Grewal, S.I.; Elgin, S.C. Transcription and RNA interference in the formation of heterochromatin. *Nature* **2007**, *447*, 399–406.
36. Verdel, A.; Jia, S.; Gerber, S.; Sugiyama, T.; Gygi, S.; Grewal, S.I.; Moazed, D. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **2004**, *303*, 672–676.
37. Creamer, K.M.; Partridge, J.F. RITS-connecting transcription, RNA interference, and heterochromatin assembly in fission yeast. *Wiley Interdiscip. Rev.* **2011**, *2*, 632–646.
38. Aravin, A.A.; Lagos-Quintana, M.; Yalcin, A.; Zavolan, M.; Marks, D.; Snyder, B.; Gaasterland, T.; Meyer, J.; Tuschl, T. The small RNA profile during *Drosophila melanogaster* development. *Dev. cell* **2003**, *5*, 337–350.
39. Lau, N.C.; Seto, A.G.; Kim, J.; Kuramochi-Miyagawa, S.; Nakano, T.; Bartel, D.P.; Kingston, R.E. Characterization of the piRNA complex from rat testes. *Science* **2006**, *313*, 363–367.
40. Aravin, A.A.; Hannon, G.J.; Brennecke, J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* **2007**, *318*, 761–764.
41. Brennecke, J.; Aravin, A.A.; Stark, A.; Dus, M.; Kellis, M.; Sachidanandam, R.; Hannon, G.J., Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **2007**, *128*, 1089–1103.
42. Houwing, S.; Kamminga, L.M.; Berezikov, E.; Cronembold, D.; Girard, A.; van den Elst, H.; Filippov, D.V.; Blaser, H.; Raz, E.; Moens, C.B.; *et al.* A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell* **2007**, *129*, 69–82.
43. Robine, N.; Lau, N.C.; Balla, S.; Jin, Z.; Okamura, K.; Kuramochi-Miyagawa, S.; Blower, M.D.; Lai, E.C. A broadly conserved pathway generates 3' UTR-directed primary piRNAs. *Curr. Biol.* **2009**, *19*, 2066–2076.
44. Khurana, J.S.; Theurkauf, W. piRNAs, transposon silencing, and *Drosophila* germline development. *J. Cell Biol.* **2010**, *191*, 905–913.

45. Chen, Y.; Pane, A.; Schupbach, T. Cutoff and aubergine mutations result in retrotransposon upregulation and checkpoint activation in *Drosophila*. *Curr. Biol.* **2007**, *17*, 637–642.
46. Khurana, J.S.; Xu, J.; Weng, Z.; Theurkauf, W.E. Distinct functions for the *Drosophila* piRNA pathway in genome maintenance and telomere protection. *PLoS Genet.* **2010**, *6*, e1001246.
47. Gunawardane, L.S.; Saito, K.; Nishida, K.M.; Miyoshi, K.; Kawamura, Y.; Nagami, T.; Siomi, H.; Siomi, M.C. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. *Science* **2007**, *315*, 1587–1590.
48. Ishizu, H.; Siomi, H.; Siomi, M.C. Biology of PIWI-interacting RNAs: New insights into biogenesis and function inside and outside of germlines. *Genes Dev.* **2012**, *26*, 2361–2373.
49. Yan, B.X.; Ma, J.X. Promoter-associated RNAs and promoter-targeted RNAs. *Cell. Mol. Life Sci.* **2012**, *69*, 2833–2842.
50. Rouget, C.; Papin, C.; Boureux, A.; Meunier, A.C.; Franco, B.; Robine, N.; Lai, E.C.; Pelisson, A.; Simonelig, M. Maternal mRNA deadenylation and decay by the piRNA pathway in the early Drosophila embryo. *Nature* **2010**, *467*, 1128–1132.
51. Lee, E.J.; Banerjee, S.; Zhou, H.; Jammalamadaka, A.; Arcila, M.; Manjunath, B.S.; Kosik, K.S. Identification of piRNAs in the central nervous system. *RNA* **2011**, *17*, 1090–1099.
52. Rajasethupathy, P.; Antonov, I.; Sheridan, R.; Frey, S.; Sander, C.; Tuschl, T.; Kandel, E.R. A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. *Cell* **2012**, *149*, 693–707.
53. Sienski, G.; Donertas, D.; Brennecke, J. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **2012**, *151*, 964–980.
54. Watanabe, T.; Tomizawa, S.; Mitsuya, K.; Totoki, Y.; Yamamoto, Y.; Kuramochi-Miyagawa, S.; Iida, N.; Hoki, Y.; Murphy, P.J.; Toyoda, A.; *et al.* Role for piRNAs and noncoding RNA in *de novo* DNA methylation of the imprinted mouse Rasgrf1 locus. *Science* **2011**, *332*, 848–852.
55. Gangaraju, V.K.; Yin, H.; Weiner, M.M.; Wang, J.; Huang, X.A.; Lin, H. *Drosophila* Piwi functions in Hsp90-mediated suppression of phenotypic variation. *Nat. genet.* **2011**, *43*, 153–158.
56. Rutherford, S.L.; Lindquist, S. Hsp90 as a capacitor for morphological evolution. *Nature* **1998**, *396*, 336–342.
57. Specchia, V.; Piacentini, L.; Tritto, P.; Fanti, L.; D'Alessandro, R.; Palumbo, G.; Pimpinelli, S.; Bozzetti, M.P. Hsp90 prevents phenotypic variation by suppressing the mutagenic activity of transposons. *Nature* **2010**, *463*, 662–665.
58. Kim, V.N. Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes. *Genes Dev.* **2006**, *20*, 1993–1997.
59. Siddiqi, S.; Matushansky, I. Piwis and piwi-interacting RNAs in the epigenetics of cancer. *J. Cell. Biochem.* **2012**, *113*, 373–380.
60. Siddiqi, S.; Terry, M.; Matushansky, I. Hiwi mediated tumorigenesis is associated with DNA hypermethylation. *PLoS One* **2012**, *7*, e33711.
61. Prescott, D.M., The DNA of ciliated protozoa. *Microbiol. Rev.* **1994**, *58*, 233–267.
62. Chalker, D.L. Dynamic nuclear reorganization during genome remodeling of *Tetrahymena*. *Biochim. Biophys. Acta* **2008**, *1783*, 2130–2136.
63. Duharcourt, S.; Lepere, G.; Meyer, E. Developmental genome rearrangements in ciliates: A natural genomic subtraction mediated by non-coding transcripts. *Trends Genet.* **2009**, *25*, 344–350.

64. Lee, J.T. Epigenetic regulation by long noncoding RNAs. *Science* **2012**, *338*, 1435–1439.
65. Pink, R.C.; Wicks, K.; Caley, D.P.; Punch, E.K.; Jacobs, L.; Carter, D.R. Pseudogenes: Pseudo-functional or key regulators in health and disease? *RNA* **2011**, *17*, 792–798.
66. Rackham, O.; Shearwood, A.M.; Mercer, T.R.; Davies, S.M.; Mattick, J.S.; Filipovska, A. Long noncoding RNAs are generated from the mitochondrial genome and regulated by nuclear-encoded proteins. *RNA* **2011**, *17*, 2085–2093.
67. Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M.C.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C.; *et al.* The transcriptional landscape of the mammalian genome. *Science* **2005**, *309*, 1559–1563.
68. Kapranov, P.; Cheng, J.; Dike, S.; Nix, D.A.; Duttagupta, R.; Willingham, A.T.; Stadler, P.F.; Hertel, J.; Hackermuller, J.; Hofacker, I.L.; *et al.* RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **2007**, *316*, 1484–1488.
69. Magistri, M.; Faghihi, M.A.; St Laurent, G., 3rd; Wahlestedt, C. Regulation of chromatin structure by long noncoding RNAs: Focus on natural antisense transcripts. *Trends Genet.* **2012**, *28*, 389–396.
70. Guttman, M.; Amit, I.; Garber, M.; French, C.; Lin, M.F.; Feldser, D.; Huarte, M.; Zuk, O.; Carey, B.W.; Cassady, J.P.; *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **2009**, *458*, 223–227.
71. Orom, U.A.; Derrien, T.; Beringer, M.; Gumireddy, K.; Gardini, A.; Bussotti, G.; Lai, F.; Zytnicki, M.; Notredame, C.; Huang, Q.; *et al.* Long noncoding RNAs with enhancer-like function in human cells. *Cell* **2010**, *143*, 46–58.
72. Kim, T.K.; Hemberg, M.; Gray, J.M.; Costa, A.M.; Bear, D.M.; Wu, J.; Harmin, D.A.; Laptewicz, M.; Barbara-Haley, K.; Kuersten, S.; *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **2010**, *465*, 182–187.
73. Katayama, S.; Tomaru, Y.; Kasukawa, T.; Waki, K.; Nakanishi, M.; Nakamura, M.; Nishida, H.; Yap, C.C.; Suzuki, M.; Kawai, J.; *et al.* Antisense transcription in the mammalian transcriptome. *Science* **2005**, *309*, 1564–1566.
74. Korneev, S.A.; Korneeva, E.I.; Lagarkova, M.A.; Kiselev, S.L.; Critchley, G.; O'Shea, M. Novel noncoding antisense RNA transcribed from human anti-NOS2A locus is differentially regulated during neuronal differentiation of embryonic stem cells. *RNA* **2008**, *14*, 2030–2037.
75. Lavorgna, G.; Dahary, D.; Lehner, B.; Sorek, R.; Sanderson, C.M.; Casari, G. In search of antisense. *Trends Biochem. Sci.* **2004**, *29*, 88–94.
76. Werner, A.; Carlile, M.; Swan, D. What do natural antisense transcripts regulate? *RNA Biol.* **2009**, *6*, 43–48.
77. He, Y.; Vogelstein, B.; Velculescu, V.E.; Papadopoulos, N.; Kinzler, K.W. The antisense transcriptomes of human cells. *Science* **2008**, *322*, 1855–1857.
78. Ravasi, T.; Suzuki, H.; Pang, K.C.; Katayama, S.; Furuno, M.; Okunishi, R.; Fukuda, S.; Ru, K.; Frith, M.C.; Gongora, M.M.; *et al.* Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res.* **2006**, *16*, 11–19.
79. Clemson, C.M.; Hutchinson, J.N.; Sara, S.A.; Ensminger, A.W.; Fox, A.H.; Chess, A.; Lawrence, J.B. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell* **2009**, *33*, 717–726.

80. Zhao, J.; Sun, B.K.; Erwin, J.A.; Song, J.J.; Lee, J.T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **2008**, *322*, 750–756.
81. Cai, X.; Cullen, B.R. The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* **2007**, *13*, 313–316.
82. Hung, T.; Chang, H.Y. Long noncoding RNA in genome regulation: Prospects and mechanisms. *RNA Biol.* **2010**, *7*, 582–585.
83. Mohammad, F.; Mondal, T.; Guseva, N.; Pandey, G.K.; Kanduri, C. Kcnq1ot1 noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1. *Development* **2010**, *137*, 2493–2499.
84. Tsai, M.C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J.K.; Lan, F.; Shi, Y.; Segal, E.; Chang, H.Y. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **2010**, *329*, 689–693.
85. Swiezewski, S.; Liu, F.; Magusin, A.; Dean, C. Cold-induced silencing by long antisense transcripts of an *Arabidopsis* polycomb target. *Nature* **2009**, *462*, 799–802.
86. Heo, J.B.; Sung, S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* **2011**, *331*, 76–79.
87. Hung, T.; Wang, Y.; Lin, M.F.; Koegel, A.K.; Kotake, Y.; Grant, G.D.; Horlings, H.M.; Shah, N.; Umbricht, C.; Wang, P.; *et al.* Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat. Genet*. **2011**, *43*, 621–629.
88. Bertani, S.; Sauer, S.; Bolotin, E.; Sauer, F. The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. *Mol. Cell* **2011**, *43*, 1040–1046.
89. Khalil, A.M.; Guttman, M.; Huarte, M.; Garber, M.; Raj, A.; Rivea Morales, D.; Thomas, K.; Presser, A.; Bernstein, B.E.; van Oudenaarden, A.; *et al.* Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11667–11672.
90. Dinger, M.E.; Amaral, P.P.; Mercer, T.R.; Pang, K.C.; Bruce, S.J.; Gardiner, B.B.; Askarian-Amiri, M.E.; Ru, K.; Solda, G.; Simons, C.; *et al.* Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res.* **2008**, *18*, 1433–1445.
91. Cabili, M.N.; Trapnell, C.; Goff, L.; Koziol, M.; Tazon-Vega, B.; Regev, A.; Rinn, J.L. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*. **2011**, *25*, 1915–1927.
92. Mercer, T.R.; Mattick, J.S. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat. Struct. Mol. Biol.* **2013**, *20*, 300–307.
93. Su, W.Y.; Xiong, H.; Fang, J.Y. Natural antisense transcripts regulate gene expression in an epigenetic manner. *Biochem. Biophys. Res. Commun.* **2010**, *396*, 177–181.
94. Imamura, T.; Yamamoto, S.; Ohgane, J.; Hattori, N.; Tanaka, S.; Shiota, K. Non-coding RNA directed DNA demethylation of Sphk1 CpG island. *Biochem. Biophys. Res. Commun.* **2004**, *322*, 593–600.
95. Berretta, J.; Pinskaya, M.; Morillon, A. A cryptic unstable transcript mediates transcriptional *trans*-silencing of the Ty1 retrotransposon in *S. cerevisiae*. *Genes Dev.* **2008**, *22*, 615–626.

96. Martianov, I.; Ramadass, A.; Serra Barros, A.; Chow, N.; Akoulitchev, A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* **2007**, *445*, 666–670.
97. Morris, K.V.; Chan, S.W.; Jacobsen, S.E.; Looney, D.J. Small interfering RNA-induced transcriptional gene silencing in human cells. *Science* **2004**, *305*, 1289–1292.
98. Napoli, S.; Pastori, C.; Magistri, M.; Carbone, G.M.; Catapano, C.V. Promoter-specific transcriptional interference and c-myc gene silencing by siRNAs in human cells. *EMBO J.* **2009**, *28*, 1708–1719.
99. Hawkins, P.G.; Santoso, S.; Adams, C.; Anest, V.; Morris, K.V. Promoter targeted small RNAs induce long-term transcriptional gene silencing in human cells. *Nucleic Acids Res.* **2009**, *37*, 2984–2995.
100. Place, R.F.; Li, L.C.; Pookot, D.; Noonan, E.J.; Dahiya, R. microRNA-373 induces expression of genes with complementary promoter sequences. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1608–1613.
101. Li, L.C.; Okino, S.T.; Zhao, H.; Pookot, D.; Place, R.F.; Urakami, S.; Enokida, H.; Dahiya, R. Small dsRNAs induce transcriptional activation in human cells. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 17337–17342.
102. Janowski, B.A.; Younger, S.T.; Hardy, D.B.; Ram, R.; Huffman, K.E.; Corey, D.R. Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nat. Chem. Biol.* **2007**, *3*, 166–173.
103. Natoli, G.; Andrau, J.C. Noncoding transcription at enhancers: General principles and functional models. *Annu. Rev. Genet.* **2012**, *46*, 1–19.
104. Orom, U.A.; Shiekhattar, R. Long non-coding RNAs and enhancers. *Curr. Opin. Genet. Dev.* **2011**, *21*, 194–198.
105. Panganiban, G.; Rubenstein, J.L. Developmental functions of the Distal-less/Dlx homeobox genes. *Development* **2002**, *129*, 4371–4386.
106. Wang, X.; Arai, S.; Song, X.; Reichart, D.; Du, K.; Pascual, G.; Tempst, P.; Rosenfeld, M.G.; Glass, C.K.; Kurokawa, R. Induced ncRNAs allosterically modify RNA-binding proteins in *cis* to inhibit transcription. *Nature* **2008**, *454*, 126–130.
107. Goodrich, J.A.; Kugel, J.F. Non-coding-RNA regulators of RNA polymerase II transcription. *Nat. Rev.* **2006**, *7*, 612–616.
108. Hastings, M.L.; Ingle, H.A.; Lazar, M.A.; Munroe, S.H. Post-transcriptional regulation of thyroid hormone receptor expression by *cis*-acting sequences and a naturally occurring antisense RNA. *J. Boil. Chem.* **2000**, *275*, 11507–11513.
109. Yin, Q.F.; Yang, L.; Zhang, Y.; Xiang, J.F.; Wu, Y.W.; Carmichael, G.G.; Chen, L.L. Long noncoding RNAs with snoRNA ends. *Mol. Cell* **2012**, *48*, 219–230.
110. Gong, C.; Maquat, L.E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **2011**, *470*, 284–288.
111. Kretz, M.; Siprashvili, Z.; Chu, C.; Webster, D.E.; Zehnder, A.; Qu, K.; Lee, C.S.; Flockhart, R.J.; Groff; A.F.; Chow, J.; *et al.* Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **2013**, *493*, 231–235.

112. Faghihi, M.A.; Modarresi, F.; Khalil, A.M.; Wood, D.E.; Sahagan, B.G.; Morgan, T.E.; Finch, C.E.; St Laurent, G., 3rd; Kenny, P.J.; Wahlestedt, C. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat. Med.* **2008**, *14*, 723–730.
113. Ebralidze, A.K.; Guibal, F.C.; Steidl, U.; Zhang, P.; Lee, S.; Bartholdy, B.; Jorda, M.A.; Petkova, V.; Rosenbauer, F.; Huang, G.; *et al.* PU.1 expression is modulated by the balance of functional sense and antisense RNAs regulated by a shared *cis*-regulatory element. *Genes Dev.* **2008**, *22*, 2085–2092.
114. Carrieri, C.; Cimatti, L.; Biagioli, M.; Beugnet, A.; Zucchelli, S.; Fedele, S.; Pesce, E.; Ferrer, I.; Collavin, L.; Santoro, C.; *et al.* Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* **2012**, *491*, 454–457.
115. Willingham, A.T.; Orth, A.P.; Batalov, S.; Peters, E.C.; Wen, B.G.; Aza-Blanc, P.; Hogenesch, J.B.; Schultz, P.G. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **2005**, *309*, 1570–1573.
116. Guo, X.; Zhang, Z.; Gerstein, M.B.; Zheng, D. Small RNAs originated from pseudogenes: *cis*- or *trans*-acting? *PLoS Comput. Biol.* **2009**, *5*, e1000449.
117. Tam, O.H.; Aravin, A.A.; Stein, P.; Girard, A.; Murchison, E.P.; Cheloufi, S.; Hodges, E.; Anger, M.; Sachidanandam, R.; Schultz, R.M.; *et al.* Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **2008**, *453*, 534–538.
118. Watanabe, T.; Totoki, Y.; Toyoda, A.; Kaneda, M.; Kuramochi-Miyagawa, S.; Obata, Y.; Chiba, H.; Kohara, Y.; Kono, T.; Nakano, T.; *et al.* Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **2008**, *453*, 539–543.
119. Mighell, A.J.; Smith, N.R.; Robinson, P.A.; Markham, A.F. Vertebrate pseudogenes. *FEBS Lett.* **2000**, *468*, 109–114.
120. Zhang, Z.D.; Frankish, A.; Hunt, T.; Harrow, J.; Gerstein, M. Identification and analysis of unitary pseudogenes: Historic and contemporary gene losses in humans and other primates. *Genome Biol.* **2010**, *11*, R26.
121. Maestre, J.; Tchenio, T.; Dhellin, O.; Heidmann, T. mRNA retroposition in human cells: Processed pseudogene formation. *EMBO J.* **1995**, *14*, 6333–6338.
122. D'Errico, I.; Gadaleta, G.; Saccone, C. Pseudogenes in metazoa: Origin and features. *Briefings Funct. Genomics Proteomics* **2004**, *3*, 157–167.
123. Werner, A.; Schmutzler, G.; Carlile, M.; Miles, C.G.; Peters, H. Expression profiling of antisense transcripts on DNA arrays. *Physiol. Genomics* **2007**, *28*, 294–300.
124. Fanarraga, M.L.; Parraga, M.; Aloria, K.; del Mazo, J.; Avila, J.; Zabala, J.C. Regulated expression of p14 (cofactor A) during spermatogenesis. *Cell Motil. Cytoskeleton* **1999**, *43*, 243–254.
125. Pavlicek, A.; Gentles, A.J.; Paces, J.; Paces, V.; Jurka, J. Retroposition of processed pseudogenes: The impact of RNA stability and translational control. *Trends Genet.* **2006**, *22*, 69–73.
126. Borsani, O.; Zhu, J.; Verslues, P.E.; Sunkar, R.; Zhu, J.K. Endogenous siRNAs derived from a pair of natural *cis*-antisense transcripts regulate salt tolerance in *Arabidopsis*. *Cell* **2005**, *123*, 1279–1291.

127. Okada, Y.; Tashiro, C.; Numata, K.; Watanabe, K.; Nakaoka, H.; Yamamoto, N.; Okubo, K.; Ikeda, R.; Saito, R.; Kanai, A.; *et al.* Comparative expression analysis uncovers novel features of endogenous antisense transcription. *Hum. Mol. Genet.* **2008**, *17*, 1631–1640.
128. Okamoto, I.; Arnaud, D.; le Baccon, P.; Otte, A.P.; Disteche, C.M.; Avner, P.; Heard, E. Evidence for *de novo* imprinted X-chromosome inactivation independent of meiotic inactivation in mice. *Nature* **2005**, *438*, 369–373.
129. Young, T.L.; Matsuda, T.; Cepko, C.L. The noncoding RNA taurine upregulated gene 1 is required for differentiation of the murine retina. *Curr. Biol.* **2005**, *15*, 501–512.
130. Ginger, M.R.; Shore, A.N.; Contreras, A.; Rijnkels, M.; Miller, J.; Gonzalez-Rimbau, M.F.; Rosen, J.M. A noncoding RNA is a potential marker of cell fate during mammary gland development. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5781–5786.
131. Rinn, J.L.; Kertesz, M.; Wang, J.K.; Squazzo, S.L.; Xu, X.; Brugmann, S.A.; Goodnough, L.H.; Helms, J.A.; Farnham, P.J.; Segal, E.; *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **2007**, *129*, 1311–1323.
132. Loewer, S.; Cabili, M.N.; Guttman, M.; Loh, Y.H.; Thomas, K.; Park, I.H.; Garber, M.; Curran, M.; Onder, T.; Agarwal, S.; *et al.* Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat. Genet.* **2010**, *42*, 1113–1117.
133. Kierszenbaum, A.L.; Tres, L.L. The acrosome-acroplaxome-manchette complex and the shaping of the spermatid head. *Arch. Histol. Cytol.* **2004**, *67*, 271–284.
134. Nolasco, S.; Bellido, J.; Goncalves, J.; Tavares, A.; Zabala, J.C.; Soares, H. The expression of tubulin cofactor A (TBCA) is regulated by a noncoding antisense Tbca RNA during testis maturation. *PLoS One* **2012**, *7*, e42536.
135. Bartel, D.P. microRNAs: Target recognition and regulatory functions. *Cell* **2009**, *136*, 215–233.
136. Lim, L.P.; Lau, N.C.; Garrett-Engele, P.; Grimson, A.; Schelter, J.M.; Castle, J.; Bartel, D.P.; Linsley, P.S.; Johnson, J.M. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **2005**, *433*, 769–773.
137. Jing, Q.; Huang, S.; Guth, S.; Zarubin, T.; Motoyama, A.; Chen, J.; di Padova, F.; Lin, S.C.; Gram, H.; Han, J. Involvement of microRNA in AU-rich element-mediated mRNA instability. *Cell* **2005**, *120*, 623–634.
138. Behm-Ansmant, I.; Rehwinkel, J.; Doerks, T.; Stark, A.; Bork, P.; Izaurralde, E. mRNA degradation by miRNAs and GW182 requires both CCR4: NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev.* **2006**, *20*, 1885–1898.
139. Llave, C.; Xie, Z.; Kasschau, K.D.; Carrington, J.C. Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* **2002**, *297*, 2053–2056.
140. Kim, D.H.; Saetrom, P.; Snove, O., Jr.; Rossi, J.J. microRNA-directed transcriptional gene silencing in mammalian cells. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 16230–16235.
141. Liu, J.; Valencia-Sanchez, M.A.; Hannon, G.J.; Parker, R. microRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat. Cell Biol.* **2005**, *7*, 719–723.
142. Pillai, R.S.; Bhattacharyya, S.N.; Artus, C.G.; Zoller, T.; Cougot, N.; Basyuk, E.; Bertrand, E.; Filipowicz, W. Inhibition of translational initiation by Let-7 microRNA in human cells. *Science* **2005**, *309*, 1573–1576.

143. Humphreys, D.T.; Westman, B.J.; Martin, D.I.; Preiss, T. microRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16961–16966.
144. Kiriakidou, M.; Tan, G.S.; Lamprinaki, S.; de Planell-Saguer, M.; Nelson, P.T.; Mourelatos, Z. An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell* **2007**, *129*, 1141–1151.
145. Eulalio, A.; Huntzinger, E.; Izaurralde, E. GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. *Nat. Struct. Mol. Biol.* **2008**, *15*, 346–353.
146. Wang, B.; Yanez, A.; Novina, C.D. microRNA-repressed mRNAs contain 40S but not 60S components. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5343–5348.
147. Thermann, R.; Hentze, M.W. *Drosophila* miR2 induces pseudo-polysomes and inhibits translation initiation. *Nature* **2007**, *447*, 875–878.
148. Maroney, P.A.; Yu, Y.; Fisher, J.; Nilsen, T.W. Evidence that microRNAs are associated with translating messenger RNAs in human cells. *Nat. Struct. Mol. Biol.* **2006**, *13*, 1102–1107.
149. Petersen, C.P.; Bordeleau, M.E.; Pelletier, J.; Sharp, P.A. Short RNAs repress translation after initiation in mammalian cells. *Mol. Cell* **2006**, *21*, 533–542.
150. Lytle, J.R.; Yario, T.A.; Steitz, J.A. Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 9667–9672.
151. Vasudevan, S.; Steitz, J.A. AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2. *Cell* **2007**, *128*, 1105–1118.
152. Orom, U.A.; Nielsen, F.C.; Lund, A.H. microRNA-10a binds the 5' UTR of ribosomal protein mRNAs and enhances their translation. *Mol. Cell* **2008**, *30*, 460–471.
153. Lee, S.; Vasudevan, S. Post-transcriptional stimulation of gene expression by microRNAs. *Adv. Exper. Med. Biol.* **2013**, *768*, 97–126.
154. Bernstein, E.; Kim, S.Y.; Carmell, M.A.; Murchison, E.P.; Alcorn, H.; Li, M.Z.; Mills, A.A.; Elledge, S.J.; Anderson, K.V.; Hannon, G.J. Dicer is essential for mouse development. *Nat. Genet.* **2003**, *35*, 215–217.
155. Leonardo, T.R.; Schultheisz, H.L.; Loring, J.F.; Laurent, L.C. The functions of microRNAs in pluripotency and reprogramming. *Nat. Cell Biol.* **2012**, *14*, 1114–1121.
156. Giraldez, A.J.; Cinalli, R.M.; Glasner, M.E.; Enright, A.J.; Thomson, J.M.; Baskerville, S.; Hammond, S.M.; Bartel, D.P.; Schier, A.F. microRNAs regulate brain morphogenesis in zebrafish. *Science* **2005**, *308*, 833–838.
157. Stark, K.L.; Xu, B.; Bagchi, A.; Lai, W.S.; Liu, H.; Hsu, R.; Wan, X.; Pavlidis, P.; Mills, A.A.; Karayiorgou, M.; *et al.* Altered brain microRNA biogenesis contributes to phenotypic deficits in a 22q11-deletion mouse model. *Nat. Genet.* **2008**, *40*, 751–760.
158. Cobb, B.S.; Nesterova, T.B.; Thompson, E.; Hertweck, A.; O'Connor, E.; Godwin, J.; Wilson, C.B.; Brockdorff, N.; Fisher, A.G.; Smale, S.T.; *et al.* T cell lineage choice and differentiation in the absence of the RNase III enzyme Dicer. *J. Exp. Med.* **2005**, *201*, 1367–1373.
159. Muljo, S.A.; Ansel, K.M.; Kanellopoulou, C.; Livingston, D.M.; Rao, A.; Rajewsky, K. Aberrant T cell differentiation in the absence of Dicer. *J. Exp. Med.* **2005**, *202*, 261–269.

160. Edbauer, D.; Neilson, J.R.; Foster, K.A.; Wang, C.F.; Seeburg, D.P.; Batterton, M.N.; Tada, T.; Dolan, B.M.; Sharp, P.A.; Sheng, M. Regulation of synaptic structure and function by FMRP-associated microRNAs miR-125b and miR-132. *Neuron* **2010**, *65*, 373–384.
161. Monticelli, S.; Ansel, K.M.; Xiao, C.; Socci, N.D.; Krichevsky, A.M.; Thai, T.H.; Rajewsky, N.; Marks, D.S.; Sander, C.; Rajewsky, K.; *et al.* microRNA profiling of the murine hematopoietic system. *Genome Biol.* **2005**, *6*, R71.
162. Rodriguez, A.; Vigorito, E.; Clare, S.; Warren, M.V.; Couttet, P.; Soond, D.R.; van Dongen, S.; Grocock, R.J.; Das, P.P.; Miska, E.A.; *et al.* Requirement of bic/microRNA-155 for normal immune function. *Science* **2007**, *316*, 608–611.
163. Boldin, M.P.; Taganov, K.D.; Rao, D.S.; Yang, L.; Zhao, J.L.; Kalwani, M.; Garcia-Flores, Y.; Luong, M.; Devrekanli, A.; Xu, J.; *et al.* miR-146a is a significant brake on autoimmunity, myeloproliferation, and cancer in mice. *J. Exp. Med.* **2011**, *208*, 1189–1201.
164. Steiner, D.F.; Thomas, M.F.; Hu, J.K.; Yang, Z.; Babiarz, J.E.; Allen, C.D.; Matloubian, M.; Blelloch, R.; Ansel, K.M. microRNA-29 regulates T-box transcription factors and interferon-gamma production in helper T cells. *Immunity* **2011**, *35*, 169–181.
165. Cheetham, S.W.; Gruhl, F.; Mattick, J.S.; Dinger, M.E. Long noncoding RNAs and the genetics of cancer. *Br. J. Cancer* **2013**, *108*, 2419–2425.
166. Munker, R.; Calin, G.A. microRNA profiling in cancer. *Clin. Sci.* **2011**, *121*, 141–158.
167. Faraoni, I.; Antonetti, F.R.; Cardone, J.; Bonmassar, E. miR-155 gene: A typical multifunctional microRNA. *Biochim. Biophys. Acta* **2009**, *1792*, 497–505.
168. Chan, J.A.; Krichevsky, A.M.; Kosik, K.S. microRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res.* **2005**, *65*, 6029–6033.
169. Volinia, S.; Calin, G.A.; Liu, C.G.; Ambs, S.; Cimmino, A.; Petrocca, F.; Visone, R.; Iorio, M.; Roldo, C.; Ferracin, M.; *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 2257–2261.
170. Xiao, C.; Srinivasan, L.; Calado, D.P.; Patterson, H.C.; Zhang, B.; Wang, J.; Henderson, J.M.; Kutok, J.L.; Rajewsky, K. Lymphoproliferative disease and autoimmunity in mice with increased miR-17-92 expression in lymphocytes. *Nat. Immunol.* **2008**, *9*, 405–414.
171. Concepcion, C.P.; Bonetti, C.; Ventura, A. The microRNA-17–92 family of microRNA clusters in development and disease. *Cancer J.* **2012**, *18*, 262–267.
172. Tripathi, V.; Ellis, J.D.; Shen, Z.; Song, D.Y.; Pan, Q.; Watt, A.T.; Freier, S.M.; Bennett, C.F.; Sharma, A.; Bubulya, P.A.; *et al.* The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* **2010**, *39*, 925–938.
173. Gutschner, T.; Hammerle, M.; Eissmann, M.; Hsu, J.; Kim, Y.; Hung, G.; Revenko, A.; Arun, G.; Stentrup, M.; Gross, M.; *et al.* The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res.* **2013**, *73*, 1180–1189.
174. Drakaki, A.; Iliopoulos, D. microRNA gene networks in oncogenesis. *Curr. Genomics* **2009**, *10*, 35–41.
175. Dalmay, T.; Edwards, D.R. microRNAs and the hallmarks of cancer. *Oncogene* **2006**, *25*, 6170–6175.
176. Gutschner, T.; Diederichs, S. The hallmarks of cancer: A long non-coding RNA point of view. *RNA Biol.* **2012**, *9*, 703–719.

177. He, L.; He, X.; Lim, L.P.; de Stanchina, E.; Xuan, Z.; Liang, Y.; Xue, W.; Zender, L.; Magnus, J.; Ridzon, D.; *et al.* A microRNA component of the p53 tumour suppressor network. *Nature* **2007**, *447*, 1130–1134.
178. Concepcion, C.P.; Han, Y.C.; Mu, P.; Bonetti, C.; Yao, E.; D'Andrea, A.; Vidigal, J.A.; Maughan, W.P.; Ogrodowski, P.; Ventura, A. Intact p53-dependent responses in miR-34-deficient mice. *PLoS Genet.* **2012**, *8*, e1002797.
179. Huarte, M.; Guttman, M.; Feldser, D.; Garber, M.; Koziol, M.J.; Kenzelmann-Broz, D.; Khalil, A.M.; Zuk, O.; Amit, I.; Rabani, M.; *et al.* A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **2010**, *142*, 409–419.
180. Redon, S.; Reichenbach, P.; Lingner, J. The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. *Nucleic Acids Res.* **2010**, *38*, 5797–5806.
181. Packer, A.N.; Xing, Y.; Harper, S.Q.; Jones, L.; Davidson, B.L. The bifunctional microRNA miR-9/miR-9* regulates REST and CoREST and is downregulated in Huntington's disease. *J. Neurosci.* **2008**, *28*, 14341–14346.
182. Cogswell, J.P.; Ward, J.; Taylor, I.A.; Waters, M.; Shi, Y.; Cannon, B.; Kelnar, K.; Kemppainen, J.; Brown, D.; Chen, C.; *et al.* Identification of miRNA changes in Alzheimer's disease brain and CSF yields putative biomarkers and insights into disease pathways. *J. Alzheimer's Disease* **2008**, *14*, 27–41.
183. Kim, J.; Inoue, K.; Ishii, J.; Vanti, W.B.; Voronov, S.V.; Murchison, E.; Hannon, G.; Abeliovich, A. A microRNA feedback circuit in midbrain dopamine neurons. *Science* **2007**, *317*, 1220–1224.
184. Contreras, J.; Rao, D.S. microRNAs in inflammation and immune responses. *Leukemia* **2012**, *26*, 404–413.
185. Cabianca, D.S.; Casa, V.; Bodega, B.; Xynos, A.; Ginelli, E.; Tanaka, Y.; Gabellini, D. A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* **2012**, *149*, 819–831.
186. Szafranski, P.; Dharmadhikari, A.V.; Brosens, E.; Gurha, P.; Kolodziejska, K.E.; Zhishuo, O.; Dittwald, P.; Majewski, T.; Mohan, K.N.; Chen, B.; *et al.* Small noncoding differentially methylated copy-number variants, including lncRNA genes, cause a lethal lung developmental disorder. *Genome Res*. **2013**, *23*, 23–33.
187. Van Dijk, M.; Thulluru, H.K.; Mulders, J.; Michel, O.J.; Poutsma, A.; Windhorst, S.; Kleiverda, G.; Sie, D.; Lachmeijer, A.M.; Oudejans, C.B. HELLP babies link a novel lincRNA to the trophoblast cell cycle. *J. Clin. Invest.* **2012**, *122*, 4003–4011.

Reprinted from *IJMS*. Cite as: Bussotti, G.; Notredame, C.; Enright, A.J. Detecting and Comparing Non-Coding RNAs in the High-Throughput Era. *Int. J. Mol. Sci.* **2013**, *14*, 15423-15458.

Review

Detecting and Comparing Non-Coding RNAs in the High-Throughput Era

Giovanni Bussotti [1,*], Cedric Notredame [2] and Anton J. Enright [1]

[1] European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK; E-Mail: aje@ebi.ac.uk

[2] Bioinformatics and Genomics Program, Centre for Genomic Regulation (CRG), Aiguader, 88, 08003 Barcelona, Spain; E-Mail: Cedric.Notredame@crg.eu

***** Author to whom correspondence should be addressed; E-Mail: giovanni.bussotti@ebi.ac.uk; Tel.: +44-1223-49-2680 (ext. 2680); Fax: +44-1223-49-4486.

Received: 31 May 2013; in revised form: 16 July 2013 / Accepted: 17 July 2013 / Published: 24 July 2013

Abstract: In recent years there has been a growing interest in the field of non-coding RNA. This surge is a direct consequence of the discovery of a huge number of new non-coding genes and of the finding that many of these transcripts are involved in key cellular functions. In this context, accurately detecting and comparing RNA sequences has become important. Aligning nucleotide sequences is a key requisite when searching for homologous genes. Accurate alignments reveal evolutionary relationships, conserved regions and more generally any biologically relevant pattern. Comparing RNA molecules is, however, a challenging task. The nucleotide alphabet is simpler and therefore less informative than that of amino-acids. Moreover for many non-coding RNAs, evolution is likely to be mostly constrained at the structural level and not at the sequence level. This results in very poor sequence conservation impeding comparison of these molecules. These difficulties define a context where new methods are urgently needed in order to exploit experimental results to their full potential. This review focuses on the comparative genomics of non-coding RNAs in the context of new sequencing technologies and especially dealing with two extremely important and timely research aspects: the development of new methods to align RNAs and the analysis of high-throughput data.

Keywords: lncRNAs; ncRNAs; high-throughput; RNA-seq; comparative biology; sequencing

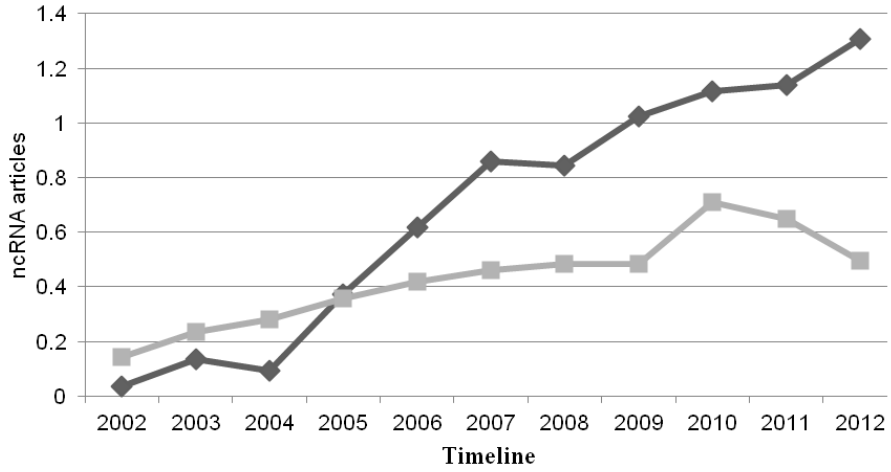
---

1. Introduction

1.1. The Non-Coding RNA (New)-World

In recent years, the non-coding RNA (ncRNA) field has rapidly expanded (Figure 1) with a rapid increase in the number of newly identified and biologically relevant ncRNAs. Just a decade ago, the number of known ncRNAs was restricted to a small amount of housekeeping genes (including ribosomal RNAs, transfer RNAs and small nucleolar RNAs) and an even more limited collection of regulatory RNAs, such as lin-4 in *Caenorhabditis elegans* [1] and Xist in mammals [2]. Since then, the number of novel ncRNAs has increased dramatically and far more is known about their function, biogenesis, length, structural and sequence features. New and ever more sophisticated high-throughput technologies, such as tiling arrays and next generation sequencing (NGS) have been applied to comprehensively profile the transcriptome of various organisms.

Figure 1. Number of publications in PubMed found using the keyword "ncRNA" (dark grey) and "regulatory RNA" (pale gray). The x-axis represents the timeline, the y-axis the number of times the words "ncRNA" and "regulatory RNA" match a publication in PubMed normalized by the total number of publications in that year (expressed as one part per ten thousand).



This wealth of data has allowed the identification of thousands of novel short ncRNAs, including PIWI interacting RNAs [3] and small nucleolar RNAs [4] and has resulted in the compilation or the update of many publicly available databases [5–10]. Furthermore, high-throughput approaches have revealed extensive and pervasive transcription of long ncRNAs (lncRNAs) [11–13], operationally defined as functional RNA longer than 200 base pairs that does not template protein synthesis. In the human genome, for instance, the GENCODE consortium annotated 9,640 lncRNA loci representing 15,512 transcripts [3,14] and in [15] the authors estimated that total number of human lncRNAs genes to be about 50,000, more than two-fold greater than the number of protein-coding genes. These discoveries were very timely in the context of growing concern for the lack of a significant correlation between the number of protein-coding genes and the commonly accepted concept of "organism complexity" [4,16,17]. It was proposed that alternative splicing and ncRNAs could be accountable for complex gene regulation architectures, meaning that the "Central Dogma" of genetic programming

enunciated by Francis Crick in 1958 (RNA is transcribed from DNA and translated into protein) [18] had to be slightly altered and at least in higher eukaryotes may be inadequate [16,17]. The biological role of most of these novel long untranslated molecules is still a controversial issue. Some authors have even raised doubts about whether these transcripts are functional at all [19]. The lack of shared discernible features hampers our ability to define lncRNA classes, thus impeding function prediction [20]. However mounting experimental evidence illustrates that lncRNAs are implicated in a variety of biological processes [21] and are linked to various diseases including cancer [22]. Additionally, the functional roles of lncRNA transcripts have been uncovered in signaling sensors [23], embryonic stem cell differentiation [11], brain function [24,25], subcellular compartmentalization and chromatin remodeling [26]. Some examples include X chromosome inactivation by Xist, the silencing of autosomal imprinted genes accomplished by Air, nuclear trafficking regulated by NRON and muscle differentiation controlled by linc-MD1 [2,27–29]. In [30] the authors identified a class of lncRNAs named ncRNA-a (ncRNA-activator) able to stimulate the expression of proximal protein-coding genes, and a recent update on ncRNA-a [31] showed that the co-activator complex Mediator plays a central role in the activation process. See [21] and [32] for more examples and lncRNAdb [33] for the central repository of known lncRNAs in eukaryotes. lncRNAs are expressed, some are spliced, they are often conserved across vertebrates, and their expression is frequently tissue- and/or cell-specific and localized to specific subcellular compartments [11,25,34]. It has been shown that lncRNAs can act both in *cis* [30,35] and in *trans* [36], some acting as precursors for short ncRNAs [37–39], while others act independently as long transcripts. As in [40] lncRNAs can be classified as “intergenic” or “genic” depending on their position/orientation with respect to protein-coding genes. lncRNAs not overlapping any protein-coding gene are tagged as intergenic and then further classified according to their transcription orientation with the closest protein-coding loci (same sense, convergent, or divergent). The genic lncRNA set are catalogued as “exonic” if overlapping a protein-coding exon. Otherwise, lncRNAs are labeled as “intronic”, when positioned within protein-coding introns or as “overlapping”, in presence of a protein-coding transcript located within the intron of the lncRNA [40].

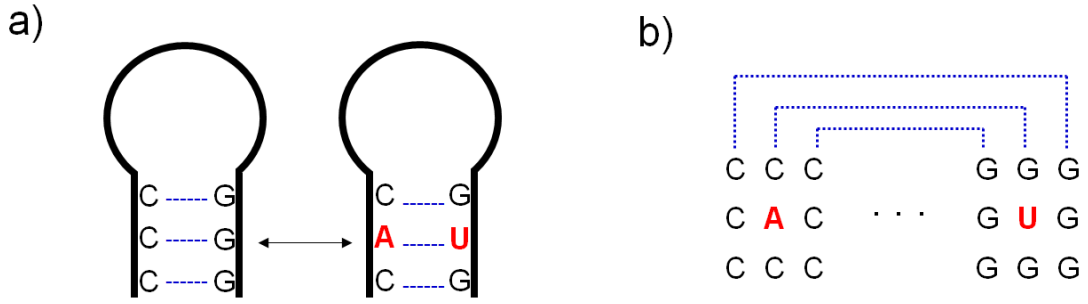
1.2. lncRNA Challenges

Although the conservation level of different lncRNAs may be not always directly comparable (e.g., the evolutionary conservation of genic lncRNAs may be biased by the presence of the protein-coding genes), overall approximately half of reported human lncRNA exhibit significant conservation across mammals [40]. These levels suggest some key cellular function, even though only a small fraction of these transcripts have so far been functionally characterized. Such functional analyses remain however, very superficial and lack precise molecular mechanisms explaining the activity of these novel transcripts. Our low level of understanding can be in part attributed to the difficulty with working experimentally with lncRNAs: detection is difficult for a combination of biological and technical aspects. The first relates to the low levels of non-coding genetic expression. After ribosomal RNA (rRNA), protein-coding mRNA represents the highest population of RNA species [41]. In previous studies [34,42,43] it has been reported that lncRNAs are on average 3 to 10 fold less expressed than mRNAs. Besides the complicated task of capturing weaker expression signals, many lncRNAs have pronounced tissue/stage specificity [43,44]. In other words, lncRNA genes can easily be left

undetected unless the correct cell type and condition are considered. One more complication for ncRNA discovery has been the difficulty of sequencing deep enough, a hurdle only recently overcome by NGS. Additionally, our ability to assemble and annotate genomes was less advanced than currently and we had simplified notions of transcriptome complexity. Most of the classical low-throughput approaches, such as RT-PCR and northern blotting, have been successfully used to analyze the expression of small numbers of genes, but they were not adequate to address the "pervasive transcription" aspect of genomes [45,46]. Furthermore, there are specific classes of ncRNAs, such as circular RNAs (circRNAs), that have been extremely hard to identify. circRNAs are a class of non-coding RNA family that were discovered more than 20 years ago [47–50]. These RNAs form circles that arise from non-canonical splicing events (also known as exon shuffling) that join a splice donor to an upstream splice acceptor to produce a circular RNA molecule. Recent studies [51,52] show that the human circRNA CDR1as, antisense to the Cerebellar Degeneration-Related protein 1 (CDR1), hosts around 70 binding sites for the miR-7 microRNA and is highly associated with the Argonaute protein Ago2 as demonstrated by PAR-CLIP and HITS-CLIP experiments [51,52]. Mainly because of their non-canonical splicing behavior, circRNAs have eluded detection by next generation sequencing until recently. These latest studies adopted a novel computational approach to identify circRNAs from high-throughput RNA-seq data and demonstrated their widespread abundance within transcriptomes [51,53].

In general, a major obstacle for ncRNA detection is the difficulty to perform informative sequence comparisons. Standard primary sequence alignment is hampered by the low complexity of the nucleic alphabet, making it difficult to produce statistically meaningful RNA alignments. Ribonucleic acid chemistry relies on just four primary residues: two purines and two pyrimidines. Consequently, RNA gene sequences do not have strong statistical signals, unlike protein-coding genes. For instance two RNA sequences must share an identity of at least ~60% to be considered significant in homology relationships prediction [54]. Below this level, common ancestry is hard to infer with certainty. By comparison, this threshold is around ~20%–35% for proteins [55]. Furthermore, ncRNA appears to be evolving rapidly [56] or are under the influence of very specific evolutionary constraints [56]. It was proposed that most ncRNAs evolve at higher mutation rates, with the maintenance of secondary structures being the main source of selection [57,58]. This assumption makes sense from an evolutionary standpoint. As ncRNAs will be left untranslated, the nucleotide sequence itself is not constrained to keep the codon reading frame. Of course many exceptions exist. Specific ncRNAs types can hold functional sequences and act via their primary sequence (*i.e.*, miRNAs). Previous reports have shown that at least some miRNA genes are well conserved across species [59–61], reinforcing the idea that sequences encoding a function evolve under purifying selection. Aside from these specific and relatively rare examples, it seems that for most known ncRNAs, evolution is limited by structural constraints [62,63]. This induces a characteristic pattern of covariance that occurs when a mutation is affecting a nucleotide pairing to another in a structured domain (Figure 2). If the mutation breaks the base pairing so that the functionality of such a domain is compromised, the matching nucleotide is favored to mutate in turn, *i.e.*, is co-varying to restore the base pairing and keep the structure unchanged.

Figure 2. RNA mutations are tightly linked to the RNA structure conservation. (**a**) Example where the mutation of a C into an A is compensated by the change G-U. The two positions are not independent, but communicating one with the other to maintain the structure unvaried; (**b**) Same hairpin as shown in (a). The presence of the compensatory mutation is highlighted by the multiple sequence comparison.



For most aligners these features of RNA are hard to account for when using standard alignment procedures that postulate positional independence and seek only to maximize identity. Furthermore, RNA can hold functional pseudo-knots. These are structural configurations where at least two RNA stem-loops are interposed one into the other. Although some comparative approaches including pseudo-knots exist [64,65], these are disregarded by most software due to reasons of computational complexity [66]. As a consequence ncRNA sequences are harder to align than proteins, a limitation that affects our ability to accurately detect and classify them. The difficulty in comparing ncRNAs calls for other information sources that alignment algorithms can use. More than ever, the issue of accurately comparing and aligning ncRNAs is of critical importance. This is precisely the problem discussed in the following section, where we review established and more recent methodologies able to make the best of available RNA information (Section 2). Next we discuss different homology based strategies for ncRNA detection (Section 3) and the analysis of high-throughput expression data (Section 4). See Table 1 for a summary of the resources described in the text.

2. Comparing Non-Coding RNAs

As mentioned, generating meaningful ncRNA alignments is a challenging task and at least in some cases, the best accuracy could be achieved by exploiting RNA structural information. However, in many situations using such information is complicated. In spite of the development of aligners that take into account the RNA secondary structure information, one major issue is the poor availability of high quality structures. The problem is at least in part due to the difficulties encountered at experimental level in crystallization. Getting crystals from RNA molecules is complicated because of their chemical specificity. The accumulation of crystals is prevented by the high RNA flexibility. RNAs have flexible structures adopting inter-domain movements and with respect to proteins have weaker tertiary interactions [67]. The polyanionic charge of the phosphate backbone makes the nucleotide sequence move much more than in proteins and this makes the packaging of crystals much harder to achieve. As a consequence, the crystals are either hard to grow or uninformative. Even when trying to resolve RNA molecules in solution using NMR, the resonance assignment is more difficult for RNA than for proteins [68]. RNAs have only 4 primary nucleosides instead of the 20 different

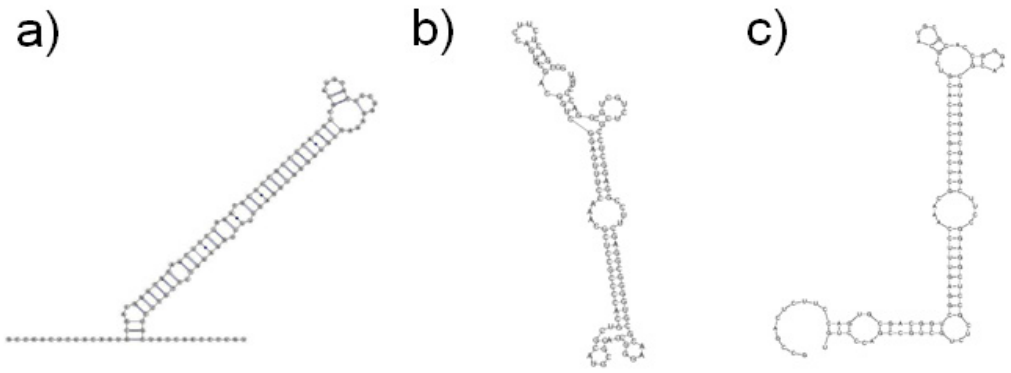
amino-acid side chains found in proteins [69]. Thus, the chemical shift dispersion is narrower in RNA than in proteins, resulting in less informative spectra [69].

2.1. RNA Structure Prediction

Because of these limitations, RNA structure is usually computationally predicted without any experimental support [70,71]. RNA secondary structure inference amounts to the computation of base-pairings that shape the *in vivo* molecule structure. The prediction is performed using primary RNA sequence data. Another possibility is including other sources of statistical information to constrain structure prediction, for instance an alignment of structurally homologous RNA sequences. Regarding single sequence RNA secondary structure predictions, there are two main groups of approaches: empirical free-energy parameters [72] and knowledge based [73–75]. The first considers a biophysical model to calculate the structure whose folding has the minimum Gibbs free energy (ΔG). In this approach, [76–80] the nearest stable folding is employed to compute the conformational stability of the Minimum Free Energy (MFE) structure. The energy parameters needed in this approach were assessed on a set of optical melting experiments on model systems [77–79]. The two most popular implementations of the MFE structure prediction algorithm are Mfold [70] and RNAfold [81] packages. The latter implements McCaskill's algorithm [82], an approach to calculate the probability of a certain secondary structure in the whole thermodynamic ensemble. This approach is based on the partition function, which sums all Boltzmann weighted free energies of each secondary structure that is possible given an RNA sequence. In this model, the confidence estimate in a particular base pair i,j is given by the sum of the probabilities of all structures containing that base pair i,j divided by the sum over all structures [83]. Knowledge based approaches rely on probabilistic models, where statistical learning procedures are used instead of empirical measurement of thermodynamic parameters. The Stochastic Context Free Grammar (SCFG) model [73] represents one popular example of such probabilistic models. The parameters used by the SCFG models are estimated on the set of RNAs with known structures (e.g., rRNA).

Prediction consistency is the main limit of both MFE and knowledge based methods [84]. (See the example in Figure 3). The percentage of known base pairs predicted correctly by the secondary structure prediction methods ranges from 40% to 75% [73–75,85]. This low figure may be the result of three confounding factors. Firstly, folding *in vivo* can be influenced by RNA chaperones [86], RNA editing [87], and even by the transcriptional process itself [88]. At present, there is no software able to account for these effects. Secondly, looking for a single structure can sometimes be inadequate. There are cases, such as the ribo-switches [89,90], where multiple functional structures can be derived from the same sequence depending on conditions such as temperature or other external factors. Standard predictors are not well suited to deal with such situations and require dedicated tools able to identify potential conformational switches [91,92]. Thirdly, RNAs might contain pseudo-knots, which are ignored by most tools due to reasons of computational complexity [66].

Figure 3. Consistency of RNA secondary structure predictions. In this example the human mir-3180 (Rfam accession id RF02010; AJ323057.1/363-249) was folded using different approaches yielding different output structures. (**a**) Secondary structure of the family as estimated by Rfam release 10.1; (**b**) RNAfold web server prediction based on Vienna RNA package version 2.0.0. [93]; (**c**) Mfold web server prediction, running Mfold version 4.6 [71].



The best secondary structure prediction accuracy can be achieved using comparative methods [66]. These apply to a set of structurally homologous RNA sequences being aligned. For some of these computation tools, the output will be the prediction of each individual homologous structure, while in other situations the result will be a unique consensus structure. The consensus structure does not exist *in vivo*, but rather it is a model that captures the common, relevant structural aspects conserved within the family.

2.2. Structure Prediction and Alignment Strategies

Due to the close relationship between sequence and structure, structure prediction and sequence alignment can be described as interdependent problems [63]. As theorized by Sankoff [94], the most suitable approach should involve the simultaneous alignment and folding of the considered sequences. Unfortunately, a strict application of this approach would be computationally prohibitive and the lack of an appropriate heuristic solution is reflected by the wealth of available alternative solutions. The web server WAR [95] is a good example. This tool allows the execution of a total of 14 different strategies to align and predict the common secondary structure of multiple RNA sequences. Over the years, so many methods have been described that some kind of classification is needed to catalogue them. Paul Gardner proposed three categories he refers to as “plans” [66,96]. In plan A, one starts with the estimation of a multiple sequence alignment and then the aligned sequences are folded jointly (as a kind of consensus). The initial alignment can be done by any standard MSA aligner (e.g., ClustalW [97], T-Coffee [98]), and the folding of the aligned sequences can be performed by a plethora of tools (e.g., RNAalifold [99], PFOLD [100], ILM [101], ConStruct [102]) optimizing a score based on compensated mutations and thermodynamics. However this strategy is effective just in a determined sequence similarity range. On one hand, sequences that are too similar do not carry any covariance or purifying selection information and are less informative from an evolutionary standpoint. On the other hand, sequences need to be similar enough to be meaningfully aligned as below the “twilight zone” of similarity sequence alignment tends to obscure the covariance signal [96]. Plan B makes it possible to

use evolutionary signals to help improve the reliability of structure predictions. This approach accounts for an underlying RNA substitution model where mutation probabilities are affected by structural dependencies. The maintenance of a 3D fold is a major evolutionary constraint influencing the acceptance of point mutations. From this perspective, a nucleotide located in the stem is not as free to mutate as a nucleotide located in a loop. Substitutions taking place in structured functional domains of RNAs can disrupt the wild-type conformation and seriously affect the molecular function. As a consequence, a nucleotide whose pairing has been disrupted by the mutation of its mate, is more likely to mutate itself so as to recover the original structure and rescue the function. Back in 1985 Sankoff developed a dynamic programming algorithm able to take into account sequence and structure of an RNA molecule simultaneously [94]. Unfortunately this algorithm is computationally expensive, with a running time equal to $O(N^{3m})$, where m is the number of sequences and N their length. This means that a pairwise comparison has the tremendous CPU cost of $O(N^6)$ which makes this algorithm inapplicable most of the times and calls for simplified heuristics. Several banded modifications of the Sankoff algorithm impose limits on the size or shape of substructures, like Dynalign [103,104], Foldalign [105,106], Stemloc [107], Consan [108]. Another example is pmmulti [109], a Sankoff algorithm variant able to perform multiple secondary structure alignments by aligning consensus base pair probability matrices. Plan C is used by programs such as R-Coffee [110] or RNAcast [111]. In these methods each individual sequence is folded separately before running the alignment. Equivalent secondary structures between two RNAs can be used to enhance the alignment accuracy. For instance, let seq1 and seq2 be two RNA sequences, x and y be two nucleotides matching in a secondary structure in seq1, and x' and y' be two nucleotides matching in a secondary structure in seq2. If x aligns to x' then implicitly y should be driven to align to y'. For example, R-Coffee uses RNAplfold [112] to compute the secondary structure of the provided sequences. After that, R-Coffee computes the multiple sequence alignment having the best agreement between sequences and structures. This pre-folding approach is especially valuable when RNAs are too different to be meaningfully aligned merely by using an off-the-shelf multiple alignment tools (*i.e.*, ClustalW [97]). Plan C is particularly well suited to situations where experimental secondary structures are available.

The situation is radically different when experimental 3D structure information is available. In this case the RNA alignment problem becomes similar to the protein structural alignment problem. This problem is nondeterministic polynomial-time complete (NP-complete) and it involves the alignment of two distance matrices. In most cases the problem can be solved in a rather satisfying way by using heuristics making the best of the geometric information contained in the PDB models. Examples of pairwise structural alignment methods for RNA are SARA [113], DIAL [114], iPARTS [115], ARTS [116] and SARSA [117]. Besides this, recently several 3D RNA structure database search programs were developed, such as LaJolla [118] and FRASS [119].

Giving an exhaustive overview of the methods available for folding and aligning structured RNA sequences is well beyond the scope of this review. Over the last twenty years, more than 30 methods have been described that deal with these issues which is an indication of the complexity of this problem, despite 25 years of research following its formal description by Sankoff.

3. Detecting ncRNA Homologues

In the ncRNA field another critical step is the collection of homologues to genes of interest. Homologues can be used in several situations, such as the detection of functional motifs, inference of possible evolutionary steps or the design of laboratory experiments. For instance, the conservation across species of a certain ncRNA can be used to estimate how likely a gene is to be functionally important. Such information can be used to prioritize experiments, e.g., knockdown experiments of the orthologous gene in a model organism. Over the last few years many different methods have been developed to approach the problem of RNA homology detection. As previously shown [120], homology search methods can be grouped in three sets: sequence-based, profiles and structure-based methods. The first and most straightforward approach to look for homologues is by comparing the nucleotide sequences. Already in 1981 Smith and Waterman developed a dynamic programming algorithm that allows for pairwise local alignment [121]. Nevertheless, this approach is CPU time demanding and implementations of this method have been until recently unpractical for large-scale database and genome wide screenings [122]. For this reason, alternative approaches such as FASTA [123] or BLAST [124] have been frequently preferred. These methods apply heuristics that boost computational speed at the cost of reduced accuracy. In both BLAST and FASTA, the underlying idea is to skip the time consuming comparison of entire query and target sequences, but rather to start identifying short conserved words in a first step called seeding. After that, more accurate time-consuming local alignments are performed. The second class of approaches are based on profiles, including HMMs. Profile HMMs are probabilistic models that are generated out of an input multiple sequence alignment where each position of the alignment is used to estimate nucleotide frequency. These models can be used to screen databases and look for homologs. Profile HMMs are heuristics having usually superior accuracy over methods based on single sequences [125,126]. However, such models have a linear architecture best suitable for modeling linear protein sequences (as opposed to secondary structure relationships). A more appropriate modeling of an RNA alignment should also consider RNA base pair interactions. The best sensitivity can be attained when applying approaches taking into account at the same time sequence similarity and secondary structures, as the Sankoff algorithm does. Unfortunately, the Sankoff algorithm is too computationally demanding, hence the need for approximate heuristic implementations of this exact algorithm. Such approximations include banded Sankoff tools [104,106,108,127], genetic algorithm implementations such as RAGA [128] and covariance models (CMs). CMs are the most commonly used method to carry out efficient database screening and can be described as special form of stochastic context free grammar (profile SCFGs). CMs were first introduced by Sean Eddy in [129] and implemented in Infernal [130]. This and other related applications such as RSEARCH [131] belong to a class of broadly used homology search tools based on the automatic learning of statistical models (the CMs) estimated from an input multiple RNA alignment decorated with the consensus secondary structure. CMs are probabilistic models that can be derived unambiguously out of a structure-annotated sequence alignment and can be used in turn to query a target sequence database to find homologs. Conceptually CMs are similar to profile HMM but able to include RNA base-pairs interactions information. The modeling of such information results in a higher complexity and CMs are represented by a tree-like model architecture, where the tree shape directly mirrors the consensus RNA structure. Unlike HMM states that only allow the emission of

matches and indels, CMs embed new states to handle paired/not-paired and directionality information. For instance, in the paired sites, deletions can involve either a single 5' or 3' nucleotide, or the complete base pair. The direction also matters for the insertions that can now concern either the 5' or 3' ends of a base pair. In order to permit multi-loops, the bifurcation states are incorporated as well. In spite of their superior accuracy, CM cannot be used in all situations and are restricted to the identification of unsplit genes. The mapping of queries composed by multiple exons is impossible due to the impossibility of aligning secondary structures to a target interrupted by introns whose position is unknown. Moreover CMs need to "learn" from a set of homologous transcripts, but the set of sequences required to train the model are not always available. There is some circularity in this problem where the CM is used to search homologs that are themselves needed to estimate the model. Another layer of complexity results from the need to assemble a multiple sequence alignment of homologous sequences, needed to train the CM. In the CM the alignment will be used for a probabilistic description of matches, mismatches, insertions and deletions. However, generating accurate RNA alignments is difficult. In Rfam [132] CMs parameters are trained on high quality alignments (seed alignment) obtained from literature with manual curation. This input is used to estimate CMs, which are then passed to Infernal for homology searches. This CM/Infernal strategy is analogous to HMM/HMMER used for Pfam [133]. An option for spotting promising sequence segments and accelerate the detection procedure is to include a pre-filtering step as done for the Rfam setup [134]. This can be accomplished by means of *ad hoc* algorithms [135], profile HMMs like ML-heuristic [126] or BLAST with relaxed expectation values (*E*-values) to avoid losing sensitivity as achieved in Rfam [136]. A number of studies have been dedicated to the optimization of BLASTn parameters for seeking RNA homologs. For instance, in one study [120,137] the effectiveness of BLAST and other popular homology search methods tuned for ncRNA screenings were benchmarked. In [138], BlastR is introduced, a method that both takes advantage of di-nucleotide conservation and BLASTp as search engine to discover distantly related homologs. BlastR can be mounted on the top of computationally demanding algorithms to serve as a pre-filtering tool. One merit of this approach is that it neither require profiles nor secondary structure information, but relies solely on information encoded in primary nucleotide sequences.

Together with sequence-based, profiles and structure-based methods, another possibility for detecting inter-species homologs involves the use of multiple genome alignments [43]. Once established reciprocity between blocks of genomes belonging to different organisms (*i.e.*, syntenic regions), coordinate transfer from one gene to its homolog is straightforward and implies the projection of corresponding positions. This has been made possible thanks to the availability of genome sequences [139–142] and the development of alignment tools able to detect orthologous genomic regions, *i.e.*, loci that proceeded from the same genomic position in the ancestral genome [143]. Although comparing ncRNAs is currently still a complicated task, there exist several bioinformatics options to workaround the poor sequence conservation and effectively perform homology based prediction of novel ncRNAs.

4. High-Throughput Technologies and Genome-Wide Annotation of ncRNAs

4.1. Approaches for the High-Throughput Expression Detection

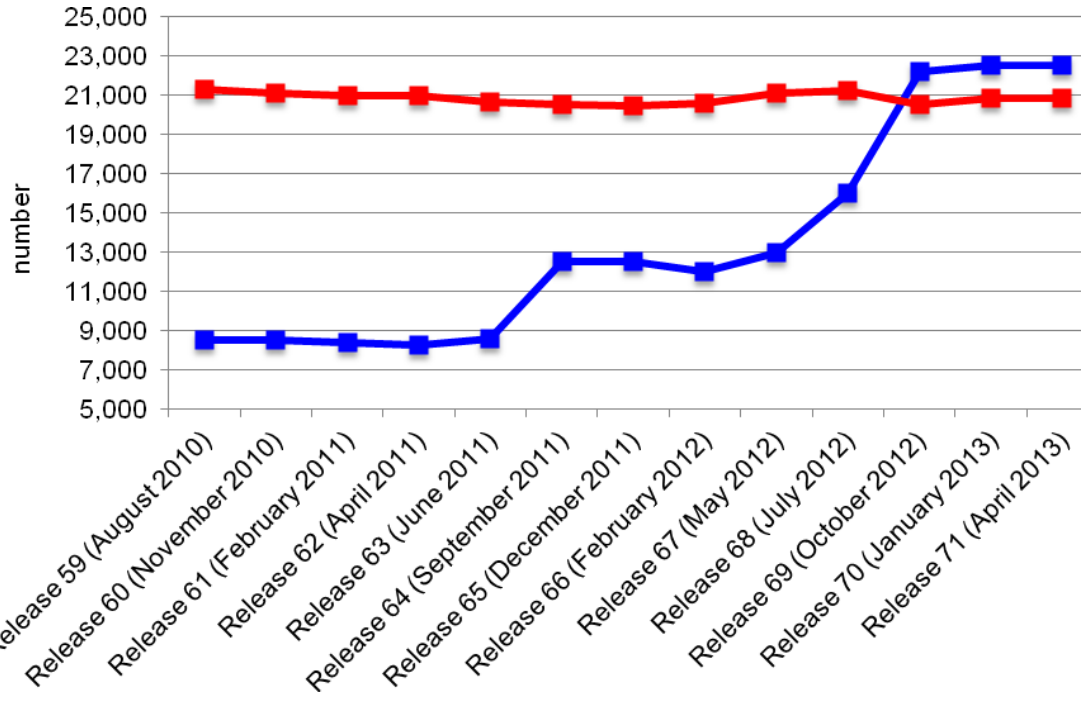
Recent technological advances have allowed the collection of an unprecedented amount of RNA sequence data coming from a wide range of organisms and conditions. For many years the main strategy for transcript discovery had been the sequencing of cloned complementary DNA (cDNA) of expressed sequence tags (ESTs) [144–146]. EST sequencing was then successfully used for the generation of large-scale expression datasets [147], and already by 1991 this approach had been utilized for human gene discovery [148]. Although it is widely acknowledged that ESTs represent a valuable resource to detect gene expression, they also came with severe limitations such as cost and sequencing requirements. Their dependence on bacterial cloning is an important source of bias and contamination that can lead to redundancy and under-representation or over-representation of host-selected transcripts [149–151]. More recently, oligonucleotide microarray technologies have made high throughput expression analysis much more practical, while the even more recently developed RNA-seq technologies promising transcriptomic analysis of unprecedented accuracy thanks to the application of NGS methods to transcriptome sequencing. Microarrays rely on a collection of nucleotide probe spots attached to a solid support. RNAs are labeled with fluorescent dyes, hybridized to the arrays, washed, and scanned with a laser [152]. Such arrays have been used for the investigation of known or predicted genes and have been until recently one of the most widespread technology for transcriptome exploration. Standard expression arrays are affected by several limitations including the hybridization and cross-hybridization artefacts [153–155], dye-based identification problems [156–160] and physical manufacturing restrictions, impeding the detection of splicing events and the discovery of unannotated genes [151]. A variant of traditional expression array is represented by tiling arrays. These are chips that use extremely densely spotted and probes representing overlapping contiguous regions of genome. Several works relying on this technology and aiming at transcript discovery have been published [38,161–164]. However tiling arrays require a substantial quantity of RNA and have further limitations affecting their sensitivity, specificity and the detection of splicing [151]. For instance, as shown in [165], microarrays lack sensitivity for genes expressed either at low or very high levels and if compared with RNA-seq have much smaller dynamic range. As a consequence, microarrays are inadequate for the quantification of both the prevailing RNA classes, and the less abundant ones. For genes with medium levels of expression, RNA-seq and microarrays return comparable results [165–167]. Still, each approach presents very specific advantages and disadvantages. A thorough comparison of these two approaches lies outside the purpose of this text (for reference, see [152,166,168]). Additional methods for high-throughput RNA discovery include the serial analysis of gene expression (SAGE) [169,170], several updated variants such as LongSAGE [171], RL-SAGE [172], SuperSAGE [173] and analogous approaches like the massively parallel signature sequencing (MPSS) [174]. In general, SAGE-like methods consist in the cloning and then the sequencing of short tags (17–25 nucleotides) coming from RNA extract. The resulting tag sequences can be compared against the source genome or a reference RNA database to attain the digital count of transcript quantities. Two other protocols that can be used in combination with high-throughput sequencing are the paired-end ditags (PETs) [175] and the rapid amplification of cDNA ends (RACE) [176–178]. Both approaches can be used to demarcate transcript

boundaries, *i.e.*, define start and end of a transcript. Such information is extremely valuable *in situ* ations where the first and last exons can be respectively 5' and 3' associated with other transcript isoforms, thus making it difficult to define gene boundaries. Similarly, the cap analysis of gene expression (CAGE) [179,180] is a technique that allows high-throughout profiling of transcriptional starts points. Another promising application for ncRNA discovery, named RNA CaptureSeq, has been recently reported [181]. This approach is able to reach unprecedented sequencing depth. RNA CaptureSeq is inspired from exome sequencing techniques and relies on the use of tiling arrays in order to enrich the population of RNAs one wants to sequence. This enrichment step allows a sequencing depth that would be impossible when dealing with the full transcriptome. Although RNA CaptureSeq is not suited to generate full transcriptome profile, it can be used to target specific genomic sites and detect transcript isoforms expressed at very low abundance. As shown in [181] RNA CaptureSeq can be used to fuel the detection of ncRNAs that are missed by genome-wide standard RNA sequencing.

4.2. Datasets

Undoubtedly, high-throughput technologies enable the tremendous possibility to get both qualitative and quantitative information on whole transcripts mass produced by cells. This has resulted in high-resolution views of RNA expression dynamics throughout different tissues and time points [182–184] and fueled the development of ncRNA specific databases, such as Rfam [132], NRED [20], lncRNAdb [33], RNAdb [185], fRNAdb [186] and NONCODE [187]. Furthermore, various groups and projects, such as RefSeq [188], GENCODE [14,189], HAVANA team [190,191], Ensembl [192] and FANTOM [193] undertook the task to comprehensively annotate functional elements, including ncRNAs, of a number of species using experimental data. The RefSeq repository houses annotations resulting from automated analyses, collaboration and manual curation [188,194]. The GENCODE pipeline combines HAVANA and Ensembl automatic annotations to annotate the human gene features generated in the context of the ENCODE project [14,45,189]. The HAVANA team has the goal to provide manually curated annotations of transcripts aligned to human, mouse and zebrafish genomes. Ensembl runs an automatic *genebuild* process including *ab initio* gene predictions and release 64 supported a total of 61 species [192]. The Ensembl *genebuild* system is adapted to every species in the set according to the data that is available. For instance Ensembl imports and merges high quality HAVANA annotations exclusively for human and mouse. The FANTOM consortium aims to provide functional annotations to the full-length cDNAs [193]. The annotations generated by these consortia are freely available through genome browsers, including UCSC [195], Ensembl [196] and VEGA [197]. As new genomic regions get annotated and new transcript sequences become publicly available, these gene sets continue to growth [14,188,194]. A recent publication [14] indicated that in the last years the number of annotated protein-coding and non-coding transcripts in GENCODE has dramatically increased. For instance, passing from GENCODE version 3c (July 2009, http://www.gencodegenes.org/archive_stats.html) to version 7 (December 2010, http://www.gencodegenes.org/archive_stats.html), the number of protein-coding transcripts increased from 68,880 to 76,052, and the number of lncRNAs jumped from 10,457 to 15,512. In terms of gene annotations, the number of known protein-coding genes has remained almost unchanged, while the ncRNA gene annotations expanded tremendously (see Figure 4).

Figure 4. Number of non-coding and protein-coding genes annotated over the last Ensembl releases. The *x*-axis indicates the number and the date of the release. The vertical axis reports the number of ncRNA (blue line) and protein-coding genes (red line).



The overall picture, however, remains blurred by inconsistent findings, suggesting that more analyses are still needed. For instance, the recent estimates reported by the ENCODE project indicate that about the 62% of human genomic bases are expressed in long transcripts, while 5.5% only of the whole genome is found within the GENCODE annotated exons [198]. This discrepancy can be in part explained by the fact that GENCODE catalogues transcripts using cDNA/EST alignments [14] rather than RNA-seq short-read data. A classic low-throughput EST sequencing operated by the Sanger technology can identify mostly high abundant transcripts [199], while deep coverage RNA-seq experiments can reveal rare but potentially regulatory transcripts. Nonetheless, ESTs are longer than RNA-seq reads, and can provide more reliable transcriptional evidence [200].

4.3. NGS Challenges

To make the most out of the extraordinary possibilities that NGS offers, it is essential to understand the current limitations. One important point is that the reads returned by standard NGS platforms are usually short (35–500 base pairs [201]) and as a consequence it becomes necessary to reassemble the full-length transcripts. Small non-coding RNAs (*i.e.*, miRNA and piRNAs) represent an exception and there is no need to reassemble them, as they are small enough to be entirely covered by the read length. Unfortunately the process of reassembling transcriptomes starting from short reads is difficult. Normally RNA-seq dataset are big (gigabases to terabases), and thus need to be handled by sufficient large memory and by multi-CPU computers able to execute the algorithms in parallel with sufficient high-performance storage to store primary, temporary and output data. Although various short-read assemblers [202–204] were successfully applied to genome assembly, these packages cannot be easily

used to reconstruct transcriptomes. Applying tools normally designed for genome reconstruction to the problem of transcriptome assembly leads to multiple complications. A key issue is that the DNA sequencing depth is supposed to be identical over the entire genome while transcriptome sequencing depth is expected to fluctuate significantly. For this reason, DNA short-read assemblers could erroneously interpret highly abundant transcripts as repetitive genomic regions. Furthermore, when using genome short-read assemblers the read strand is not taken into account. On the contrary, when available, a transcriptome assembler should exploit the strand information to unravel possible antisense expressions on different strands. Finally, the transcript modeling is involved as transcript variants coming from the same gene can share exons and are difficult to resolve unambiguously [199].

It is possible to work out the transcriptome assembly following a reference-based approach, a *de novo* assembly or combinations of each [199]. The first considers the initial mapping of the reads on a reference genome, and then the usage of transcript assemblers. To the end of labeling each read with the genomic location they come from, a new class of software, generally referred as read mapper, has recently shown up. In this context, the availability and the quality of the underlying reference genome are critical. Besides that, when dealing with massive amount of short-read data the CPU and the memory costs can be challenging, and several algorithms are being tailored to achieve best mapping efficiency [205–211]. Other important issues relate to the mapping of reads crossing exon-junction boundaries [212,213] and the uncertainty or lack of accuracy in read alignments. For most downstream applications, the accurate positioning of the reads back to the source genome is crucial. To improve the mapping accuracy, the process can take into account the read quality information [214,215]. The quality scores, introduced by the Phred algorithm [216,217], indicate the reliability of each base call in each read in a log-likelihood scale. Since bases with reduced quality scores have an increased probability to be sequencing errors, a read mapper should either use less severe penalization for mismatches at positions with low base-call quality, or not align such positions at all. The information about the quality score is particularly relevant when mapping reads of large size. This is a result of the fact that 3' ends of longer reads are affected by sequencing errors at higher rates [215]. Besides choosing a threshold for accepted mismatches, other important and sometimes arbitrary decisions regard the split mapping and multiple mapping reads. The first refers to reads that could not be aligned to the reference genome unless split in subparts. Such reads could either highlight the presence of an unreported exon-junction boundary, or be sequencing artefacts. The second indicates reads that align multiple times across the reference genome. This mapping uncertainty is caused by repeated elements and may results in flawed expression establishments. On one hand, removing multiple mapping reads from the analysis would imply an underestimation of the expression of genes embedding repeats. On the other hand, considering multiple mapping reads would lead to artefactual expression measurements. Once mapped the reads, additional issues concern the application of transcript assemblers. Several computational tools have been developed with the purpose of reconstructing transcripts in their entire length, *i.e.*, annotating exon-intron transcript structures. These methods include Cufflinks [218], Isolazo [219] and Scripture [42]. In [220] the authors have shown that variations across transcript assemblers can be source of confusion, with low consistency across methods and a high number of false positives [200,219]. Transcript assemblers seem to have a better agreement when reconstructing protein-coding transcripts [43] with the agreement dropping dramatically when modeling large intervening ncRNAs (lincRNAs). For instance, Cufflinks and

Scripture share only 46% agreement for lincRNA transcript models [43]. Such discrepancies are caused by the differences in how each assembler reconstruct lowly expressed transcripts [43]. In other words, about half of the isoforms estimated by a method in areas with low read density do not correspond to isoforms called by the other method. This poor agreement between transcript assemblers highlights the need for further improvements, calling for the development of new algorithms to accurately represent low abundant transcripts.

Another possibility to assemble a transcriptome from-short reads is *de novo* assembly of transcripts. This strategy does not require any reference genome and is therefore independent on the correct alignment of the reads to the splice sites. Advantages of this approach are that it is less reliant on accurate genome annotation and can be applied to organisms without sequenced or fully annotated genomes. Examples of applications adopting this strategy are described in [221–223]. Nevertheless, the application of *de novo* assembly to complex transcriptomes (e.g., higher eukaryotes) is complicated by the dataset sizes and the dense network of alternatively spliced variants. Furthermore, *de novo* transcriptome assemblers need much deeper sequencing than reference-based assemblers and are largely affected by sequencing errors [199].

Once transcriptome dataset is generated, there are additional complications in the downstream analysis if trying to distinguish genuine ncRNAs from mRNAs. Currently, this issue is becoming increasingly important as many researchers are only interested in one or the other. The most straightforward procedure would be to compare a newly generated transcriptome against existing gene annotations. However in most cases annotations are far from complete and the great majority of genes they include are protein-coding. As a consequence, in a normal RNA-seq experiment a substantial fraction of read contigs map outside of annotated exons [198]. Previously unreported transcripts can be either classified as ncRNA or mRNA according to the protein-coding potential they have. The *in-silico* assignment of a transcript to one of these two groups is not always trivial and may require dedicated expert curation [190]. Some transcript isoforms might insert coding exons and therefore could be partially translated, *i.e.*, generating small peptides. There are further ambiguities for coding transcripts whose untranslated structured molecules are also functional as ncRNAs [224] and for genes having both coding and non-coding isoforms [225]. A commonly used approach to predict the coding potential involves the codon substitution frequency (CSF) estimation [226,227]. This measure is based on an input multiple alignment of orthologous sequences. The CSF score deems a region to be coding depending on how the sequences of the multiple alignment evolved, *i.e.*, showing distinctive mutation patterns, as are expected in coding and non-coding loci. A coding region is expected to embed prevalently conservative amino acid substitutions and synonymous codon substitutions, while showing low occurrence of nonsense and missense mutations. Although CSF has been successfully applied in various research projects [226,228,229], the score is not always easy to estimate with the availability of trustworthy orthologues being the main limiting factor when dealing with new transcriptome datasets. Issues include scarcity or even the absence of orthologs, erroneous insert of pseudogenes in the set and absence of informative variations. For instance, as shown in [40] many putative human lncRNAs are not found in other species, and cannot be analyzed using CSF. Besides this, primate specific lncRNAs rarely show sufficient changes to highlight sense/nonsense mutations patterns. In addition to CSF, other strategies not relying on evolutionary signatures can be effectively used to predict if a transcript is going to be translated into protein or not. For example, there are dedicated

BLAST flavors including BLASTx and RPS-Blast [124,230] that can be used to identify transcripts whose translational product possesses a match in protein databases such as Pfam [133] and UniProt [231]. Other algorithms include CPC (Coding Potential Calculator), a support vector machine (SVM) classifier including both Open Reading Frame (ORF) and homology predictions features [232], PORTRAIT (Prediction of transcriptomic ncRNA by ab initio methods) a SVM classifier not using homology information [233] and CPAT (Coding Potential Assessment Tool), a logistic regression model built with four sequence features including ORF predictions [234]. Unfortunately, bioinformatics predictions can easily return mistaken assignments when dealing with ncRNAs closely related to coding mRNAs, and result in some confusion when transferring annotation across species, or within a genome. Such observations may wrongly suggest pseudogenization events or a turnover between proteins and ncRNAs.

4.4. Other Approaches

Over the last few years, other approaches alternative or complementary to RNA-seq have been attempted to generate high-throughput ncRNA annotations. In 2009, Mitchell Guttman and co-workers published the first of a series of analysis that recently came out linking lncRNA detection to histone modifications [13]. In this work, the authors pioneered a chromatin-state based method to identify well-defined transcriptional units occurring between known protein-coding genes. Their analysis relied on the observation by [235] that promoters of genes expressed by the RNA polymerase II (Pol II) are signed by trimethylation of lysine 4 of histone H3 (H3K4me3) while the transcribed area is marked by trimethylation of lysine 36 of histone H3 (H3K36me3). Following this observation, the authors did chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) [235] to generate profiles of chromatin states. This approach revealed 1600 mouse lincRNAs, corresponding to H3K4me3-H3K36me3 chromatin domains and lying outside of protein-coding regions. The prediction reliability has been estimated by additional analysis showing that lincRNAs are more conserved than neutrally evolving sequences and that most of experimentally tested loci were found to be expressed [13]. An alternative strategy used for ncRNA detection involves a combination of different high-throughput data sources and their integration using bioinformatics [236]. This approach, named incRNA, relies on a machine learning method and has been applied to the genome-wide identification of *Caenorhabditis elegans* ncRNAs. incRNA combines predicted and experimental data for a total of nine different information sources. These include the expression data coming from various developmental stages and conditions, as well as the GC content, the predictions of RNA secondary structure folding energy, the prediction of evolutionary conserved DNA sequence and secondary structure. These results illustrate how the integration of multiple information sources ends in highly accurate predictions of novel ncRNA genes.

Recently, a number of works reporting a massive quantity of novel ncRNA genes in various species have been published [40,43,44,237,238]. Such rapid growth has been possible thanks to the parallel development of new and ever more sophisticated bioinformatics approaches. Nevertheless, such analyses remain superficial with uncertainties of different type and degree affecting most predictions. For example, the homology search pipeline described in [40] is not sensitive enough to map rapidly evolving lncRNAs, hence the limit to play comprehensive evolutionary study. Moreover such lncRNA predictions should be taken with care, not just because they are not experimentally verified, but also

because they are far from representing the complete genome-wide lncRNA figure. For validation purpose, some works provide the number of predicted lncRNA supported by expression evidences. For instance in [13] the authors confirmed by tiling array the expression of ~70% lncRNA predictions. In other cases as in [44], RT-PCR has been used to validate 15 newly identified lncRNAs. On the short run available transcription data is expected to increase very rapidly, and the necessity to accurately and quickly validate ncRNAs is becoming more pressing than ever.

Table 1. A summary of methods, datasets and browsers for non-coding RNA analysis. The first column indicates the resource type. The second column the resource name. The third column reports the PubMed ID when available, if not the web address. The fourth column provides a brief description of the resource.

	Resource	Pubmed ID	Description
Comparing ncRNAs (Section 2)	Mfold	6163133	Single sequence RNA secondary structure prediction.
	RNAfold	12824340	
	WAR	18492721	WEB server allowing the execution of different alignment methods
	RNAalifold	12079347	Folding previously aligned RNAs (Plan A)
	PFOLD	12824339	
	ILM	14693809	
	Construct	10518612	
	Dynalign	11902836	Sankoff derived algorithm for the simultaneous alignment and secondary structure prediction (Plan B)
	Foldalign	9278497	
	Stemloc	15790387	
	Consan	16952317	
	pmmulti	15073017	
	R-Coffee	18420654	Aligners taking into account previously estimated secondary structure (Plan C)
	RNAcast	16020472	
	SARA	18689811	3D structure alignment method
	DIAL	17567620	
	iPARTS	20507908	
	ARTS	16204124	
	SARSA	18502774	
	LaJolla	http://www.mdpi.com/1999-4893/2/2/692	
	FRASS	20553602	

Table 1. *Cont.*

	Resource	Pubmed ID	Description
Detecting ncRNAs (section 3)	ML-heuristic	16267089	Profile HMM
	RAGA	9358168	Genetic algorithm
	RSEARCH	14499004	Covariance model
	Infernal	12095421	
	BlastR	21624887	BLAST-based dinucleotide homology search
Datasets and browsers (section 4)	ENCODE	22955616	Consortium
	Ensembl	22086963	
	FANTOM	11217851	
	HAVANA	http://www.sanger.ac.uk/research/projects/vertebrategenome/havana/	Annotation team
	GENCODE	22955987	Project for the annotation of all human gene features
	UCSC	12045153	Genome browser
	VEGA	18003653	
	RefSeq	18927115	Collection of DNA, transcripts, and proteins
	Rfam	12520045	ncRNA database
	NRED	18829717	
	lncRNAdb	21112873	
	RNAdb	17145715	
	fRNAdb	17099231	
	NONCODE	15608158	

5. Discussion and Conclusions

ncRNA functional characterization is a rapidly expanding research area. In the past few years, it has become clear that the majority of the transcripts in cells are more than mere intermediates between the hereditary information encoded in DNA and the mechanical operative component represented by proteins. Indeed, it appears that numerous transcripts may not be translated at all while still being involved in critical biological functions such as cell differentiation and chromatin remodeling. Taking together 15 human cell lines, the cumulative coverage of transcribed regions is ~62% and ~75% of the whole human genome for processed and primary transcripts, respectively [239]. This "pervasive transcription" is strikingly high, especially when considering that a mere 3% of the human genome codes for protein-coding exons. [198]. Numerous novel, previously uncharacterized RNA species have been recently detected. A sizeable fraction of them are defined as lncRNA, *i.e.*, functional molecules longer than 200 nucleotides that do not show any coding potential. Some of these molecules are spliced, capped, differentially expressed in tissues/cells or developmental stages and tend to be more conserved across species than would result from neutral evolution. For these reasons and because of the increasing number of transcripts whose function has been experimentally validated, it is believed that many of these new ncRNAs belong to an important, relatively unexplored class of regulatory elements. Thanks to ongoing improvements in sequencing technologies it has become possible to

collect a significant amount of these uncharacterized transcripts. The latest generation of sequencing technologies makes it possible to perform large scale sequencing of entire transcriptomes. This technique, known as RNA-seq has already had a dramatic impact on our perception of the human transcription landscape [183,239]. Similar studies have been carried out in a number of genetic model organisms including rodents [44,151], plants [240], insects [184], worms [241] and yeasts [242]. In [243] the author argues that RNA-seq represents the most promising technology for transcriptome research. The main strength of RNA-seq approaches are the high degree of dynamic range they offers, returning better sensitivities than microarrays without the need of *a priori* speculation regarding the genomic loci being transcribed [244]. If the pace of scientific progress is maintained and if costs keep decreasing, one can reasonably expect this technology to rapidly become a key component of personalized medicine, especially when considering the new venues of development that are currently being considered [152,245].

From a functional perspective, much remains to be done for accurate characterization and functional analysis of ncRNAs. To infer the function of novel ncRNAs one possibility is looking for functional motifs. This can be done by running motif finders algorithms to predict structurally conserved and potentially functional motifs [246–252]. Furthermore, the functional characterization of a novel ncRNA can be aided by the detection of protein-RNA binding motifs and the identification of protein interaction partners. Experimental approaches suited for this include RIP (Rna Immunoprecipitation) and CLIP (Cross-Linking and ImmunoPrecipitation) [253]. Comparative studies also offer a very efficient way to have functional insights and prioritizing analysis. They can be used to predict function by homology, assess phylogenetic relationships, detect functional motifs or classify related molecules in order to identify families. A major challenge when tackling ncRNA comparisons results from the remarkable variability of traits and functions. Considering sizes only, ncRNA molecules can be as short as a miRNA (~22 nt) and up to ~17 kb long in the case of Xist [2]. Another difficulty when comparing ncRNAs is that most of these genes have poorly conserved sequences. Such diversity challenges our ability to compare, classify and search with conventional alignment tools. In addition ncRNA genes have no equivalent of codon bias and ORFs that help powering the statistical component of machine learning approaches when doing protein prediction [254]. The strongest signal contained by RNA sequences is usually evolutionarily conserved secondary structures. Many efficient algorithms exist that are able to predict potential structures using MFE or SCFG computations. Unfortunately, these predictions ignore the contribution of the environment and are not always accurate enough to significantly improve alignment accuracy and homology modeling. Emerging technologies allowing the high-throughput generation of experimentally derived secondary structures [255] will hopefully help addressing this problem. Unfortunately, taking into account secondary structures while comparing sequences is a challenging procedure, too intensive from a computational point of view to be practical in most circumstances [108]. This makes it is difficult to compare mono-exonic genes while taking the secondary structure into account, and totally impossible when the transcripts are multi-exonic (*i.e.,* the secondary structures are interrupted by introns). It has been shown [40,237] that BLAST can be effectively used for lncRNA homolog prediction, in combination with splicing informed heuristics such as exonerate [256] or GeneWise [257]. This strategy is not new, and similar approaches have already been used for the discovery of protein-coding homologs [258–260]. As one would expect, homology based RNA searches are severely limited by the capacity to

align distant homologues. For instance, when searching the human lncRNA complement against mammalian genomes [40] or when using an estimated pig complement [237], the authors only managed to find, beyond primates, less than 50% of the query genes across cow, mouse or dog. This result may reflect a high turnover, but the conservation/disappearance patterns, poorly correlated to phylogenetic history, are most likely indicative of a limited detection capacity. Other confounding factors include misassembled or partially sequenced genomes. Additional analysis would be needed to validate the Blast/exonerate mapping approach. At this stage, it is therefore impossible, without further experiments, to establish whether lncRNA queries that failed to map are really absent in the target species or undetected. In this context, high quality templates, such as the GENCODE queries used in [40], offer better likelihood to return precise annotation. In the same publication it is also predicted that sizeable fraction of the human lncRNAs is primate specific [40]. This result is in agreement with a recently published study [44] where the authors identified lncRNAs expressed in rodents' adult liver, and then compared the expression of the orthologous genomic regions. This work illustrates that loss of lncRNA transcription among rodents is associated with loss of sequence constraints and that many lncRNA genes seems to be species or lineage specific. Another application of homology based approaches is the possibility to identify novel human lncRNA genes candidates by using non-human templates as query [237]. As shown in the paper [237], there are 131 pig lncRNAs mapping to unannotated regions of the human genome. This result suggests that although human is probably one of the most extensively annotated higher-eukaryote, extra improvements might be achieved using data gathered in other non-model organisms. In [40] the authors also extend the lncRNA conservation study to a multiple genome alignments strategy based on PhastCons conservation scores. The analysis is in agreement with previous reports [13,30] and confirms that lncRNAs sequences are less constrained than those of protein-coding genes. Remarkably, it was shown that the distribution of lncRNA exons conservation is bimodal, with a fraction substantially approximate to ancestral repeats, and another group appreciably shifted toward the protein-coding set. This indicates that some lncRNA are under a selection as strong as that seen for proteins and suggests that a sizeable fraction of lncRNA genes are probably functional. The fraction of lncRNAs having a mutation rate almost indistinguishable from repeats suggests that at least some lncRNAs (close to a third) may be transcriptional noise. However, despite this abundance of lncRNA sequences that do not appear to be under selection, the transcript product itself might still have a biological role and as shown in [261,262] the transcription process itself of some ncRNA can bear regulative functions.

Despite the difficulties encountered when comparing ncRNAs, homology search of ncRNAs can be successfully used to detect new genes. New and ever more sophisticated algorithms will help addressing the challenges brought by NGS technologies. The ultimate goal is the creation of thorough transcriptome annotations and unbiased expression profiling of each individual transcript. It is still too early to tell. However, if they live up to their promises and expectation, the discovery of this new large class of RNAs may well define one of the turning points of modern biology.

Acknowledgments

The authors wish to acknowledge Roderic Guigó for the helpful comments and suggestions.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Lee, R.C.; Feinbaum, R.L.; Ambros, V. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **1993**, *75*, 843–854.
2. Brown, C.J.; Hendrich, B.D.; Rupert, J.L.; Lafreniere, R.G.; Xing, Y.; Lawrence, J.; Willard, H.F. The human XIST gene: Analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **1992**, *71*, 527–542.
3. Farazi, T.A.; Juranek, S.A.; Tuschl, T. The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development* **2008**, *135*, 1201–1214.
4. Bachellerie, J.P.; Cavaille, J.; Huttenhofer, A. The expanding snoRNA world. *Biochimie* **2002**, *84*, 775–790.
5. Barrett, T.; Suzek, T.O.; Troup, D.B.; Wilhite, S.E.; Ngau, W.C.; Ledoux, P.; Rudnev, D.; Lash, A.E.; Fujibuchi, W.; Edgar, R. NCBI GEO: Mining millions of expression profiles—database and tools. *Nucleic Acids Res.* **2005**, *33*, D562–D566.
6. Parkinson, H.; Sarkans, U.; Shojatalab, M.; Abeygunawardena, N.; Contrino, S.; Coulson, R.; Farne, A.; Lara, G.G.; Holloway, E.; *et al.* ArrayExpress—A public repository for microarray gene expression data at the EBI. *Nucleic Acids Res.* **2005**, *33*, D553–D555.
7. Griffiths-Jones, S.; Grocock, R.J.; van Dongen, S.; Bateman, A.; Enright, A.J. miRBase: MicroRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* **2006**, *34*, D140–D144.
8. Fraser, B.A.; Weadick, C.J.; Janowitz, I.; Rodd, F.H.; Hughes, K.A. Sequencing and characterization of the guppy (Poecilia reticulata) transcriptome. *BMC Genomics* **2011**, *12*, 202.
9. Tuda, J.; Mongan, A.E.; Tolba, M.E.; Imada, M.; Yamagishi, J.; Xuan, X.; Wakaguri, H.; Sugano, S.; Sugimoto, C.; Suzuki, Y. Full-parasites: database of full-length cDNAs of apicomplexa parasites, 2010 update. *Nucleic Acids Res.* **2011**, *39*, D625–D631.
10. Mamidala, P.; Wijeratne, A.J.; Wijeratne, S.; Kornacker, K.; Sudhamalla, B.; Rivera-Vega, L.J.; Hoelmer, A.; Meulia, T.; Jones, S.C.; Mittapalli, O. RNA-Seq and molecular docking reveal multi-level pesticide resistance in the bed bug. *BMC Genomics* **2012**, *13*, 6.
11. Dinger, M.E.; Amaral, P.P.; Mercer, T.R.; Pang, K.C.; Bruce, S.J.; Gardiner, B.B.; Askarian-Amiri, M.E.; Ru, K.; Solda, G.; Simons, C.; *et al.* Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res.* **2008**, *18*, 1433–1445.
12. Okazaki, Y.; Furuno, M.; Kasukawa, T.; Adachi, J.; Bono, H.; Kondo, S.; Nikaido, I.; Osato, N.; Saito, R.; Suzuki, H.; *et al.* Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **2002**, *420*, 563–573.
13. Guttman, M.; Amit, I.; Garber, M.; French, C.; Lin, M.F.; Feldser, D.; Huarte, M.; Zuk, O.; Carey, B.W.; Cassady, J.P.; *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **2009**, *458*, 223–227.

14. Harrow, J.; Frankish, A.; Gonzalez, J.M.; Tapanari, E.; Diekhans, M.; Kokocinski, F.; Aken, B.L.; Barrell, D.; Zadissa, A.; Searle, S.; *et al.* GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res.* **2012**, *22*, 1760–1774.
15. Managadze, D.; Lobkovsky, A.E.; Wolf, Y.I.; Shabalina, S.A.; Rogozin, I.B.; Koonin, E.V. The vast, conserved mammalian lincRNome. *PLoS Comput. Biol.* **2013**, *9*, e1002917.
16. Mattick, J.S.; Gagen, M.J. The evolution of controlled multitasked gene networks: The role of introns and other noncoding RNAs in the development of complex organisms. *Mol. Biol. Evol.* **2001**, *18*, 1611–1630.
17. Mattick, J.S. Non-coding RNAs: The architects of eukaryotic complexity. *EMBO Rep.* **2001**, *2*, 986–991.
18. Crick, F.H. On protein synthesis. *Symp. Soc. Exp. Biol.* **1958**, *12*, 138–163.
19. Wang, J.; Zhang, J.; Zheng, H.; Li, J.; Liu, D.; Li, H.; Samudrala, R.; Yu, J.; Wong, G.K. Mouse transcriptome: Neutral evolution of "non-coding" complementary DNAs. *Nature* **2004**, *431*, doi:10.1038/nature03016.
20. Dinger, M.E.; Pang, K.C.; Mercer, T.R.; Crowe, M.L.; Grimmond, S.M.; Mattick, J.S. NRED: A database of long noncoding RNA expression. *Nucleic Acids Res.* **2009**, *37*, D122–D126.
21. Mattick, J.S. The genetic signatures of noncoding RNAs. *PLoS Genet.* **2009**, *5*, e1000459.
22. Wapinski, O.; Chang, H.Y. Long noncoding RNAs and human disease. *Trends Cell. Biol.* **2011**, *21*, 354–361.
23. Wang, X.; Song, X.; Glass, C.K.; Rosenfeld, M.G. The long arm of long noncoding RNAs: roles as sensors regulating gene transcriptional programs. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, a003756.
24. Satterlee, J.S.; Barbee, S.; Jin, P.; Krichevsky, A.; Salama, S.; Schratt, G.; Wu, D.Y. Noncoding RNAs in the brain. *J. Neurosci*. **2007**, *27*, 11856–11859.
25. Mercer, T.R.; Dinger, M.E.; Sunkin, S.M.; Mehler, M.F.; Mattick, J.S. Specific expression of long noncoding RNAs in the mouse brain. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 716–721.
26. Kaikkonen, M.U.; Lam, M.T.; Glass, C.K. Non-coding RNAs as regulators of gene expression and epigenetics. *Cardiovasc. Res.* **2011**, *90*, 430–440.
27. Braidotti, G.; Baubec, T.; Pauler, F.; Seidl, C.; Smrzka, O.; Stricker, S.; Yotova, I.; Barlow, D.P. The Air noncoding RNA: An imprinted *cis*-silencing transcript. *Cold Spring Harb. Symp. Quant. Biol.* **2004**, *69*, 55–66.
28. Willingham, A.T.; Orth, A.P.; Batalov, S.; Peters, E.C.; Wen, B.G.; Aza-Blanc, P.; Hogenesch, J.B.; Schultz, P.G. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **2005**, *309*, 1570–1573.
29. Cesana, M.; Cacchiarelli, D.; Legnini, I.; Santini, T.; Sthandier, O.; Chinappi, M.; Tramontano, A.; Bozzoni, I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **2011**, *147*, 358–369.
30. Ørom, U.A.; Derrien, T.; Beringer, M.; Gumireddy, K.; Gardini, A.; Bussotti, G.; Lai, F.; Zytnicki, M.; Notredame, C.; Huang, Q.; *et al.* Long noncoding RNAs with enhancer-like function in human cells. *Cell* **2010**, *143*, 46–58.

31. Lai, F.; Orom, U.A.; Cesaroni, M.; Beringer, M.; Taatjes, D.J.; Blobel, G.A.; Shiekhattar, R. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* **2013**, *494*, 497–501.
32. Rinn, J.L.; Chang, H.Y. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **2012**, *81*, 145–166.
33. Amaral, P.P.; Clark, M.B.; Gascoigne, D.K.; Dinger, M.E.; Mattick, J.S. lncRNAdb: A reference database for long noncoding RNAs. *Nucleic Acids Res.* **2011**, *39*, D146–D151.
34. Ravasi, T.; Suzuki, H.; Pang, K.C.; Katayama, S.; Furuno, M.; Okunishi, R.; Fukuda, S.; Ru, K.; Frith, M.C.; Gongora, M.M.; *et al.* Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res.* **2006**, *16*, 11–19.
35. Wang, X.; Arai, S.; Song, X.; Reichart, D.; Du, K.; Pascual, G.; Tempst, P.; Rosenfeld, M.G.; Glass, C.K.; Kurokawa, R. Induced ncRNAs allosterically modify RNA-binding proteins in *cis* to inhibit transcription. *Nature* **2008**, *454*, 126–130.
36. Rinn, J.L.; Kertesz, M.; Wang, J.K.; Squazzo, S.L.; Xu, X.; Brugmann, S.A.; Goodnough, L.H.; Helms, J.A.; Farnham, P.J.; Segal, E.; *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **2007**, *129*, 1311–1323.
37. Rodriguez, A.; Griffiths-Jones, S.; Ashurst, J.L.; Bradley, A. Identification of mammalian microRNA host genes and transcription units. *Genome Res.* **2004**, *14*, 1902–1910.
38. Kapranov, P.; Cheng, J.; Dike, S.; Nix, D.A.; Duttagupta, R.; Willingham, A.T.; Stadler, P.F.; Hertel, J.; Hackermuller, J.; Hofacker, I.L.; *et al.* RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **2007**, *316*, 1484–1488.
39. Ogawa, Y.; Sun, B.K.; Lee, J.T. Intersection of the RNA interference and X-inactivation pathways. *Science* **2008**, *320*, 1336–1341.
40. Derrien, T.; Johnson, R.; Bussotti, G.; Tanzer, A.; Djebali, S.; Tilgner, H.; Guernec, G.; Martin, D.; Merkel, A.; Knowles, D.G.; *et al.* The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.* **2012**, *22*, 1775–1789.
41. Chen, Z.; Duan, X. Ribosomal RNA depletion for massively parallel bacterial RNA-sequencing applications. *Methods Mol. Biol.* **2011**, *733*, 93–103.
42. Guttman, M.; Garber, M.; Levin, J.Z.; Donaghey, J.; Robinson, J.; Adiconis, X.; Fan, L.; Koziol, M.J.; Gnirke, A.; Nusbaum, C.; *et al. Ab. initio* reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat. Biotechnol.* **2010**, *28*, 503–510.
43. Cabili, M.N.; Trapnell, C.; Goff, L.; Koziol, M.; Tazon-Vega, B.; Regev, A.; Rinn, J.L. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **2011**, *25*, 1915–1927.
44. Kutter, C.; Watt, S.; Stefflova, K.; Wilson, M.D.; Goncalves, A.; Ponting, C.P.; Odom, D.T.; Marques, A.C. Rapid turnover of long noncoding rnas and the evolution of gene expression. *PLoS Genet.* **2012**, *8*, e1002841.
45. ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **2007**, *447*, 799–816.

46. Clark, M.B.; Amaral, P.P.; Schlesinger, F.J.; Dinger, M.E.; Taft, R.J.; Rinn, J.L.; Ponting, C.P.; Stadler, P.F.; Morris, K.V.; Morillon, A.; *et al.* The reality of pervasive transcription. *PLoS Biol.* **2011**, *9*, doi:10.1371/journal.pbio.1000625.
47. Capel, B.; Swain, A.; Nicolis, S.; Hacker, A.; Walter, M.; Koopman, P.; Goodfellow, P.; Lovell-Badge, R. Circular transcripts of the testis-determining gene Sry in adult mouse testis. *Cell* **1993**, *73*, 1019–1030.
48. Cocquerelle, C.; Mascrez, B.; Hétuin, D.; Bailleul, B. Mis-splicing yields circular RNA molecules. *FASEB J.* **1993**, *7*, 155–160.
49. Nigro, J.M.; Cho, K.R.; Fearon, E.R.; Kern, S.E.; Ruppert, J.M.; Oliner, J.D.; Kinzler, K.W.; Vogelstein, B. Scrambled exons. *Cell* **1991**, *64*, 607–613.
50. Zaphiropoulos, P.G. Exon skipping and circular RNA formation in transcripts of the human cytochrome P-450 2C18 gene in epidermis and of the rat androgen binding protein gene in testis. *Mol. Cell. Biol.* **1997**, *17*, 2985–2993.
51. Memczak, S.; Jens, M.; Elefsinioti, A.; Torti, F.; Krueger, J.; Rybak, A.; Maier, L.; Mackowiak, S.D.; Gregersen, L.H.; Munschauer, M.; *et al.* Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **2013**, *495*, 333–338.
52. Hansen, T.B.; Jensen, T.I.; Clausen, B.H.; Bramsen, J.B.; Finsen, B.; Damgaard, C.K.; Kjems, J. Natural RNA circles function as efficient microRNA sponges. *Nature* **2013**, *495*, 384–388.
53. Jeck, W.R.; Sorrentino, J.A.; Wang, K.; Slevin, M.K.; Burd, C.E.; Liu, J.; Marzluff, W.F.; Sharpless, N.E. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **2013**, *19*, 141–157.
54. Capriotti, E.; Marti-Renom, M.A. Quantifying the relationship between sequence and three-dimensional structure conservation in RNA. *BMC Bioinforma.* **2010**, *11*, 322.
55. Rost, B. Twilight zone of protein sequence alignments. *Protein Eng.* **1999**, *12*, 85–94.
56. Pang, K.C.; Frith, M.C.; Mattick, J.S. Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends Genet.* **2006**, *22*, 1–5.
57. Bernhart, S.H.; Hofacker, I.L. From consensus structure prediction to RNA gene finding. *Brief Funct. Genomic Proteomic* **2009**, *8*, 461–471.
58. Sun, Y.; Aljawad, O.; Lei, J.; Liu, A. Genome-scale NCRNA homology search using a Hamming distance-based filtration strategy. *BMC Bioinforma.* **2012**, *13*, S12.
59. Bentwich, I.; Avniel, A.; Karov, Y.; Aharonov, R.; Gilad, S.; Barad, O.; Barzilai, A.; Einat, P.; Einav, U.; Meiri, E.; *et al.* Identification of hundreds of conserved and nonconserved human microRNAs. *Nat. Genet.* **2005**, *37*, 766–770.
60. Berezikov, E.; van Tetering, G.; Verheul, M.; van de Belt, J.; van Laake, L.; Vos, J.; Verloop, R.; van de Wetering, M.; Guryev, V.; Takada, S.; *et al.* Many novel mammalian microRNA candidates identified by extensive cloning and RAKE analysis. *Genome Res.* **2006**, *16*, 1289–1298.
61. Guerra-Assuncao, J.A.; Enright, A.J. Large-scale analysis of microRNA evolution. *BMC Genomics* **2012**, *13*, 218.
62. Missal, K.; Rose, D.; Stadler, P.F. Non-coding RNAs in *Ciona intestinalis*. *Bioinformatics* **2005**, *21*, ii77–ii78.
63. Lindgreen, S.; Gardner, P.P.; Krogh, A. MASTR: Multiple alignment and structure prediction of non-coding RNAs using simulated annealing. *Bioinformatics* **2007**, *23*, 3304–3311.

64. Sperschneider, J.; Datta, A.; Wise, M.J. Predicting pseudoknotted structures across two RNA sequences. *Bioinformatics* **2012**, *28*, 3058–3065.
65. Wong, T.K.F.; Wan, K.-L.; Hsu, B.-Y.; Cheung, B.W.Y.; Hon, W.-K.; Lam, T.-W.; Yiu, S.-M. RNASAlign: RNA structural alignment system. *Bioinformatics* **2011**, *27*, 2151–2152.
66. Gardner, P.P.; Giegerich, R. A comprehensive comparison of comparative RNA structure prediction approaches. *BMC Bioinforma.* **2004**, *5*, 140.
67. Ravindran, P.P.; Heroux, A.; Ye, J.D. Improvement of the crystallizability and expression of an RNA crystallization chaperone. *J. Biochem.* **2011**, *150*, 535–543.
68. Furtig, B.; Richter, C.; Wohnert, J.; Schwalbe, H. NMR spectroscopy of RNA. *Chembiochem* **2003**, *4*, 936–962.
69. Tzakos, A.G.; Grace, C.R.; Lukavsky, P.J.; Riek, R. NMR techniques for very large proteins and rnas in solution. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 319–342.
70. Zuker, M.; Stiegler, P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.* **1981**, *9*, 133–148.
71. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **2003**, *31*, 3406–3415.
72. Mathews, D.H.; Turner, D.H.; Zuker, M. RNA secondary structure prediction. *Curr. Protoc. Nucleic Acid Chem.* **2007**, doi:10.1002/0471142700.nc1102s28.
73. Dowell, R.D.; Eddy, S.R. Evaluation of several lightweight stochastic context-free grammars for RNA secondary structure prediction. *BMC Bioinforma.* **2004**, *5*, 71.
74. Dima, R.I.; Hyeon, C.; Thirumalai, D. Extracting stacking interaction parameters for RNA from the data set of native structures. *J. Mol. Biol.* **2005**, *347*, 53–69.
75. Do, C.B.; Woods, D.A.; Batzoglou, S. CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics* **2006**, *22*, e90–e98.
76. Xia, T.; SantaLucia, J.J.; Burkard, M.E.; Kierzek, R.; Schroeder, S.J.; Jiao, X.; Cox, C.; Turner, D.H. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* **1998**, *37*, 14719–14735.
77. Mathews, D.H.; Sabina, J.; Zuker, M.; Turner, D.H. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **1999**, *288*, 911–940.
78. Mathews, D.H.; Disney, M.D.; Childs, J.L.; Schroeder, S.J.; Zuker, M.; Turner, D.H. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7287–7292.
79. Lu, Z.J.; Turner, D.H.; Mathews, D.H. A set of nearest neighbor parameters for predicting the enthalpy change of RNA secondary structure formation. *Nucleic Acids Res.* **2006**, *34*, 4912–4924.
80. Lu, Z.J.; Gloor, J.W.; Mathews, D.H. Improved RNA secondary structure prediction by maximizing expected pair accuracy. *RNA* **2009**, *15*, 1805–1813.
81. Hofacker, I.L.; Fontana W.; Stadler, P.F.; Bonhoeffer S.; Tacker M.; Schuster, P. Fast folding and comparison of rna secondary structures. *Monatshefte f. Chem.* **1994**, *125*, 167–188.
82. McCaskill, J.S. The equilibrium partition function and base pair binding probabilities for RNA secondary structure. *Biopolymers* **1990**, *29*, 1105–1119.

83. Durbin, R.; Eddy, S.R.; Krogh, A.; Mitchison, G. *Biological Sequence Analysis. Probabilistic Models of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, UK, 1998.
84. Deigan, K.E.; Li, T.W.; Mathews, D.H.; Weeks, K.M. Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 97–102.
85. Doshi, K.J.; Cannone, J.J.; Cobaugh, C.W.; Gutell, R.R. Evaluation of the suitability of free-energy minimization using nearest-neighbor energy parameters for RNA secondary structure prediction. *BMC Bioinforma.* **2004**, *5*, 105.
86. Herschlag, D. RNA chaperones and the RNA folding problem. *J. Biol. Chem.* **1995**, *270*, 20871–20874.
87. Brennicke, A.; Marchfelder, A.; Binder, S. RNA editing. *FEMS Microbiol. Rev.* **1999**, *23*, 297–316.
88. Pan, T.; Sosnick, T. RNA folding during transcription. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 161–175.
89. Mandal, M.; Breaker, R.R. Gene regulation by riboswitches. *Nat. Rev. Mol. Cell. Biol.* **2004**, *5*, 451–463.
90. Soukup, J.K.; Soukup, G.A. Riboswitches exert genetic control through metabolite-induced conformational change. *Curr. Opin. Struct. Biol.* **2004**, *14*, 344–349.
91. Bengert, P.; Dandekar, T. Riboswitch finder—A tool for identification of riboswitch RNAs. *Nucleic Acids Res.* **2004**, *32*, W154–W159.
92. Voss, B.; Meyer, C.; Giegerich, R. Evaluating the predictability of conformational switching in RNA. *Bioinformatics* **2004**, *20*, 1573–1582.
93. Hofacker, I.L. Vienna RNA secondary structure server. *Nucleic Acids Res.* **2003**, *31*, 3429–3431.
94. Sankoff, D. Simultaneous solution of the RNA folding, alignment and protosequence problems. *SIAM J. Appl. Math.* **1985**, *45*, 810–825.
95. Torarinsson, E.; Lindgreen, S. WAR: Webserver for aligning structural RNAs. *Nucleic Acids Res.* **2008**, *36*, W79–W84.
96. Bremges, A.; Schirmer, S.; Giegerich, R. Fine-tuning structural RNA alignments in the twilight zone. *BMC Bioinforma.* **2010**, *11*, 222.
97. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680.
98. Notredame, C.; Higgins, D.G.; Heringa, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **2000**, *302*, 205–217.
99. Hofacker, I.L.; Fekete, M.; Stadler, P.F. Secondary structure prediction for aligned RNA sequences. *J. Mol. Biol.* **2002**, *319*, 1059–1066.
100. Knudsen, B.; Hein, J. Pfold: RNA secondary structure prediction using stochastic context-free grammars. *Nucleic Acids Res.* **2003**, *31*, 3423–3428.
101. Ruan, J.; Stormo, G.D.; Zhang, W. An iterated loop matching approach to the prediction of RNA secondary structures with pseudoknots. *Bioinformatics* **2004**, *20*, 58–66.
102. Luck, R.; Graf, S.; Steger, G. ConStruct: A tool for thermodynamic controlled prediction of conserved secondary structure. *Nucleic Acids Res.* **1999**, *27*, 4208–4217.

103. Mathews, D.H.; Turner, D.H. Dynalign: An algorithm for finding the secondary structure common to two RNA sequences. *J. Mol. Biol.* **2002**, *317*, 191–203.
104. Mathews, D.H. Predicting a set of minimal free energy RNA secondary structures common to two sequences. *Bioinformatics* **2005**, *21*, 2246–2253.
105. Gorodkin, J.; Heyer, L.J.; Stormo, G.D. Finding the most significant common sequence and structure motifs in a set of RNA sequences. *Nucleic Acids Res.* **1997**, *25*, 3724–3732.
106. Havgaard, J.H.; Lyngso, R.B.; Stormo, G.D.; Gorodkin, J. Pairwise local structural alignment of RNA sequences with sequence similarity less than 40%. *Bioinformatics* **2005**, *21*, 1815–1824.
107. Holmes, I. Accelerated probabilistic inference of RNA structure evolution. *BMC Bioinforma.* **2005**, *6*, 73.
108. Dowell, R.D.; Eddy, S.R. Efficient pairwise RNA structure prediction and alignment using sequence alignment constraints. *BMC Bioinforma.* **2006**, *7*, 400.
109. Hofacker, I.L.; Bernhart, S.H.; Stadler, P.F. Alignment of RNA base pairing probability matrices. *Bioinformatics* **2004**, *20*, 2222–2227.
110. Wilm, A.; Higgins, D.G.; Notredame, C. R-Coffee: A method for multiple alignment of non-coding RNA. *Nucleic Acids Res.* **2008**, *36*, e52.
111. Reeder, J.; Giegerich, R. Consensus shapes: An alternative to the Sankoff algorithm for RNA consensus structure prediction. *Bioinformatics* **2005**, *21*, 3516–3523.
112. Bernhart, S.H.; Hofacker, I.L.; Stadler, P.F. Local RNA base pairing probabilities in large sequences. *Bioinformatics* **2006**, *22*, 614–615.
113. Capriotti, E.; Marti-Renom, M.A. RNA structure alignment by a unit-vector approach. *Bioinformatics* **2008**, *24*, i112–i118.
114. Ferre, F.; Ponty, Y.; Lorenz, W.A.; Clote, P. DIAL: A web server for the pairwise alignment of two RNA three-dimensional structures using nucleotide, dihedral angle and base-pairing similarities. *Nucleic Acids Res.* **2007**, *35*, W659–W668.
115. Wang, C.W.; Chen, K.T.; Lu, C.L. iPARTS: An improved tool of pairwise alignment of RNA tertiary structures. *Nucleic Acids Res.* **2010**, *38*, W340–W347.
116. Dror, O.; Nussinov, R.; Wolfson, H. ARTS: Alignment of RNA tertiary structures. *Bioinformatics* **2005**, *21*, ii47–ii53.
117. Chang, Y.F.; Huang, Y.L.; Lu, C.L. SARSA: A web tool for structural alignment of RNA using a structural alphabet. *Nucleic Acids Res.* **2008**, *36*, W19–W24.
118. Bauer, R.P.; Rother, K.P.; Moor, P.P.; Reinert, K.P.; Steinke, T.P.; Bujnicki, J.P.; Preissner, R.P. Fast structural alignment of biomolecules using a hash table, N-grams and string descriptors. *Algorithms* **2009**, *2*, 692–709.
119. Kirillova, S.; Tosatto, S.C.; Carugo, O. FRASS: The web-server for RNA structural comparison. *BMC Bioinforma.* **2010**, *11*, 327.
120. Freyhult, E.K.; Bollback, J.P.; Gardner, P.P. Exploring genomic dark matter: A critical assessment of the performance of homology search methods on noncoding RNA. *Genome Res.* **2007**, *17*, 117–125.
121. Smith, T.F.; Waterman, M.S. Identification of common molecular subsequences. *J. Mol. Biol.* **1981**, *147*, 195–197.

122. Rognes, T. Faster Smith-Waterman database searches with inter-sequence SIMD parallelisation. *BMC Bioinforma.* **2011**, *12*, 221.
123. Lipman, D.J.; Pearson, W.R. Rapid and sensitive protein similarity searches. *Science* **1985**, *227*, 1435–1441.
124. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410.
125. Eddy, S.R. Profile hidden Markov models. *Bioinformatics* **1998**, *14*, 755–763.
126. Weinberg, Z.; Ruzzo, W.L. Sequence-based heuristics for faster annotation of non-coding RNA families. *Bioinformatics* **2006**, *22*, 35–39.
127. Holmes, I. A probabilistic model for the evolution of RNA structure. *BMC Bioinforma.* **2004**, *5*, 166.
128. Notredame, C.; O'Brien, E.A.; Higgins, D.G. RAGA: RNA sequence alignment by genetic algorithm. *Nucleic Acids Res.* **1997**, *25*, 4570–4580.
129. Eddy, S.R.; Durbin, R. RNA sequence analysis using covariance models. *Nucleic Acids Res.* **1994**, *22*, 2079–2088.
130. Eddy, S.R. A memory-efficient dynamic programming algorithm for optimal alignment of a sequence to an RNA secondary structure. *BMC Bioinforma.* **2002**, *3*, 18.
131. Klein, R.J.; Eddy, S.R. RSEARCH: Finding homologs of single structured RNA sequences. *BMC Bioinforma.* **2003**, *4*, 44.
132. Griffiths-Jones, S.; Bateman, A.; Marshall, M.; Khanna, A.; Eddy, S.R. Rfam: An RNA family database. *Nucleic Acids Res.* **2003**, *31*, 439–441.
133. Finn, R.D.; Tate, J.; Mistry, J.; Coggill, P.C.; Sammut, S.J.; Hotz, H.R.; Ceric, G.; Forslund, K.; Eddy, S.R.; Sonnhammer, E.L.; *et al.* The Pfam protein families database. *Nucleic Acids Res.* **2008**, *36*, D281–D288.
134. Griffiths-Jones, S.; Moxon, S.; Marshall, M.; Khanna, A.; Eddy, S.R.; Bateman, A. Rfam: Annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* **2005**, *33*, D121–D124.
135. Zhang, S.; Borovok, I.; Aharonowitz, Y.; Sharan, R.; Bafna, V. A sequence-based filtering method for ncRNA identification and its application to searching for riboswitch elements. *Bioinformatics* **2006**, *22*, e557–e565.
136. Gardner, P.P.; Daub, J.; Tate, J.G.; Nawrocki, E.P.; Kolbe, D.L.; Lindgreen, S.; Wilkinson, A.C.; Finn, R.D.; Griffiths-Jones, S.; Eddy, S.R.; *et al.* Rfam: Updates to the RNA families database. *Nucleic Acids Res.* **2009**, *37*, D136–D140.
137. Roshan, U.; Chikkagoudar, S.; Livesay, D.R. Searching for evolutionary distant RNA homologs within genomic sequences using partition function posterior probabilities. *BMC Bioinforma.* **2008**, *9*, 61.
138. Bussotti, G.; Raineri, E.; Erb, I.; Zytnicki, M.; Wilm, A.; Beaudoing, E.; Bucher, P.; Notredame, C. BlastR—Fast and accurate database searches for non-coding RNAs. *Nucleic Acids Res.* **2011**, *39*, 6886–6895.
139. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; *et al.* Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921.

140. Venter, J.C.; Adams, M.D.; Myers, E.W.; Li, P.W.; Mural, R.J.; Sutton, G.G.; Smith, H.O.; Yandell, M.; Evans, C.A.; Holt, R.A.; *et al.* The sequence of the human genome. *Science* **2001**, *291*, 1304–1351.
141. Aparicio, S.; Chapman, J.; Stupka, E.; Putnam, N.; Chia, J.M.; Dehal, P.; Christoffels, A.; Rash, S.; Hoon, S.; Smit, A.; *et al.* Whole-genome shotgun assembly and analysis of the genome of Fugu rubripes. *Science* **2002**, *297*, 1301–1310.
142. Waterston, R.H.; Lindblad-Toh, K.; Birney, E.; Rogers, J.; Abril, J.F.; Agarwal, P.; Agarwala, R.; Ainscough, R.; Alexandersson, M.; An, P.; *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **2002**, *420*, 520–562.
143. Schwartz, S.; Kent, W.J.; Smit, A.; Zhang, Z.; Baertsch, R.; Hardison, R.C.; Haussler, D.; Miller, W. Human-mouse alignments with BLASTZ. *Genome Res.* **2003**, *13*, 103–107.
144. Boguski, M.S.; Tolstoshev, C.M.; Bassett, D.E.J. Gene discovery in dbEST. *Science* **1994**, *265*, 1993–1994.
145. Dias Neto, E.; Correa, R.G.; Verjovski-Almeida, S.; Briones, M.R.; Nagai, M.A.; da Silva, W.J.; Zago, M.A.; Bordin, S.; Costa, F.F.; Goldman, G.H.; *et al.* Shotgun sequencing of the human transcriptome with ORF expressed sequence tags. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3491–3496.
146. Gerhard, D.S.; Wagner, L.; Feingold, E.A.; Shenmen, C.M.; Grouse, L.H.; Schuler, G.; Klein, S.L.; Old, S.; Rasooly, R.; Good, P.; *et al.* The status, quality, and expansion of the NIH full-length cDNA project: The Mammalian Gene Collection (MGC). *Genome Res.* **2004**, *14*, 2121–2127.
147. Boguski, M.S.; Lowe, T.M.; Tolstoshev, C.M. dbEST—Database for "expressed sequence tags." *Nat. Genet.* **1993**, *4*, 332–333.
148. Adams, M.D.; Kelley, J.M.; Gocayne, J.D.; Dubnick, M.; Polymeropoulos, M.H.; Xiao, H.; Merril, C.R.; Wu, A.; Olde, B.; Moreno, R.F.; *et al.* Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science* **1991**, *252*, 1651–1656.
149. Bonaldo, M.F.; Lennon, G.; Soares, M.B. Normalization and subtraction: Two approaches to facilitate gene discovery. *Genome Res.* **1996**, *6*, 791–806.
150. Nagaraj, S.H.; Gasser, R.B.; Ranganathan, S. A hitchhiker's guide to expressed sequence tag (EST) analysis. *Brief. Bioinforma.* **2007**, *8*, 6–21.
151. Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **2008**, *5*, 621–628.
152. Malone, J.H.; Oliver, B. Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol.* **2011**, *9*, 34.
153. Eklund, A.C.; Turner, L.R.; Chen, P.; Jensen, R.V.; deFeo, G.; Kopf-Sill, A.R.; Szallasi, Z. Replacing cRNA targets with cDNA reduces microarray cross-hybridization. *Nat. Biotechnol.* **2006**, *24*, 1071–1073.
154. Okoniewski, M.J.; Miller, C.J. Hybridization interactions between probesets in short oligo microarrays lead to spurious correlations. *BMC Bioinforma.* **2006**, *7*, 276.
155. Casneuf, T.; van de Peer, Y.; Huber, W. *In situ* analysis of cross-hybridisation on microarrays and the inference of expression correlation. *BMC Bioinforma.* **2007**, *8*, 461.
156. Cox, W.G.; Beaudet, M.P.; Agnew, J.Y.; Ruth, J.L. Possible sources of dye-related signal correlation bias in two-color DNA microarray assays. *Anal. Biochem.* **2004**, *331*, 243–254.

157. Dombkowski, A.A.; Thibodeau, B.J.; Starcevic, S.L.; Novak, R.F. Gene-specific dye bias in microarray reference designs. *FEBS Lett.* **2004**, *560*, 120–124.
158. Rosenzweig, B.A.; Pine, P.S.; Domon, O.E.; Morris, S.M.; Chen, J.J.; Sistare, F.D. Dye bias correction in dual-labeled cDNA microarray gene expression measurements. *Environ. Health Perspect.* **2004**, *112*, 480–487.
159. Dobbin, K.K.; Kawasaki, E.S.; Petersen, D.W.; Simon, R.M. Characterizing dye bias in microarray experiments. *Bioinformatics* **2005**, *21*, 2430–2437.
160. Martin-Magniette, M.L.; Aubert, J.; Cabannes, E.; Daudin, J.J. Evaluation of the gene-specific dye bias in cDNA microarray experiments. *Bioinformatics* **2005**, *21*, 1995–2000.
161. Bertone, P.; Stolc, V.; Royce, T.E.; Rozowsky, J.S.; Urban, A.E.; Zhu, X.; Rinn, J.L.; Tongprasit, W.; Samanta, M.; Weissman, S.; *et al.* Global identification of human transcribed sequences with genome tiling arrays. *Science* **2004**, *306*, 2242–2246.
162. Cheng, J.; Kapranov, P.; Drenkow, J.; Dike, S.; Brubaker, S.; Patel, S.; Long, J.; Stern, D.; Tammana, H.; Helt, G.; *et al.* Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* **2005**, *308*, 1149–1154.
163. Royce, T.E.; Rozowsky, J.S.; Bertone, P.; Samanta, M.; Stolc, V.; Weissman, S.; Snyder, M.; Gerstein, M. Issues in the analysis of oligonucleotide tiling microarrays for transcript mapping. *Trends Genet.* **2005**, *21*, 466–475.
164. Kapranov, P.; Willingham, A.T.; Gingeras, T.R. Genome-wide transcription and the implications for genomic organization. *Nat. Rev. Genet.* **2007**, *8*, 413–423.
165. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **2009**, *10*, 57–63.
166. Marioni, J.C.; Mason, C.E.; Mane, S.M.; Stephens, M.; Gilad, Y. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* **2008**, *18*, 1509–1517.
167. Fu, X.; Fu, N.; Guo, S.; Yan, Z.; Xu, Y.; Hu, H.; Menzel, C.; Chen, W.; Li, Y.; Zeng, R.; *et al.* Estimating accuracy of RNA-Seq and microarrays with proteomics. *BMC Genomics* **2009**, *10*, 161.
168. Ozsolak, F.; Milos, P.M. RNA Sequencing: Advances, challenges and opportunities. *Nat. Rev. Genet.* **2011**, *12*, 87–98.
169. Velculescu, V.E.; Zhang, L.; Vogelstein, B.; Kinzler, K.W. Serial analysis of gene expression. *Science* **1995**, *270*, 484–487.
170. Harbers, M.; Carninci, P. Tag-based approaches for transcriptome research and genome annotation. *Nat. Methods* **2005**, *2*, 495–502.
171. Saha, S.; Sparks, A. B.; Rago, C.; Akmaev, V.; Wang, C.J.; Vogelstein, B.; Kinzler, K.W.; Velculescu, V.E. Using the transcriptome to annotate the genome. *Nat. Biotechnol.* **2002**, *20*, 508–512.
172. Gowda, M.; Jantasuriyarat, C.; Dean, R.A.; Wang, G.L. Robust-LongSAGE (RL-SAGE): A substantially improved LongSAGE method for gene discovery and transcriptome analysis. *Plant. Physiol.* **2004**, *134*, 890–897.
173. Matsumura, H.; Ito, A.; Saitoh, H.; Winter, P.; Kahl, G.; Reuter, M.; Kruger, D.H.; Terauchi, R. SuperSAGE. *Cell. Microbiol.* **2005**, *7*, 11–18.

174. Brenner, S.; Johnson, M.; Bridgham, J.; Golda, G.; Lloyd, D.H.; Johnson, D.; Luo, S.; McCurdy, S.; Foy, M.; Ewan, M.; *et al.* Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* **2000**, *18*, 630–634.
175. Ng, P.; Wei, C.L.; Sung, W.K.; Chiu, K.P.; Lipovich, L.; Ang, C.C.; Gupta, S.; Shahab, A.; Ridwan, A.; Wong, C.H.; *et al.* Gene identification signature (GIS) analysis for transcriptome characterization and genome annotation. *Nat. Methods* **2005**, *2*, 105–111.
176. Schaefer, B.C. Revolutions in rapid amplification of cDNA ends: New strategies for polymerase chain reaction cloning of full-length cDNA ends. *Anal. Biochem.* **1995**, *227*, 255–273.
177. Kapranov, P.; Drenkow, J.; Cheng, J.; Long, J.; Helt, G.; Dike, S.; Gingeras, T.R. Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays. *Genome Res.* **2005**, *15*, 987–997.
178. Olivarius, S.; Plessy, C.; Carninci, P. High-throughput verification of transcriptional starting sites by Deep-RACE. *BioTechniques* **2009**, *46*, 130–132.
179. Kodzius, R.; Kojima, M.; Nishiyori, H.; Nakamura, M.; Fukuda, S.; Tagami, M.; Sasaki, D.; Imamura, K.; Kai, C.; Harbers, M.; *et al.* CAGE: Cap analysis of gene expression. *Nat. Methods* **2006**, *3*, 211–222.
180. Shiraki, T.; Kondo, S.; Katayama, S.; Waki, K.; Kasukawa, T.; Kawaji, H.; Kodzius, R.; Watahiki, A.; Nakamura, M.; Arakawa, T.; *et al.* Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 15776–15781.
181. Mercer, T.R.; Gerhardt, D.J.; Dinger, M.E.; Crawford, J.; Trapnell, C.; Jeddeloh, J.A.; Mattick, J.S.; Rinn, J.L. Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nat. Biotechnol.* **2011**, *30*, 99–104.
182. Mathavan, S.; Lee, S.G.; Mak, A.; Miller, L.D.; Murthy, K.R.; Govindarajan, K.R.; Tong, Y.; Wu, Y.L.; Lam, S.H.; Yang, H.; *et al.* Transcriptome analysis of zebrafish embryogenesis using microarrays. *PLoS Genet.* **2005**, *1*, 260–276.
183. Wang, E.T.; Sandberg, R.; Luo, S.; Khrebtukova, I.; Zhang, L.; Mayr, C.; Kingsmore, S.F.; Schroth, G.P.; Burge, C.B. Alternative isoform regulation in human tissue transcriptomes. *Nature* **2008**, *456*, 470–476.
184. Graveley, B.R.; Brooks, A.N.; Carlson, J.W.; Duff, M.O.; Landolin, J.M.; Yang, L.; Artieri, C.G.; van Baren, M.J.; Boley, N.; Booth, B.W.; *et al.* The developmental transcriptome of Drosophila melanogaster. *Nature* **2011**, *471*, 473–479.
185. Pang, K.C.; Stephen, S.; Dinger, M.E.; Engström, P.G.; Lenhard, B.; Mattick, J.S. RNAdb 2.0—An expanded database of mammalian non-coding RNAs. *Nucleic Acids Res.* **2007**, *35*, D178–D182.
186. Kin, T.; Yamada, K.; Terai, G.; Okida, H.; Yoshinari, Y.; Ono, Y.; Kojima, A.; Kimura, Y.; Komori, T.; Asai, K. fRNAdb: A platform for mining/annotating functional RNA candidates from non-coding RNA sequences. *Nucleic Acids Res.* **2007**, *35*, D145–D148.
187. Liu, C.; Bai, B.; Skogerbø, G.; Cai, L.; Deng, W.; Zhang, Y.; Bu, D.; Zhao, Y.; Chen, R. NONCODE: An integrated knowledge database of non-coding RNAs. *Nucleic Acids Res.* **2005**, *33*, D112–D125.
188. Pruitt, K.D.; Tatusova, T.; Klimke, W.; Maglott, D.R. NCBI reference Sequences: Current status, policy and new initiatives. *Nucleic Acids Res.* **2009**, *37*, D32–D36.

189. Harrow, J.; Denoeud, F.; Frankish, A.; Reymond, A.; Chen, C.K.; Chrast, J.; Lagarde, J.; Gilbert, J.G.; Storey, R.; Swarbreck, D.; *et al.* GENCODE: Producing a reference annotation for ENCODE. *Genome Biol.* **2006**, *7*, S4 1–S4 9.
190. Havana team. Available online: http://www.sanger.ac.uk/research/projects/vertebrategenome/havana/ (accessed 6 March 2013).
191. Loveland, J.E.; Gilbert, J.G.; Griffiths, E.; Harrow, J.L. Community gene annotation in practice. *Database (Oxford)* **2012**, *2012*, doi:10.1093/database/bas009.
192. Flicek, P.; Amode, M.R.; Barrell, D.; Beal, K.; Brent, S.; Carvalho-Silva, D.; Clapham, P.; Coates, G.; Fairley, S.; Fitzgerald, S.; *et al.* Ensembl 2012. *Nucleic Acids Res.* **2012**, *40*, D84–D90.
193. Kawai, J.; Shinagawa, A.; Shibata, K.; Yoshino, M.; Itoh, M.; Ishii, Y.; Arakawa, T.; Hara, A.; Fukunishi, Y.; Konno, H.; *et al.* Functional annotation of a full-length mouse cDNA collection. *Nature* **2001**, *409*, 685–690.
194. Pruitt, K.D.; Tatusova, T.; Brown, G.R.; Maglott, D.R. NCBI reference sequences (RefSeq): Current status, new features and genome annotation policy. *Nucleic Acids Res.* **2012**, *40*, D130–D135.
195. Kent, W.J.; Sugnet, C.W.; Furey, T.S.; Roskin, K.M.; Pringle, T.H.; Zahler, A.M.; Haussler, D. The human genome browser at UCSC. *Genome Res.* **2002**, *12*, 996–1006.
196. Stalker, J.; Gibbins, B.; Meidl, P.; Smith, J.; Spooner, W.; Hotz, H.R.; Cox, A.V. The Ensembl web site: Mechanics of a genome browser. *Genome Res.* **2004**, *14*, 951–955.
197. Wilming, L.G.; Gilbert, J.G.; Howe, K.; Trevanion, S.; Hubbard, T.; Harrow, J.L. The vertebrate genome annotation (Vega) database. *Nucleic Acids Res.* **2008**, *36*, D753–D760.
198. ENCODE Project Consortium An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74.
199. Martin, J.A.; Wang, Z. Next-generation transcriptome assembly. *Nat. Rev. Genet.* **2011**, *12*, 671–682.
200. Rogers, M.F.; Thomas, J.; Reddy, A.S.; Ben-Hur, A. SpliceGrapher: Detecting patterns of alternative splicing from RNA-Seq data in the context of gene models and EST data. *Genome Biol.* **2012**, *13*, R4.
201. Metzker, M.L. Sequencing technologies—the next generation. *Nat. Rev. Genet.* **2010**, *11*, 31–46.
202. Butler, J.; MacCallum, I.; Kleber, M.; Shlyakhter, I.A.; Belmonte, M.K.; Lander, E.S.; Nusbaum, C.; Jaffe, D.B. ALLPATHS: *de novo* assembly of whole-genome shotgun microreads. *Genome Res.* **2008**, *18*, 810–820.
203. Zerbino, D.R.; Birney, E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **2008**, *18*, 821–829.
204. Simpson, J.T.; Wong, K.; Jackman, S.D.; Schein, J.E.; Jones, S.J.; Birol, I. ABySS: A parallel assembler for short read sequence data. *Genome Res.* **2009**, *19*, 1117–1123.
205. Li, R.; Li, Y.; Kristiansen, K.; Wang, J. SOAP: Short oligonucleotide alignment program. *Bioinformatics* **2008**, *24*, 713–714.
206. Lin, H.; Zhang, Z.; Zhang, M.Q.; Ma, B.; Li, M. ZOOM! Zillions of oligos mapped. *Bioinformatics* **2008**, *24*, 2431–2437.
207. Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **2009**, *10*, R25.

208. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760.
209. Schatz, M.C. CloudBurst: Highly sensitive read mapping with MapReduce. *Bioinformatics* **2009**, *25*, 1363–1369.
210. Ahmadi, A.; Behm, A.; Honnalli, N.; Li, C.; Weng, L.; Xie, X. Hobbes: Optimized gram-based methods for efficient read alignment. *Nucleic Acids Res.* **2012**, *40*, e41.
211. Derrien, T.; Estelle, J.; Marco Sola, S.; Knowles, D.G.; Raineri, E.; Guigo, R.; Ribeca, P. Fast computation and applications of genome mappability. *PLoS One* **2012**, *7*, e30377.
212. Cloonan, N.; Xu, Q.; Faulkner, G.J.; Taylor, D.F.; Tang, D.T.; Kolle, G.; Grimmond, S.M. RNA-MATE: A recursive mapping strategy for high-throughput RNA-sequencing data. *Bioinformatics* **2009**, *25*, 2615–2616.
213. Trapnell, C.; Pachter, L.; Salzberg, S.L. TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **2009**, *25*, 1105–1111.
214. Li, H.; Ruan, J.; Durbin, R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.* **2008**, *18*, 1851–1858.
215. Smith, A.D.; Xuan, Z.; Zhang, M.Q. Using quality scores and longer reads improves accuracy of Solexa read mapping. *BMC Bioinforma.* **2008**, *9*, 128.
216. Ewing, B.; Green, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **1998**, *8*, 186–194.
217. Ewing, B.; Hillier, L.; Wendl, M.C.; Green, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **1998**, *8*, 175–185.
218. Trapnell, C.; Williams, B.A.; Pertea, G.; Mortazavi, A.; Kwan, G.; van Baren, M.J.; Salzberg, S.L.; Wold, B.J.; Pachter, L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **2010**, *28*, 511–515.
219. Li, W.; Feng, J.; Jiang, T. IsoLasso: A LASSO regression approach to RNA-Seq based transcriptome assembly. *J. Comput. Biol.* **2011**, *18*, 1693–1707.
220. Palmieri, N.; Nolte, V.; Suvorov, A.; Kosiol, C.; Schlötterer, C. Evaluation of different reference based annotation strategies using RNA-Seq—A case study in *Drososphila pseudoobscura*. *PLoS One* **2012**, doi:10.1371/journal.pone.0046415.
221. Garg, R.; Patel, R.K.; Tyagi, A.K.; Jain, M. *De novo* assembly of chickpea transcriptome using short reads for gene discovery and marker identification. *DNA Res.* **2011**, *18*, 53–63.
222. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652.
223. Jager, M.; Ott, C.E.; Grunhagen, J.; Hecht, J.; Schell, H.; Mundlos, S.; Duda, G.N.; Robinson, P.N.; Lienau, J. Composite transcriptome assembly of RNA-seq data in a sheep model for delayed bone healing. *BMC Genomics* **2011**, *12*, 158.
224. Keiler, K.C. Biology of trans-translation. *Annu. Rev. Microbiol.* **2008**, *62*, 133–151.
225. Novikova, I.V.; Hennelly, S.P.; Sanbonmatsu, K.Y. Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. *Nucleic Acids Res.* **2012**, *40*, 5034–5051.

226. Clamp, M.; Fry, B.; Kamal, M.; Xie, X.; Cuff, J.; Lin, M.F.; Kellis, M.; Lindblad-Toh, K.; Lander, E.S. Distinguishing protein-coding and noncoding genes in the human genome. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19428–19433.
227. Lin, M.F.; Jungreis, I.; Kellis, M. PhyloCSF: A comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics* **2011**, *27*, i275–i282.
228. Lin, M.F.; Carlson, J.W.; Crosby, M.A.; Matthews, B.B.; Yu, C.; Park, S.; Wan, K.H.; Schroeder, A.J.; Gramates, L.S.; St Pierre, S.E.; *et al.* Revisiting the protein-coding gene catalog of Drosophila melanogaster using 12 fly genomes. *Genome Res.* **2007**, *17*, 1823–1836.
229. Liao, Q.; Liu, C.; Yuan, X.; Kang, S.; Miao, R.; Xiao, H.; Zhao, G.; Luo, H.; Bu, D.; Zhao, H.; *et al.* Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network. *Nucleic Acids Res.* **2011**, *39*, 3864–3878.
230. Marchler-Bauer, A.; Panchenko, A.R.; Shoemaker, B.A.; Thiessen, P.A.; Geer, L.Y.; Bryant, S.H. CDD: A database of conserved domain alignments with links to domain three-dimensional structure. *Nucleic Acids Res.* **2002**, *30*, 281–283.
231. UniProt Consortium. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* **2012**, *40*, D71–D75.
232. Kong, L.; Zhang, Y.; Ye, Z.-Q.; Liu, X.-Q.; Zhao, S.-Q.; Wei, L.; Gao, G. CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* **2007**, *35*, W345–W349.
233. Arrial, R.T.; Togawa, R.C.; Brigido, M.M. Screening non-coding RNAs in transcriptomes from neglected species using PORTRAIT: Case study of the pathogenic fungus Paracoccidioides brasiliensis. *BMC Bioinforma.* **2009**, *10*, 239.
234. Wang, L.; Park, H.J.; Dasari, S.; Wang, S.; Kocher, J.-P.; Li, W. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Res.* **2013**, *41*, e74.
235. Mikkelsen, T.S.; Ku, M.; Jaffe, D.B.; Issac, B.; Lieberman, E.; Giannoukos, G.; Alvarez, P.; Brockman, W.; Kim, T.K.; Koche, R.P.; *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **2007**, *448*, 553–560.
236. Lu, Z.J.; Yip, K.Y.; Wang, G.; Shou, C.; Hillier, L.W.; Khurana, E.; Agarwal, A.; Auerbach, R.; Rozowsky, J.; Cheng, C.; *et al.* Prediction and characterization of noncoding RNAs in *C. elegans* by integrating conservation, secondary structure, and high-throughput sequencing and array data. *Genome Res.* **2011**, *21*, 276–285.
237. Esteve-Codina, A.; Kofler, R.; Palmieri, N.; Bussotti, G.; Notredame, C.; Perez-Enciso, M. Exploring the gonad transcriptome of two extreme male pigs with RNA-seq. *BMC Genomics* **2011**, *12*, 552.
238. Nam, J.W.; Bartel, D. Long non-coding RNAs in *C. elegans. Genome Res.* **2012**, *22*, 2529–2540.
239. Djebali, S.; Davis, C.A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F.; *et al.* Landscape of transcription in human cells. *Nature* **2012**, *489*, 101–108.
240. Eveland, A.L.; McCarty, D.R.; Koch, K.E. Transcript profiling by 3'-untranslated region sequencing resolves expression of gene families. *Plant Physiol.* **2008**, *146*, 32–44.

241. Hillier, L.W.; Reinke, V.; Green, P.; Hirst, M.; Marra, M.A.; Waterston, R.H. Massively parallel sequencing of the polyadenylated transcriptome of *C. elegans*. *Genome Res.* **2009**, *19*, 657–666.
242. Nagalakshmi, U.; Wang, Z.; Waern, K.; Shou, C.; Raha, D.; Gerstein, M.; Snyder, M. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **2008**, *320*, 1344–1349.
243. Shendure, J. The beginning of the end for microarrays? *Nat. Methods* **2008**, *5*, 585–587.
244. Morozova, O.; Hirst, M.; Marra, M.A. Applications of new sequencing technologies for transcriptome analysis. *Annu. Rev. Genomics Hum. Genet.* **2009**, *10*, 135–151.
245. Auer, P.L.; Doerge, R.W. Statistical design and analysis of RNA sequencing data. *Genetics* **2010**, *185*, 405–416.
246. Hiller, M.; Pudimat, R.; Busch, A.; Backofen, R. Using RNA secondary structures to guide sequence motif finding towards single-stranded regions. *Nucleic Acids Res.* **2006**, *34*, e117.
247. Chang, T.-H.; Huang, H.-D.; Chuang, T.-N.; Shien, D.-M.; Horng, J.-T. RNAMST: Efficient and flexible approach for identifying RNA structural homologs. *Nucleic Acids Res.* **2006**, *34*, W423–W428.
248. Yao, Z.; Weinberg, Z.; Ruzzo, W.L. CMfinder—A covariance model based RNA motif finding algorithm. *Bioinformatics* **2006**, *22*, 445–452.
249. Ji, Y.; Xu, X.; Stormo, G.D. A graph theoretical approach for predicting common RNA secondary structure motifs including pseudoknots in unaligned sequences. *Bioinformatics* **2004**, *20*, 1591–1602.
250. Riccitelli, N.J.; Lupták, A. Computational discovery of folded RNA domains in genomes and *in vitro* selected libraries. *Methods* **2010**, *52*, 133–140.
251. Gautheret, D.; Major, F.; Cedergren, R. Pattern searching/alignment with RNA primary and secondary structures: An effective descriptor for tRNA. *Comput. Appl. Biosci.* **1990**, *6*, 325–331.
252. Gautheret, D.; Lambert, A. Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. *J. Mol. Biol.* **2001**, *313*, 1003–1011.
253. König, J.; Zarnack, K.; Luscombe, N.M.; Ule, J. Protein-RNA interactions: New genomic technologies and perspectives. *Nat. Rev. Genet.* **2011**, *13*, 77–83.
254. Rivas, E.; Eddy, S.R. Secondary structure alone is generally not statistically significant for the detection of noncoding RNAs. *Bioinformatics* **2000**, *16*, 583–605.
255. Kertesz, M.; Wan, Y.; Mazor, E.; Rinn, J.L.; Nutter, R.C.; Chang, H.Y.; Segal, E. Genome-wide measurement of RNA secondary structure in yeast. *Nature* **2010**, *467*, 103–107.
256. Slater, G.S.; Birney, E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinforma.* **2005**, *6*, 31.
257. Birney, E.; Clamp, M.; Durbin, R. GeneWise and genomewise. *Genome Res.* **2004**, *14*, 988–995.
258. Eyras, E.; Reymond, A.; Castelo, R.; Bye, J.M.; Camara, F.; Flicek, P.; Huckle, E.J.; Parra, G.; Shteynberg, D.D.; Wyss, C.; *et al.* Gene finding in the chicken genome. *BMC Bioinforma.* **2005**, *6*, 131.
259. Mariotti, M.; Guigo, R. Selenoprofiles: Profile-based scanning of eukaryotic genome sequences for selenoprotein genes. *Bioinformatics* **2010**, *26*, 2656–2663.

260. Vieira, F.G.; Rozas, J. Comparative genomics of the odorant-binding and chemosensory protein gene families across the Arthropoda: Origin and evolutionary history of the chemosensory system. *Genome Biol. Evol.* **2011**, *3*, 476–490.
261. Latos, P.A.; Pauler, F.M.; Koerner, M.V.; Şenergin, H.B.; Hudson, Q.J.; Stocsits, R.R.; Allhoff, W.; Stricker, S.H.; Klement, R.M.; Warczok, K.E.; *et al.* Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Science* **2012**, *338*, 1469–1472.
262. Santoro, F.; Pauler, F.M. Silencing by the imprinted Airn macro lncRNA: Transcription is the answer. *Cell. Cycle* **2013**, *12*, 711–712.

Reprinted from *IJMS*. Cite as: Zhang, C.; Li, G.; Wang, J.; Zhu, S.; Li, H. Cascading *cis*-Cleavage on Transcript from *trans*-Acting siRNA-Producing Locus 3. *Int. J. Mol. Sci.* **2013**, *14*, 14689-14699.

Article

Cascading *cis*-Cleavage on Transcript from *trans*-Acting siRNA-Producing Locus 3

Changqing Zhang [1,†,*], Guangping Li [2,†], Jin Wang [3], Shinong Zhu [1] and Hailing Li [1]

[1] College of Horticulture, Jinling Institute of Technology, Nanjing 210038, China; E-Mails: zsn@jit.edu.cn (S.Z.); lihailing@jit.edu.cn (H.L.)

[2] College of Forest Resources and Environment, Nanjing Forestry University, Nanjing 210037, China; E-Mail: liguangping108@sina.com

[3] State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, Nanjing 210093, China; E-Mail: jwang@nju.edu.cn

[†] These authors contributed equally to this work.

[*] Author to whom correspondence should be addressed; E-Mail: zhang_chq2002@sohu.com; Tel./Fax: +86-025-8539-3314.

Received: 2 May 2013; in revised form: 24 June 2013 / Accepted: 4 July 2013 / Published: 12 July 2013

Abstract: The production of small RNAs (sRNAs) from phased positions set by microRNA-directed cleavage of *trans*-acting-siRNA-producing locus (TAS) transcript has been characterized extensively; however, the production of sRNAs from non-phased positions remains unknown. We report three *cis*-cleavages that occurred in *TAS3* transcripts in *Vitis vinifera*, by combining high-throughput sRNA deep sequencing information with evolutional conservation and genome-wide RNA degradome analysis. The three *cis*-cleavages can be deciphered to generate an orderly cleavage cascade, and can also produce distinct phasing patterns. Each of the patterns, either upstream or downstream of the *cis*-cleaved position, had a set of sRNAs arranged in 21-nucleotide increments. Part of the cascading *cis*-cleavages was also conserved in *Arabidopsis thaliana*. Our results will enhance the understanding of the production of sRNAs from non-phased positions that are not set by microRNA-directed cleavage.

Keywords: *trans*-acting siRNA; *cis*-cleavage; ta-siRNA-producing locus; miRNA

1. Introduction

In plants, many endogenous small RNAs (sRNAs), including microRNAs (miRNAs), heterochromatic small interfering RNAs (siRNAs), natural antisense siRNAs, and *trans*-acting siRNAs (ta-siRNAs), play important roles in regulating gene expression networks [1,2]. The sRNAs are also valuable tools for functional genomics studies. Usually, the sRNAs silence gene expression by either degrading mRNA or repressing translation but, in a few cases, they also generate a population of secondary siRNAs. The ta-siRNAs are secondary siRNAs that are produced by a miRNA-targeted trigger that bridges the pathways of miRNA and siRNA regulation. ta-siRNAs can regulate plant development, metabolism, and responses to biotic and abiotic stresses, and thus have received more attention in the recent decade [3–5].

During the biogenesis of ta-siRNA, a single-stranded RNA is transcribed from a ta-siRNA-producing locus (TAS) and then cleaved by a phase-initiator (a miRNA or, in some cases, a ta-siRNA). Then, RNA-dependent RNA polymerase 6 (RDR6)-dependent conversion of the resulting fragments into double-stranded RNA and its subsequent cleavage by dicer-like 4 (DCL4) at every ~21 nucleotide (nt) relative to the phase-initiator cleavage site generates ~21-nt phased sRNAs. Some of the phased sRNAs become ta-siRNAs by binding argonaute (AGO) proteins to direct a *trans*-cleavage of targeted mRNAs [5–7]. Plant TASs can be classified into at least eight families, based on initiator-dependence, sequence similarity, and target gene identity. *TAS1* and *TAS2* are targets of miR173 and their ta-siRNAs can target the pentatricopeptide repeat genes [8]. *TAS3* is a target of miR390 and its ta-siRNA can target the auxin response factor gene family [8]. The initiator of *TAS4* is miR828, and the *TAS4* ta-siRNA can target the MYB transcription factor gene family [9]. *TAS5* is triggered by miR482 and its ta-siRNA can target the Bs4 resistance gene [10]. miR156 and miR529 initiate *TAS6*, which targets an mRNA that encodes a zinc finger protein [11]. miR828 initiates *TAS7*, which can target 13 genes, including genes that encode the leucine-rich receptor protein kinase-like protein and a calcium-transporting ATPase [12]. *At1g63130* is a pentatricopeptide repeat gene that was reported to be cleaved by *TAS2*-derived ta-siR2140 [13]. *TAS3* is flanked by two miR390 binding sites; one of which can be cleaved by the interaction of miR390 and AGO7, and another that is non-cleavable. Both binding sites are critical for the biogenesis of the *TAS3* ta-siRNAs. In contrast, other TASs have only a single miRNA binding site and are cleaved by the interaction of miRNA/ta-siRNA and AGO1. Recently, AGO2 has been reported to mediate *cis*-cleavage of *TAS1c*, although its slicer activity has not been demonstrated so far [14]. Taken together, it might be expected that many of the sequenced sRNAs could be mapped onto the phased positions set by the phase-initiator. Yet, while many of the sRNAs were successfully mapped, unexpectedly, many were mapped onto non-phased positions; that is, the intervals between the phased positions [12,13]. These sRNAs have been called "non-phased sRNAs" and how they are produced remains unclear.

The recent publication of the degradome library generated from cleaved mRNA fragments and sRNA libraries generated by high-throughput deep sequencing has enabled the study of all the cleavages that occur in a TAS transcript [15,16]. Here, we studied *cis*-cleavage of grapevine *TAS3* and found a cascading *cis*-cleavage, which produce sRNAs from the so-called non-phased positions and broaden the known scope of non-phased sRNA production.

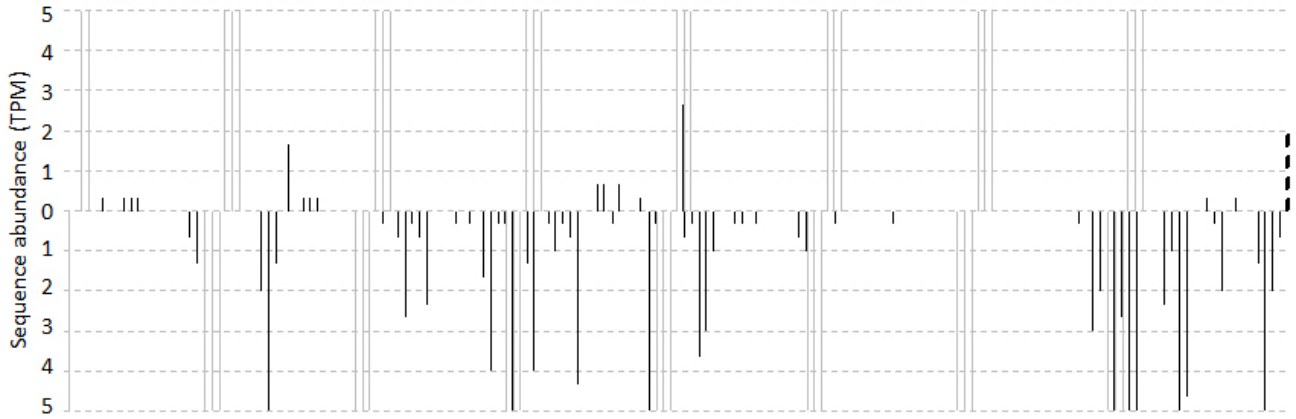
2. Results and Discussion

2.1. Overview of Small RNA Distribution on TAS3 *from* Vitis vinifera

Previously we reported that the *TAS3* from *Vitis vinifera* (vvi*TAS3*) can be targeted by vvi-miR390 to trigger ta-siRNA production in grapevine [12]. Here, to determine the distribution of sRNAs on the vvi*TAS3* transcript, a sRNA library from grapevine leaves (GEO: GSM458927) was used. To improve mapping confidence, only the sRNAs that mapped to a single site on the whole *V. vinifera* genome were used, because they could be attributed with certainty to a particular locus.

As a result of the mapping, we detected 131 unique sRNAs, representing 3969 reads, which matched perfectly to vvi*TAS3* (Figure 1). The 5' ends of the reads occupied 79 positions on the transcript. Only 14 (18%) of the positions were found to belong to phased positions set by vvi-miR390 when a 1-nt offset from the phased positions was allowed. After filtering out sRNAs that had TPM (tags per million) values of five or less, some unique sRNAs that mapped to non-phased positions remained (Figure 1). The percentage of phased sRNA positions increased from 18% to 40%. Together, these results showed that some sRNAs are really generated from non-phased positions, and might even be functional because of the relatively high levels at which they are often expressed.

Figure 1. Abundance distribution of sRNAs mapped to vvi*TAS3*. The number of reads with a 5' end at each position is plotted. Bars above the sequence represent sense reads; those below represent antisense reads. The cleavage site of vvi-miR390 is marked by a vertical dotted line. Vertical gray lines indicate the miRNA-set 21-nt phased cleavage, allowing a 1-nt offset. Regions with TPMs greater than five are not shown.



2.2. Computational Prediction and Validation of cis-*Cleavages*

Recent reports have shown that many functional siRNAs belong to a class of 21–22-nt 5'U/A sRNAs [14]. Therefore, we filtered out the potential siRNAs from the mapped sRNAs by limiting the length of the reads to 21 nt and the 5' end to U/A. As a result, we detected 35 sRNAs that were mapped to the antisense strand that passed the rule.

Additionally, it has been reported that many functional cleaved positions tend to be conserved through evolution, and they have been found to be highly conserved in alignments of genomic sequences from different species [8]. To identify conserved positions in the vvi*TAS3* transcript,

we compiled a dataset of *TAS3* sequences from eight dicotyledonous plants; namely, *V. vinifera*, *Ricinus communis*, *Populus trichocarpa*, *Arabidopsis thaliana*, *Malus domestica*, *Fragaria vesca*, *Prunus persica*, and *Glycine max*, and aligned them using ClustalX2 [17] with the default parameters (Figure 2). The multiple sequence alignment showed that there are no insertions or deletions among the *TAS3* sequences, except for a one-nucleotide deletion in vvi*TAS3* between position 116 and 117. The conserved positions were filtered by requiring each position to be conserved in at least six of the species, and to correspond to the 10th position of the 35 candidate *cis*-acting sRNAs. We found that 19 of the 35 sRNAs passed these rules.

Figure 2. Alignment of *TAS3* DNA sequences from eight dicotyledonous plant species. The numbers in the last row indicate the number of species with the same nucleotide as vvi*TAS3* at each of the marked positions. Only numbers over six are shown. vvi, *Vitis vinifera*; rco, *Ricinus communis*; ptr, *Populus trichocarpa*; ath, *Arabidopsis thaliana*; mdo, *Malus domestica*; fve, *Fragaria vesca*; ppe, *Prunus persica*; gma, *Glycine max*.

```
Position 1         10        20        30        40        50        60        70        80
vviTAS3  tcTTCTTGACCTTGTAAGACCTTTTCTTGACCTTGTAAGACCCTGTTGTGCGCATTTCTTCTTCGTTTTCCATCTTGTCATGAAC
rcoTAS3    TTCTTGACCTTGTAAGACCTTTTCTTGACCTTGTAAGACCCCATTTTGTGTTTATTTTTTCTGTTTTTTGTGTCCTTATAAAC
ptrTAS3    TTCTTGACCTTGTAAGACCTTTTCTTGACCTTGTAAGACCCCATTTTTTGTTCTTTGTTTCTATTTTTTGTATTATCATCAAC
athTAS3    TTCTTGACCTTGTAAGGCCTTTTCTTGACCTTGTAAGACCCCATCTCTTTCTAAACGTTTTATTATTTTCTCGTTTTACAGAT
mdoTAS3    TTCTTGACCTTGTAAGACCTTTTCTTGACCTTGTAAGACGTCACGCCGGGTCTGTTGTCTTTGTTTTGTGTTTTGGTTCAAGC
fveTAS3    TTCTTGACCTTGTAAGACCTTTTCTTGACCTTGTAAGATCTCACGCCGCGTCTGTTGTCTTTCTCTTCGCATTTCTAACACAC
ppeTAS3    TTCTTGACCTTGTAAGACCTTTTCTTGACCTTGTAAGATGTCACGCCGGGTCTGTTGTCTTCGTTTGTTGTTTTGGTCCAAGC
gmaTAS3    TTCTTGACCTTGTAAGACCTTTTCTTGACCTTGTAAGACCTCACACCCTATCTCTTCTCTTTGTTTTTGCTTTTGTGGAAGAC
           888888888888888878888888888888888888866         6  6 7  8 86  8687   7 77 6     67

Position      90        100       110       120       130       140       150       160       170
vviTAS3  TCTATACCAGCACCGACTGACATAGAGTTCG-GCCATCGTGTTCCCCGTTTCCGCCCAACTCATGTTCTCCTTCCTTGTCTATCCCtc
rcoTAS3  TCTTTATCAATATTTATTGATATAGAGTTTACAATTCTATTTCTCCATTTTCCGCCCAACTCATCTTCTCCTTCCTTGTCTATCCC
ptrTAS3  TTTACATCAAAATGTCTTGATGTAGAGTTTTGACTAATTTTTTCCCCGTTTCCGTCCAACTCATCTTCTCCTTCCTTGTCTATCCC
athTAS3  TCTATTCTATCTCTTCTCAATATAGAATAGATATCTATCTCTACCTCTAATTCGTTCGAGTCATTTTCTCCTACCTTGTCTATCCC
mdoTAS3  TTTGTGTCGGAATTCACTGATATAGAGTTTGGCCTCCGTTCCTCATCGTTACCGACCAACTCTTCTTCTCCTTCCTTGTCTATCCC
fveTAS3  CCTGCACCGGAATTCTCCGATGTAGAGTTTGGCCTCCGCTTCTCCCCGCTCCCGCCATACCCTTCTTCTCCTTCCTTGTCTATCCC
ppeTAS3  TCTGTGTGAGAATTATCTCATGTAGAGCTTGGCCAGCGTTCCTCCTTGTTGCCGCCCAACTCATCTTCTCCTTCCTTGTCTATCCC
gmaTAS3  CCTGTATCACTATCCACTGATATAGAGTTTGATCTCCTTCTTTCCCCGTTACCACCCAACTCATACTTTCCTTCCTTGTCTATCCC
         668 6  66  7     668  8888777    6     7  677 6667 787 776877868 787888878888888888888
```

Finally, using a parallel sequencing sRNA library that contained degradome tags from grapevine leaves [18], we validated the predicted *cis*-cleaved positions on the vvi*TAS3* transcript by requiring that the *cis*-cleaved positions overlapped with the 5' end of the RNA degradation fragment mapped onto the TAS. As a result, three of the positions, 63, 85, and 138, were validated. The corresponding *cis*-acting siRNAs (ca-siRNAs) were three 21-nt 5'U ca-siRNAs and one 22-nt 5'U ca-siRNA (Table 1). It has been reported that the size of the sRNAs and the 5'-terminal nucleotide are critical for the sorting of AGO. AGO1 binds 21-nt 5'U sRNAs, but in some cases, it also binds 22-nt 5'U sRNAs [14]. In Arabidopsis, the 22-nt 5'U 3'D10(-) from *TAS1c* has been reported to mediate its *cis*-cleavage by binding to AGO1, and the 21-nt 5'U 3'D6(-) from *TAS1c* has also been shown to mediate *TAS1c* cleavage by binding to AGO1; however, in this case, the cleaved site is not its original site [14]. In this study, we found three ca-siRNAs that were 21-nt 5'U sRNAs and one that was a 22-nt 5'U sRNA. These results implied that the four ca-siRNAs were all loaded to AGO1.

Table 1. Validated *cis*-cleaved positions on vvi*TAS3* transcript and their ca-siRNAs.

***cis*-cleaved position**	**ca-siRNA**	
	Name	**Sequence**
63	ca-siRM72	UGGAAAACGAAGAAGAAAUGC
85	ca-siRM94	UGGUAUAGAGUUCAUGACAAG
138	ca-siRM147	UGAGUUGGGCGGAAACGGGGA UGAGUUGGGCGGAAACGGGGAA

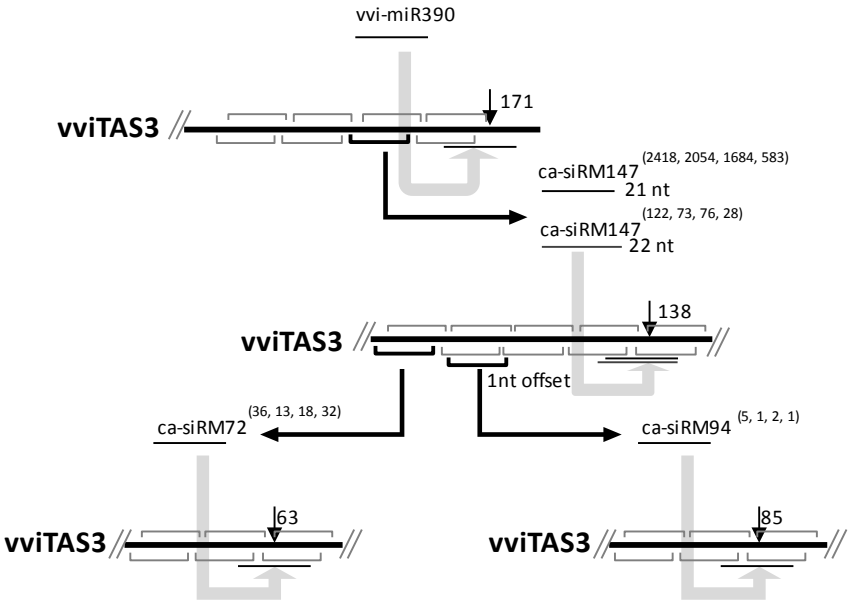
It has been demonstrated that two miR390 binding sites located on each side of *TAS3* are critical for *TAS3* ta-siRNAs biogenesis [8]. Therefore, we looked for ca-siRNA targeting sites on vvi*TAS3* and the flanking 300 bp upstream and downstream of the gene where the two miR390 binding sites were located. We found one targeted site on vvi*TAS3* for each of the ca-siRNAs. This finding suggested that *cis*-cleavage might use a different mechanism from the mechanism used by miR390 to initiate the cleavage of vvi*TAS3*.

We used the same method and criteria to examine the antisense strand that was not targeted by miR390. Surprisingly, no ca-siRNA targeting sites were detected on the antisense strand of vvi*TAS3*. The asymmetrical distribution between the targeted and non-targeted strand might imply that the non-targeted strand is readily in a double-stranded RNA form and is constantly processed by DCL4 and, therefore, protected from *cis*-cleavage. In addition, it might support the hypothetically biological function of *cis*-cleavage, *i.e.*, inactivating TAS transcription to feedback control ta-siRNA's production, as the sense strand was a template strand synthesizing antisense strand, so it would be more effective when the *cis*-cleavages preferred to occur in sense strand rather than in antisense strand.

2.3. Cascading cis-*Cleavages*

After identifying the *cis*-cleaved positions and their ca-siRNAs, we investigated how ca-siRNA production is triggered. It has been shown that ca-siRM147 can be triggered by miR390 [12], but for other ca-siRNAs the triggers remain unclear because they are out of the register set by miR390. To identify possible initiators, we investigated the distribution of *cis*-cleaved position and the locations of ca-siRNA 5' ends on the vvi*TAS3* transcript. We found that the 5' end of ca-siRM72 occurred precisely at the register set by ca-siRM147 in which the cleaved site of ca-siRM147 located on sense strand was 65 nt away from the 5' end of ca-siRM72 on antisense strand. The 5' end of ca-siRM94 was offset by 1 nt from the register set by ca-siRM147. Because the cleavage of phasing sRNAs often occurs within 1–2 nt of the phased position [8,13], we propose that the production of ca-siRM147, ca-siRM72, and ca-siRM94 is triggered by miR390, ca-siRM147, and ca-siRM147, respectively. The cascading *cis*-cleavage that we have proposed is shown schematically in Figure 3.

Figure 3. Cleavage cascades generated by *cis*-cleavages of siRNA on the sense strand of vvi*TAS3*. The four numbers in the brackets indicate the raw abundance of ca-siRNA in grapevine berries, leaves, inflorescences, and tendrils, respectively.



2.4. The Accumulated Levels of ca-siRNAs

When we analyzed the accumulated levels of the ca-siRNAs in the grapevine leaf, berry (GEO: GSM458930), inflorescence (GEO: GSM458929), and tendril (GEO: GSM458928) libraries, we found that the levels were in agreement with the cleavage cascades. The abundance of the ca-siRNAs that were located upstream of the cascade was always higher than the abundance of the ca-siRNAs that were downstream. For example, in the grapevine leaf library, the 21-nt ca-siRM147 located upstream had 2054 sequenced reads, while ca-siRM72, located downstream of ca-siRM147 cleavage, had only 13 sequenced reads. Moreover, the ca-siRNAs that occurred precisely at the register were always more abundant than those that occurred out of the register. For example, in the grapevine leaf library, ca-siRM72 had 13 sequenced reads, while ca-siRM94, which was shifted by 1 nt from the phased positions set by ca-siRM147, had only one sequenced read. Similar results were obtained for the other three tissues (Figure 3).

2.5. cis-*Cleavages Produced sRNAs in Increments of Approximately 21 nt*

To test whether or not *cis*-cleavages can also produce phased sRNAs in increments of approximately 21 nt, we searched for sRNAs with 5' ends that overlapped the predicted phased and non-phased positions by allowing an offset of 1 nt. As expected, each *cis*-cleavage had a set of corresponding phased sRNAs arranged in ~21-nt increments upstream and downstream of the cleavage position (Table 2).

Table 2. Phased patterns set by different *cis*-cleavages.

cis-cleaved position	Upstream of *cis*-cleaved position			Downstream of *cis*-cleaved position		
	Number of phased sRNAs	Number of non-phased sRNAs	*p* value *	Number of phased sRNAs	Number of non-phased sRNAs	*p*-value
63	2	12	1.11×10^{-1}	3	47	4.92×10^{-1}
85	5	27	**7.54×10^{-3}**	4	30	1.00×10^{-1}
138	8	41	**4.71×10^{-4}**	3	16	7.43×10^{-2}

* *p*-values less than 0.01 are in bold font.

To test whether or not the phased patterns produced from *cis*-cleavages were statistically significant, we developed an improved equation (see Experimental Section) by modifying previous algorithms [12,13,19] to evaluate the phasing pattern set by *cis*-cleavage. First, the new equation is not constrained by the previous 231-bp length requirement [12,13,19], but requires only a multiple of 21 nt, which provides a more accurate TAS evaluation, especially for TASs that are longer or shorter than 231 bp. Second, our equation uses a variable *s* to reflect the maximum offset from a phase position [12,19], making the evaluation more flexible. This equation could be applied to TAS identification in the future. Using our improved algorithm [12,13,19], we determined that two *cis*-cleavages had significant phasing patterns (*p*-value < 0.01). When the number of sRNAs located in the phased positions set by ca-siRNAs was counted, we found that 54 unique sRNA located in non-phased positions set by miRNA390 were included in the phasing patterns. These results suggested that some common processes might be used for both miRNA-mediated ta-siRNA production and the *cis*-cleavage of siRNA.

2.6. The Conservation of ca-siRNAs and Cascading cis-*Cleavages*

To examine the conservation of *cis*-cleavages further, we first looked for the presence of ca-siRNAs in the sRNA datasets of *V. vinifera*, *A. thaliana*, *M. domestica*, and *P. persica* downloaded from the Gene Expression Omnibus (GEO) or the plant MPSS databases. We found that although the accumulation levels of the four ca-siRNAs varied in the different species, they were expressed in all four species, except for 22-nt ca-siRNA147, which was not detected in *P. persica* (peach) (Table 3).

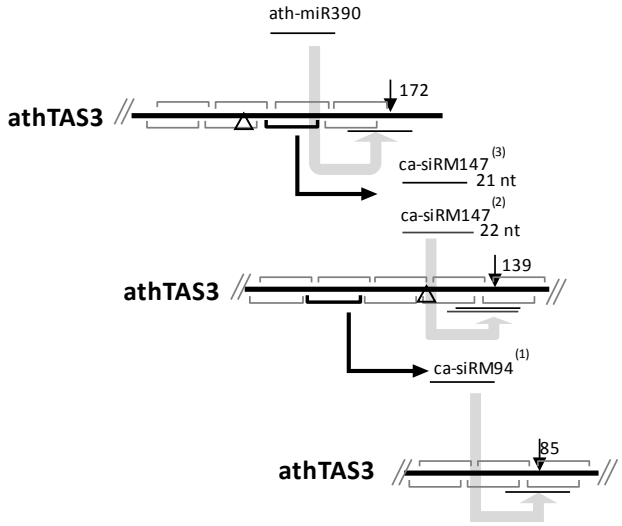
Table 3. Conservation of ca-siRNAs in the sRNA datasets of four species.

ca-siRNAs		Average of normalized abundance (TPM)			
Name	Length	Grapevine	Apple	Peach	Arabidopsis
ca-siRM72	21 nt	6	2	1	1
ca-siRM94	21 nt	1	3	14	1
ca-siRM147	21 nt	419	32	9	2
	22 nt	18	1	0	2

We then evaluated the corresponding *cis*-cleavages based on the Col7d samples (GEO: GSE20197) and the degradome library (GEO: GSM280227) from Arabidopsis, and found that, two *cis*-cleavages occurred in positions 85 and 139 (equivalent to position 138 in vvi*TAS3* because of the nucleotide deletion in the vvi*TAS3* sequence) were also validated on ath*TAS3* (Figure 4). The accumulated levels of ca-siRNAs in Col7d sample were also in agreement with the cleavage cascades. In which, the 21-nt and

22-nt ca-siRM147 located upstream had, respectively, three and two sequenced reads, while ca-siRM94, located downstream of ca-siRM147 cleavage, had one sequenced reads. In a previous study [5], it was suggested that *cis*-cleavage occurred on position 139 in ath*TAS3*, although the ca-siRNA and functional sRNAs were not found. Here, we detected the ca-siRNA and a secondary ca-siRNA product (ca-siRNA94), which we believe provides enough evidence to establish the *cis*-cleavage on *TAS3*. The findings reported here for ath*TAS3* indicate that cascading *cis*-cleavage is conserved.

Figure 4. Cleavage cascades generated by *cis*-cleavages of siRNA on the sense strand of ath*TAS3*. The triangle indicates the position of the nucleotide deletion in vviTAS3. The number in the bracket indicates the raw abundance of ca-siRNA in Col7d (GEO: GSE20197).



3. Experimental Section

3.1. Sources of sRNA Libraries

In this study, we used four deep sequencing sRNA datasets; namely, two degradome library and two sRNA libraries. All the datasets were downloaded from the Gene Expression Omnibus (GEO). The GEO accession numbers for these libraries are given in the Results section.

3.2. Evaluation of Phasing Patterns Set by cis-*Cleavage*

Once a *cis*-cleaved position was determined, the numbers of phased and non-phased positions were counted upstream and downstream of the cleavage sites respectively. Phased positions refer to positions arranged in 21-nt increments relative to the cleavage position as well as to positions shifted by *s* nt relative to the positions of 21-nt increments. Non-phased positions are all the other positions. The *p*-value of each detected phasing pattern was calculated based on a random hyper-geometric distribution using an improved equation based on previously used algorithms [12,13,19,20]:

$$P(k_1) = \sum_{x=k_1}^{\min\left(k_1+k_2,\frac{2L}{21}-1\right)} \frac{\binom{(2L-1)-(\frac{2L}{21}-1)\times(2s-1)}{k_2}\binom{\frac{2L}{21}-1}{k_1}}{\binom{(2L-1)-(\frac{2L}{21}-1)\times 2s}{k_1+k_2}} \quad (1)$$

Where L is the length of the detected pattern and is a multiple of 21, K_1 is the number of phased positions having sRNA hits, K_2 is the number of non-phased positions having sRNA hits, and s is the maximum allowed offset from the phase position.

3.3. Expressional Conservation of ca-siRNAs

The expressional conservation of the grapevine ca-siRNAs was investigated by performing a search against 84 sRNA libraries from grapevine, apple, Arabidopsis, and peach. The sRNA libraries of grapevine (GEO: GSE18405) and apple (GEO: GSE36065) were downloaded from the GEO and the sRNA libraries of Arabidopsis and peach were used from the MPSS databases [21]. The normalized abundance is the raw expression value divided by the total number of signatures and multiplied by 1,000,000.

4. Conclusions

In this work, we reexamined the distribution of sRNAs on vvi*TAS3* using a stringent threshold that used only the sRNAs that mapped to a single site on the whole *V. vinifera* genome and that had normalized abundant values of one or more TPM. Our results showed that the non-phased positions were indeed located by some of the uniquely mapped sRNAs. We identified three *cis*-cleavages that directed by four ca-siRNAs at positions 63, 85, and 138 on vvi*TAS3* by combining computational predictions and validation. We found that three *cis*-cleavages, together with their ca-siRNAs, formed a cascading *cis*-cleavage. The accumulated levels of four ca-siRNAs in the berry, leaf, inflorescence, and tendril libraries of *V. vinifera* also agreed with the cascade. A comparative analysis showed that the expression levels of the four ca-siRNAs were conserved among grapevine, apple, peach, and Arabidopsis, and part of the *cis*-cascade was also identified in Arabidopsis. We also found that sRNAs were located at the phased positions set by ca-siRNA. These results broaden the known scope of non-phased sRNA production. We also developed an improved equation by modifying previous algorithms to evaluate the phasing pattern set by *cis*-cleavage. It could be applied to TAS identification in the future.

Acknowledgments

This work was supported by the National Science Foundation of China (Grant No. 31171273) and the Qing Lan Project of Jiangsu Province, China.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Axtell, M.J. Classification and comparison of small RNAs from plants. *Annu. Rev. Plant Biol.* **2013**, *64*, 137–159.
2. Chen, X. Small RNAs and their roles in plant development. *Annu. Rev. Cell Dev. Biol.* **2009**, *25*, 21–44.

3. Peragine, A.; Yoshikawa, M.; Wu, G.; Albrecht, H.L.; Poethig, R.S. SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in *Arabidopsis*. *Genes Dev.* **2004**, *18*, 2368–2379.
4. Vazquez, F.; Vaucheret, H.; Rajagopalan, R.; Lepers, C.; Gasciolli, V.; Mallory, A.C.; Hilbert, J.L.; Bartel, D.P.; Crete, P. Endogenous trans-acting siRNAs regulate the accumulation of *Arabidopsis* mRNAs. *Mol. Cell* **2004**, *16*, 69–79.
5. Allen, E.; Xie, Z.; Gustafson, A.M.; Carrington, J.C. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* **2005**, *121*, 207–221.
6. Gasciolli, V.; Mallory, A.C.; Bartel, D.P.; Vaucheret, H. Partially redundant functions of *Arabidopsis* DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr. Biol.* **2005**, *15*, 1494–1500.
7. Yoshikawa, M.; Peragine, A.; Park, M.Y.; Poethig, R.S. A pathway for the biogenesis of trans-acting siRNAs in *Arabidopsis*. *Genes Dev.* **2005**, *19*, 2164–2175.
8. Axtell, M.J.; Jan, C.; Rajagopalan, R.; Bartel, D.P. A two-hit trigger for siRNA biogenesis in plants. *Cell* **2006**, *127*, 565–577.
9. Rajagopalan, R.; Vaucheret, H.; Trejo, J.; Bartel, D.P. A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev.* **2006**, *20*, 3407–3425.
10. Li, F.; Orban, R.; Baker, B. SoMART: A web server for plant miRNA, tasiRNA and target gene analysis. *Plant J.* **2012**, *70*, 891–901.
11. Arif, M.A.; Fattash, I.; Ma, Z.; Cho, S.H.; Beike, A.K.; Reski, R.; Axtell, M.J.; Frank, W. DICER-LIKE3 activity in *Physcomitrella patens* DICER-LIKE4 mutants causes severe developmental dysfunction and sterility. *Mol. Plant* **2012**, *5*, 1281–1294.
12. Zhang, C.; Li, G.; Wang, J.; Fang, J. Identification of trans-acting siRNAs and their regulatory cascades in grapevine. *Bioinformatics* **2012**, *28*, 2561–2568.
13. Chen, H.M.; Li, Y.H.; Wu, S.H. Bioinformatic prediction and experimental validation of a microRNA-directed tandem trans-acting siRNA cascade in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3318–3323.
14. Rajeswaran, R.; Aregger, M.; Zvereva, A.S.; Borah, B.K.; Gubaeva, E.G.; Pooggin, M.M. Sequencing of RDR6-dependent double-stranded RNAs reveals novel features of plant siRNA biogenesis. *Nucleic Acids Res.* **2012**, *40*, 6241–6254.
15. Howell, M.D.; Fahlgren, N.; Chapman, E.J.; Cumbie, J.S.; Sullivan, C.M.; Givan, S.A.; Kasschau, K.D.; Carrington, J.C. Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in Arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting. *Plant Cell* **2007**, *19*, 926–942.
16. German, M.A.; Pillay, M.; Jeong, D.H.; Hetawal, A.; Luo, S.; Janardhanan, P.; Kannan, V.; Rymarquis, L.A.; Nobuta, K.; German, R.; *et al*. Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat. Biotechnol.* **2008**, *26*, 941–946.
17. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; *et al*. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948.

18. Pantaleo, V.; Szittya, G.; Moxon, S.; Miozzi, L.; Moulton, V.; Dalmay, T.; Burgyan, J. Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. *Plant J. Cell Mol. Biol.* **2010**, *62*, 960–976.
19. Dai, X.; Zhao, P.X. pssRNAMiner: A plant short small RNA regulatory cascade analysis server. *Nucleic Acids Res.* **2008**, *36*, W114–W118.
20. Zhang, C.; Wang, J.; Hua, X.; Fang, J.; Zhu, H.; Gao, X. A mutation degree model for the identification of transcriptional regulatory elements. *BMC Bioinforma.* **2011**, *12*, doi:10.1186/1471-2105-12-262.
21. Nakano, M.; Nobuta, K.; Vemaraju, K.; Tej, S.S.; Skogen, J.W.; Meyers, B.C. Plant MPSS databases: Signature-based transcriptional resources for analyses of mRNA and small RNA. *Nucleic Acids Res.* **2006**, 34, D731–D735.

Reprinted from *IJMS*. Cite as: Kang, S.-M.; Choi, J.-W.; Hong, S.-H.; Lee, H.-J. Up-Regulation of microRNA* Strands by Their Target Transcripts. *Int. J. Mol. Sci.* **2013**, *14*, 13231-13240.

Communication

Up-Regulation of microRNA* Strands by Their Target Transcripts

Sung-Min Kang [1], Ji-Woong Choi [1], Su-Hyung Hong [1] and Heon-Jin Lee [1,2,*]

[1] Department of Oral Microbiology, School of Dentistry, Kyungpook National University, Daegu 700-412, Korea; E-Mails: dkdkdk43@naver.com (S.-M.K.); cjwboyboy@naver.com (J.-W.C.); hongsu@knu.ac.kr (S.-H.H.)

[2] Brain Science and Engineering Institute, Kyungpook National University, Daegu 700-412, Korea

* Author to whom correspondence should be addressed; E-Mail: heonlee@knu.ac.kr; Tel.: +82-53-660-6832; Fax: +82-53-425-6025.

Received: 27 April 2013; in revised form: 29 May 2013 / Accepted: 17 June 2013 / Published: 26 June 2013

Abstract: During microRNA (miRNA) biogenesis, one strand of a 21–23 nucleotide RNA duplex is preferentially selected for entry into an RNA-induced silencing complex (RISC). The other strand, known as the miRNA* species, is typically thought to be degraded. Previous studies have provided miRNA* selection models, but it remains unclear how the dominance of one arm arises during the biogenesis of miRNA. Using miRNA sponge-like methods, we cloned four tandem target sequences (artificial target) of miR-7b* and then measured miR-7b* expression levels after transfection of the artificial target. miR-7b* levels were found to significantly increase after transfection of the artificial target. We postulate that the abundance of target transcripts drives miRNA arm selection.

Keywords: microRNA (miRNA); pre-miRNA; RISC; Drosha; Dicer; Argonaute (Ago)

1. Introduction

MicroRNAs (miRNAs) are noncoding 21–23 nucleotide (nt) strands and constitute an evolutionarily conserved class of pleiotropically acting small RNAs. miRNAs usually control posttranscriptional processes, such as, sequence-specific interactions with 3' untranslated regions (UTRs) of cognate mRNA targets in animals [1]. Nucleotides at positions 2–8 are considered to be important for pairing with target messenger RNAs and are referred to as “seed” sequences [2,3].

A miRNA gene is first transcribed into a primary miRNA (pri-miRNA), then processed into a ~70 nt hairpin precursor miRNA (pre-miRNA) by the RNase III enzyme Drosha and double-stranded RNA-binding domain protein, DGCR8 [4]. Pre-miRNA is then cleaved to generate the ~22 nt miRNA:miRNA* duplex by another RNase III enzyme, Dicer [5]. One strand of the duplex is loaded onto Argonaute (Ago) protein to produce RNA-induced silencing complex (RISC). Furthermore, it has been suggested that, whereas all Ago proteins participate in the stabilization of mature miRNA, only Ago2 (which has endonuclease activity) cleaves the miRNA* strand and activates RISC [6–9]. This suggestion led to the observation that Ago1 facilitated RISC-mediated translational repression and Ago2-RISC led to target mRNA cleavage [8,10]. However, Ago proteins have also been demonstrated to have dual roles, for example, Ago proteins increased the abundance of mature miRNAs, and decreased miRNA expression was observed in a cell line from an Ago2 knockout mouse [11].

The nomenclature originated because one arm, the miRNA, of the RNA duplex preferentially accumulates and the opposite arm, miRNA*, degrades. Another nomenclature often used is miR-3p/miR-5p, which refers to the direction of the mature miRNA strand. 3p and 5p miRNAs usually exhibit partial complementary overlap and have different target genes, despite being produced from the same transcript [12]. However, it remains unclear how dominance of one arm arises during the biogenesis of miRNA. Previous models suggest that the choice of the dominant miRNA arm is based on two mechanisms, that is, on the thermodynamic stability and structural properties of the processed duplex [13,14], or on energy-independent protein-mediated selection by Ago2, an endonuclease that cleaves complementary siRNA strands to facilitate RISC loading of the siRNA strand [6].

However, in a recent study, it was suggested the hairpin arm that makes dominant miRNA differs in different tissues, at different times of development, and between species [15]. In human gastric cancer, miRNA hairpin arm (3p or 5p) selection exhibits different tissue expression preferences in healthy and tumor tissues [16]. Furthermore, some miRNA precursors are processed to produce significant amounts of mature miRNAs from both arms and both miRNAs might regulate target transcripts [17]. These findings suggest the existence of another mechanism for controlling the selection of mature miRNAs.

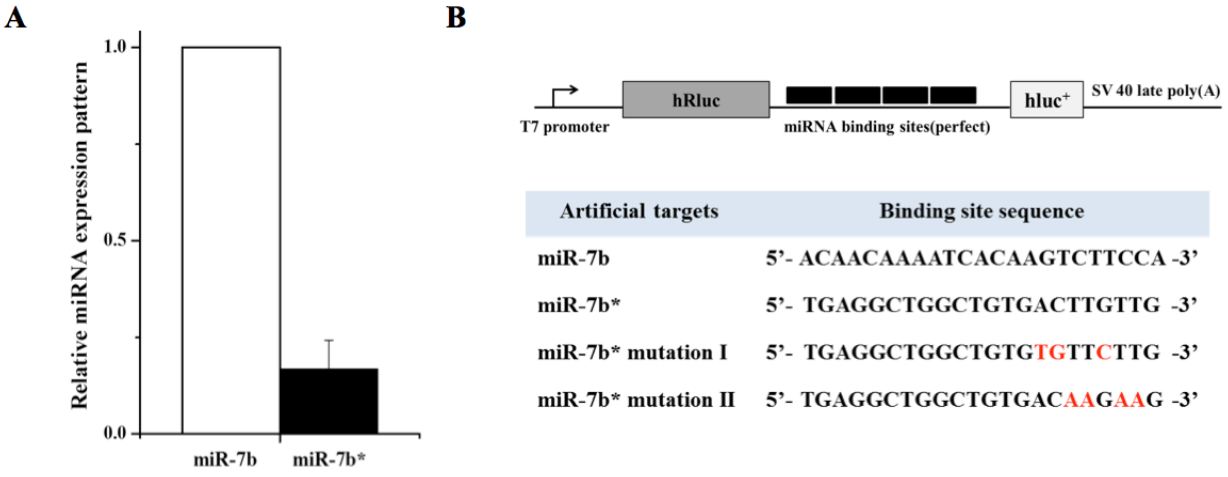
We expressed an artificial target of miR-7b* that normally presents less than its mature miRNA, miR-7b in order to investigate the effect of the target mRNA on miRNA*. The expression of artificial targets can force the accumulation of miR-7b* rather than miR-7b, which suggests that target abundance might be a critical prerequisite of miRNA* strand stabilization.

2. Results and Discussion

The importance of the influences of target mRNA abundance and turnover rates on miRNA activity have been discussed [18,19]. Recently, Chatterjee *et al.* suggested the term “target-mediated miRNA protection (TMMP)”, and showed that target mRNAs in *C. elegans* can protect their cognate miRNAs from degradation *in vivo* [20]. However, little is known about the decay of miRNA*, the other arm of the same hairpin precursor. We hypothesized that target mRNAs of high abundance may drive miRNA arm selection, and in a previous study using miRNA sponge-like methods [21], we cloned multiple target sequences of miR-7b and miR-7b* (miRNA artificial targets; Figure 1B). It has been reported that hyperosmolar stimulation induces miR-7b in the hypothalamus and that the neuronal marker Fos expression is inhibited by miR-7b [22]. Usually, miR-7b is dominant (approximately eight fold higher than miR-7b*) in AtT-20 mouse pituitary cells (Figure 1A). However, in the present study, both

qRT-PCR and Northern blotting clearly showed dramatic elevated expression of endogenous miR-7b* by the miR-7b* artificial target. The seed-sequence mutated artificial targets (miR-7b* mutation I and II) reduced the miR-7b* up-regulating effect (Figure 2A,B), suggesting that the up-regulation of miR-7b* occurred in a sequence specific manner (see Figure 1B for artificial and seed mutated targets). Furthermore, luciferase assays showed that miR-7b* mimic oligonucleotide strongly suppressed luciferase activity by binding to its artificial target with a perfect complementary match: this effect was reduced by the seed mutated artificial targets (Figure 1C). We observed the same effects for miR-338-3p (dominant or guide strand) and for miR-338-5p (miR-338* or passenger strand) after artificial targets transfection (Figure 3). Interestingly, when the artificial targets of the dominant strands (miR-7b and miR-338-3p) were transfected, the elevation effect was not as great as that of the artificial target of the non-dominant strands (miR-7b* and miR-338-5p) (Figures 2A,B and 3C). This suggests there might be a certain threshold for this differential regulation or some unknown mechanism that overrides miRNA arm selection.

Figure 1. The expression patterns of miR-7b and miR-7b*. (**A**) Basal levels of miR-7b and miR-7b* expression in AtT-20 cells were analyzed by qRT-PCR; (**B**) Design of miRNA artificial targets. *Renilla* (hR*luc*) containing miRNA artificial targets were constructed by inserting multiple miRNA binding sites into the 3' UTR of a hR*luc* reporter gene driven by a T7 promoter. The figure shows the nucleotide sequence of the miRNA artificial targets: the red letters are a mutated sequence; (**C**) Luciferase assay using the miRNA artificial target reporter constructs. COS-7 cells were co-transfected with the miR-7b* mimic oligonucleiotide and miRNA artificial target (1. miR-7b, 2. miR-7b*, 3. miR-7b* mutation I, 4. miR-7b* mutation II, 5. miR-338-3p). *Renilla* luciferase activity was normalized against *firefly* luciferase activity and fold changes are compared to miR-338-3p. The results shown are the means of three independent transfections (error bars indicate standard deviations).



Artificial targets	Binding site sequence
miR-7b	5'- ACAACAAAATCACAAGTCTTCCA -3'
miR-7b*	5'- TGAGGCTGGCTGTGACTTGTTG -3'
miR-7b* mutation I	5'- TGAGGCTGGCTGTGTGTTCTTG -3'
miR-7b* mutation II	5'- TGAGGCTGGCTGTGACAAGAAG -3'

Figure 1. *Cont.*

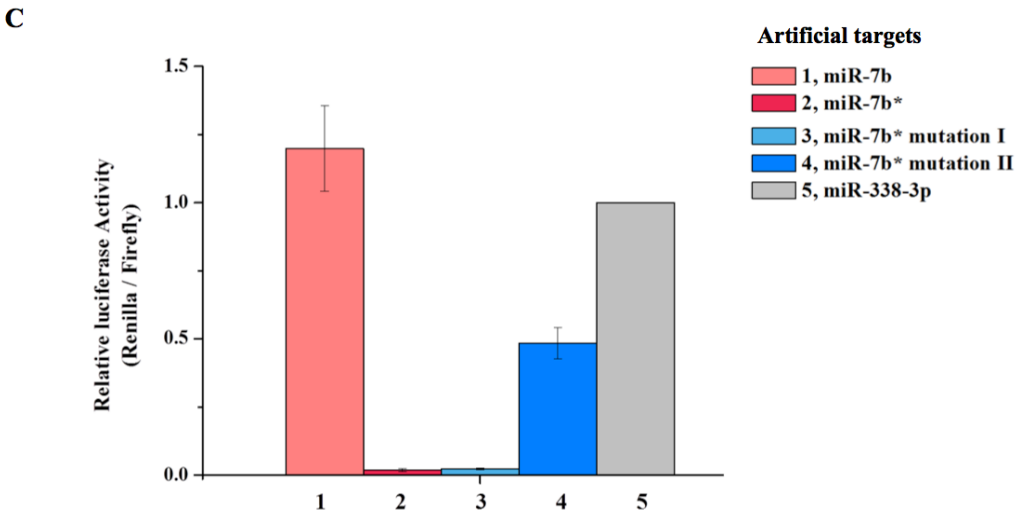


Figure 2. Changed expression levels of miR-7b and miR-7b* after transfection with multi-transcript artificial targets. (**A**) 24 h after transfection of miRNA artificial targets (1 and 1'. null, 2 and 2'. empty vector, 3 and 3'. miR-7b*, 4 and 4'. miR-7b* mutation I, 5 and 5'. miR-7b* mutation II), AtT-20 total RNA was extracted and analyzed by northern blotting using probes specific for miR-7b and miR-7b*, respectively. Lane 3' shows that miR-7b* expression was significantly increased following transfection of the miR-7b* artificial target. In contrast, miR-7b expression was similar in non-transfected and miR-7b* artificial targets transfected samples; (**B**) Twenty-four hours after various artificial targets (1 and 1'. null, 2 and 2'. empty vector, 3 and 3'. miR-7b, 4 and 4'. miR-7b*, 5 and 5'. miR-7b* mutation I, 6 and 6'. miR-7b* mutation II) were transfected, miR-7b (left) and miR-7b* (right) expression levels were analyzed by qRT-PCR. miRNA expression levels were normalize to U6. Fold changes are expressed *versus* empty target of each (2 or 2'); (**C**) Twenty-four hours after various artificial targets (1 and 1'. empty vector, 2 and 2'. miR-7b*, 3 and 3'. miR-7b* mutation I, 4 and 4'. miR-7b* mutation II) were transfected, AtT-20 lysates were immunoprecipitated (IP) using anti-Ago2 antibody. IPs were analyzed by immunoblotting with the same anti-Ago2 antibody (upper panel). Co-immunoprecipitated RNA was extracted and analyzed by qRT-PCR (lower panel). Fold changes are expressed *versus* the empty target for each (1 or 1'). Error bars in graphs indicate standard deviations; *p*-values are also indicated.

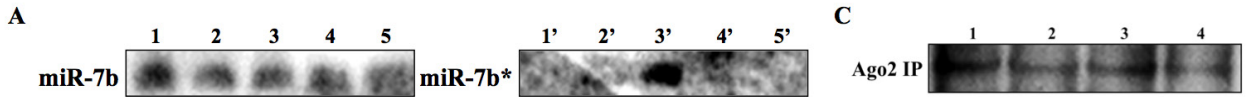


Figure 2. *Cont.*

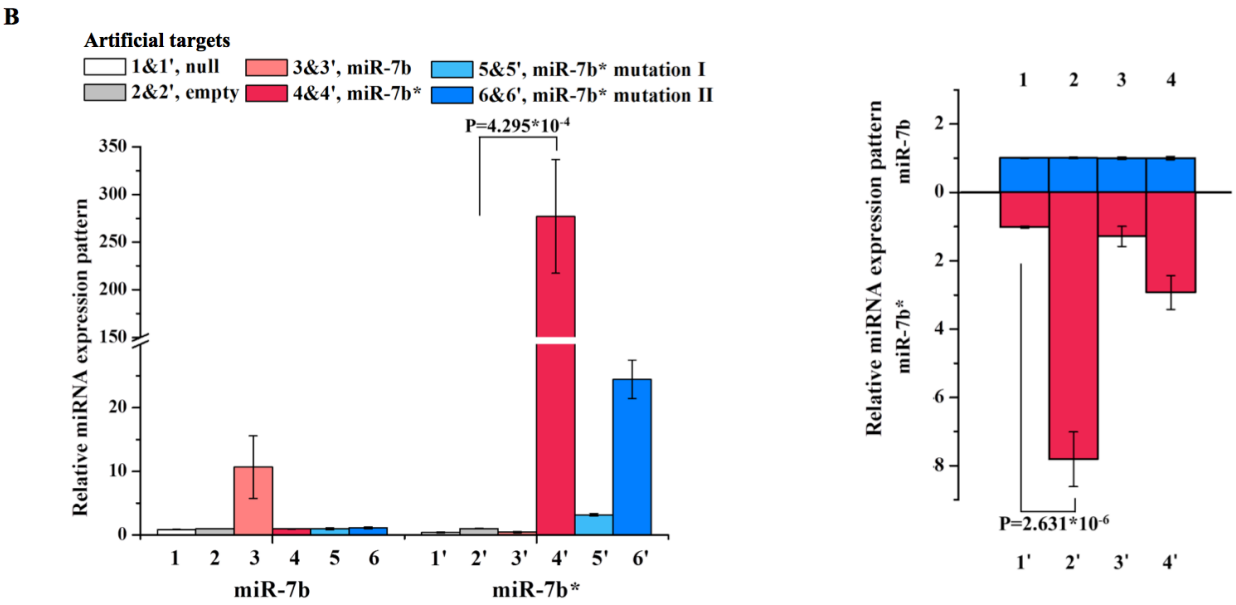


Figure 3. miR-338-3p and miR-338-5p expressions. (**A**) miR-338-3p and miR-338-5p expression were analyzed by qRT-PCR of the total RNA extracted from AtT-20 cells. miR-338-3p levels were about twice as high as miR-338-5p levels in AtT-20 cells; (**B**) Schematic representation of miRNA artificial targets. Construct map of the miR-338-3p and miR-338-5p artificial targets with miR-338-3p and miR-338-5p artificial target sequences below; (**C**) Both miR-338-3p and miR-338-5p expression was up-regulated by each of artificial targets. However, the change of miR-338-3p was not as great as miR-338-5p (miR-338*) by their artificial target. The expression levels of miR-338-5p were significantly elevated by miR-338-5p artificial target transfection but not by transfection of the miR-338-3p artificial target or empty vector (1 and 1'. null, 2 and 2'. empty vector, 3 and 3'. miR-338-3p, 4 and 4'. miR-338-5p). Fold changes are expressed *versus* the empty target for each (2 or 2'). Error bars in graphs indicate standard deviations; *p*-values are also indicated.

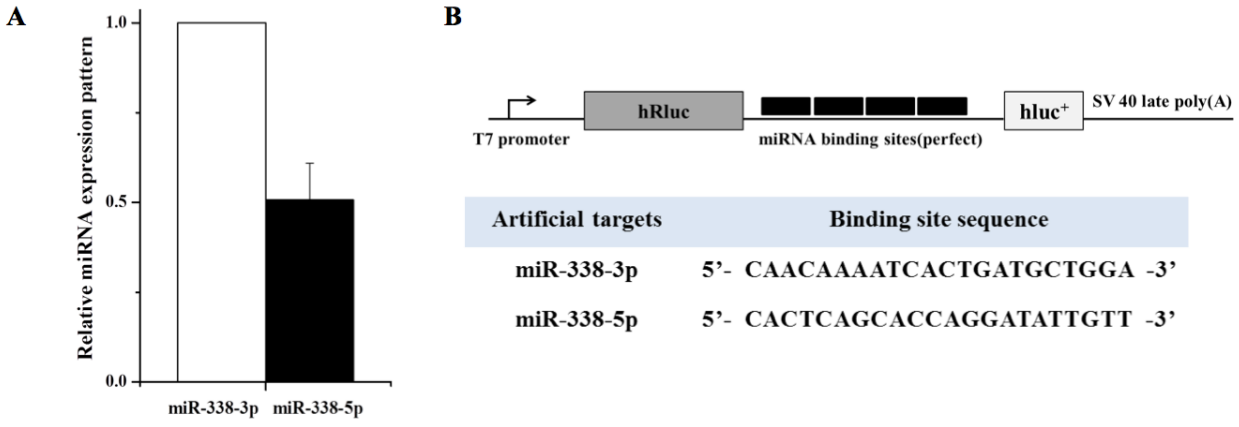


Figure 3. *Cont.*

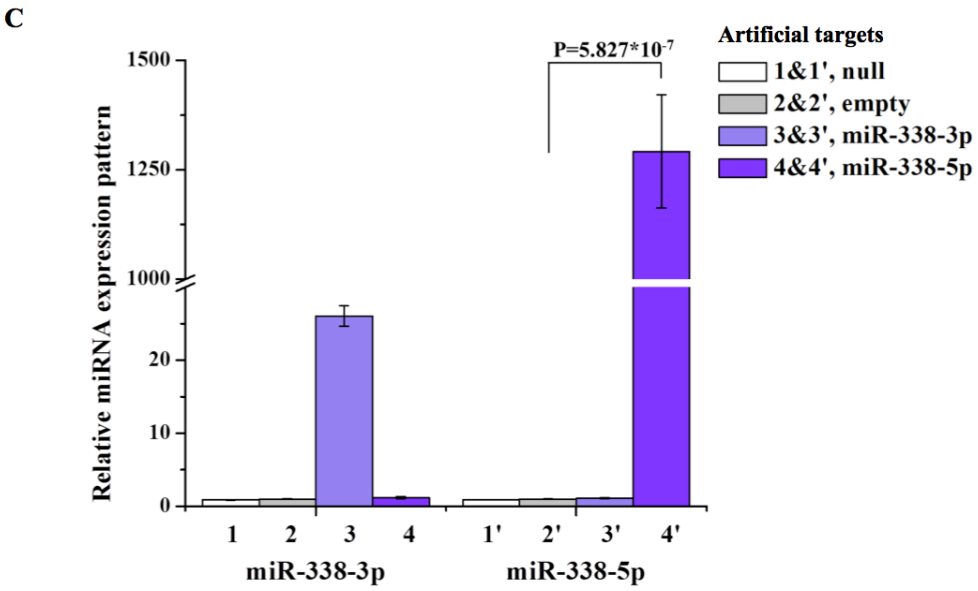
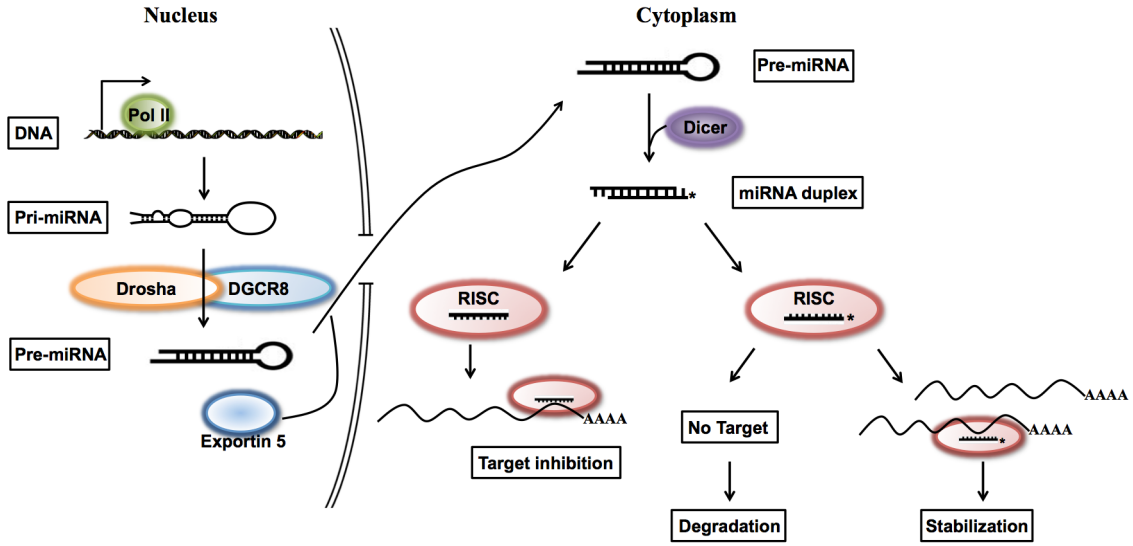


Figure 4. Schematic diagram of miRNA arm selection. Pri-miRNA is transcribed by RNA polymerase II and processed into pre-miRNA by Drosha/DGCR8 in the nucleus. Then pre-miRNA is processed into a duplex form of mRNA that is unwound during RISC assembly in the cytoplasm. Selection of the miRNA* sequence is determined by target transcript abundances. When miRNA* target transcripts are sufficiently abundant the miRNA* sequence is stabilized and acts on target transcripts in the same manner as miRNA.



Since Ago proteins are key players in small RNA-mediated RNA silencing pathways [23], and Ago2 mediates RNA cleavage by harboring RNA catalytic activity in human and mouse [24,25], we performed an Ago2-immunoprecipitation assay to check whether up-regulated miR-7b could interact with Ago2. We found that up-regulated miR-7b* interacted with Ago2 (Figure 2C), which suggests that miR-7b* may have the same inhibitory function as mature miRNAs, similar to the regulatory activity of miRNA* function without degradation [17]. We tried to examine the effect of TMMP on miR-7b* by using endogenous miR-7b* target transcripts, but failed to identify targets of miR-7b

experimentally due to the difficulty of predicting miRNA targets. Based on our results, we postulate mature miRNA arm selection is influenced by the abundance of miRNA target transcripts, and that selection may occur as duplex miRNAs, incorporated into the RISC complex (Figure 4), unwind from RISC assembly [26].

3. Experimental Section

3.1. Cell Culture

COS-7 cells (a monkey kidney fibroblast cell line) were obtained from the Korea Cell Line Bank (Seoul). COS-7 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM/High Glucose) (Hyclone, Logan, UT, USA) containing 10% fetal bovine serum (FBS) (Hyclone, Logan, UT, USA) and 1× antibiotic-antimycotic (GIBCO, Grand Island, NY, USA), and then incubated at 37 °C in a 5% CO_2 atmosphere. AtT-20 cells (a mouse pituitary tumor cell line) were obtained from the American Type Culture Collection (ATCC, CCL-89), and were grown in Kaighn's Modification of Ham's F-12 (F-12K) medium (Hyclone) supplemented with 15% horse serum, 2.5% FBS and 1× antibiotic-antimycotic. These cells were used in the non-luciferase assays.

3.2. Construction of Artificial miRNA Target

We annealed, ligated, and cloned oligonucleotides for miRNA binding sites with four tandem target sequences (Figure 1B) into psiCEHCK2 vector (Promega, Madison, WI, USA) downstream of the *Renilla* luciferase (hR*luc*) coding sequence.

3.3. Luciferase Assays

COS-7 cells were plated the day before transfection and transfected in triplicate with Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) and 800 ng of various artificial target plasmids and 25 nM of miR-7b* mimic oligonucleotide (Bioneer, Daejeon, Korea). All assays were performed 24 h after transfection using the dual luciferase assay (Promega, Madison, WI, USA), according to the manufacturer's protocol. All experiments were performed in triplicate.

3.4. Isolation of miRNA and Quantitative Real-Time PCR (qRT-PCR) Analysis

Total RNAs were extracted from the artificial target transfected AtT-20 cells using QIAzol (Qiagen, Valencia, CA, USA) using a modification of the manufacturer's instructions, and then treated with DNaseI (Ambion, Foster City, CA, USA). qRT-PCR of miRNAs was conducted on an ABI 7500 real-time PCR system using TaqMan Universal PCR Master Mix, miRNA Expression Assay primer, and probe sets (Applied Biosystems, Foster City, CA, USA). U6 RNA (a small nuclear RNA) was used as an internal cDNA loading control. Threshold cycle times (*C*t) were obtained and relative gene expressions were calculated using the comparative cycle time method.

3.5. Immunoprecipitation

For the immunoprecipitation of endogenous Ago2, AtT-20 cells were grown on 10 cm dishes and harvested at 24 h after miRNA transfections. Cells were then incubated with lysis buffer for 20 min on ice, homogenized, and centrifuged at 12,000 rpm for 20 min at 4 °C. Supernatants were incubated with anti-Ago2 antibody (Sigma Aldrich, St. Louis, MO, USA) with constant rotation for one day at 4 °C. Then, 20 μL of protein G Sepharose® beads (Sigma Aldrich, St. Louis, MO, USA) were added and incubated with rotation for 4 h at 4 °C. Beads were then washed three times with lysis buffer.

3.6. Isolation of Ago2-Associated miRNA

To measure amounts of RISC-associated miRNAs, cell lysates were prepared from AtT-20 cells after the transfection of miRNA artificial targets. Ago2-miRNA complex was immunoprecipitated from lysates using anti-Ago2 antibody and total RNA was isolated from immunoprecipitates using QIAzol reagent (Qiagen, Valencia, CA, USA). The miRNAs levels were measured by quantitative RT-PCR and normalized against U6 levels in cell lysates.

3.7. Northern Blot Analysis for miRNA

Ten μg aliquots of AtT-20 total RNA isolated from AtT-20 cells using QIAzol (Qiagen, Valencia, CA, USA), according to the manufacturer's instructions, was separated on 15% TBE-urea gels (Invitrogen, Carlsbad, CA, USA) and electro-transferred to Nylon+ membranes (Invitrogen, Carlsbad, CA, USA). Hybridizations were carried out in North2South® hybridization buffer (Invitrogen) at 37 °C using miR-7b* probe (Bioneer, Daejeon, Korea), or at 42 °C using miR-7b LNA (Locked Nucleic Acid) probe (Exiqon, Vedbaek, Denmark).

3.8. Statistical Analysis

All data are presented as means ± standard deviations (SD). Significant variation analysis was conducted to calculate the parametric two-tailed non-paired *t*-test. All analyses were performed using Origin 8.0 (OriginLab, Northampton, MA, USA), and *p*-values of ≤ 0.05 were considered statistically significant.

4. Conclusions

Despite extensive study of miRNA, it remains largely unclear how one miRNA arm becomes less dominant (often referred as miRNA*) during the miRNA maturation process. In this study, we introduced an artificial target of miR-7b* in order to check miR-7b* stability. Transfection of the miR-7b* artificial target led to a dramatic up-regulation of miR-7b*, but did not have much effect on miR-7b, the dominant sequence of miR-7b hairpin precursor. A similar phenomenon was observed in miR-338-3p and miR-338-5p (miR-338*). Therefore, we postulate that selection of the miRNA arm might be decided by the mechanism "target-mediated miRNA protection (TMMP)" and TMMP is probably more selective to miRNA* strands.

Acknowledgments

The authors sincerely thank Scott Young at the NIH (Bethesda, MD, USA) for his critical reading of the manuscript. This research was supported by the Basic Science Research Program of the Korean National Research Foundation (NRF) funded by the Korean Ministry of Education, Science and Technology (Grant nos. 2011-0028240 and 2011-0014086).

Conflict of Interest

The authors declare no conflict of interest.

References

1. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297.
2. Kim, Y.K.; Heo, I.; Kim, V.N. Modifications of small RNAs and their associated proteins. *Cell* **2010**, *143*, 703–709.
3. Lewis, B.P.; Burge, C.B.; Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **2005**, *120*, 15–20.
4. Bushati, N.; Cohen, S.M. microRNA Functions. *Annu. Rev. Cell Dev. Biol.* **2007**, *23*, 175–205.
5. Lee, Y.; Jeon, K.; Lee, J.T.; Kim, S.; Kim, V.N. MicroRNA maturation: Stepwise processing and subcellular localization. *EMBO J.* **2002**, *21*, 4663–4670.
6. Rand, T.A.; Petersen, S.; Du, F.; Wang, X. Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell* **2005**, *123*, 621–629.
7. Matranga, C.; Tomari, Y.; Shin, C.; Bartel, D.P.; Zamore, P.D. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **2005**, *123*, 607–620.
8. Wang, H.W.; Noland, C.; Siridechadilok, B.; Taylor, D.W.; Ma, E.; Felderer, K.; Doudna, J.A.; Nogales, E. Structural insights into RNA processing by the human RISC-loading complex. *Nat. Struct. Mol. Biol.* **2009**, *16*, 1148–1153.
9. Leuschner, P.J.; Ameres, S.L.; Kueng, S.; Martinez, J. Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Rep.* **2006**, *7*, 314–320.
10. Iwasaki, S.; Kawamata, T.; Tomari, Y. Drosophila argonaute1 and argonaute2 employ distinct mechanisms for translational repression. *Mol. Cell.* **2009**, *34*, 58–67.
11. Diederichs, S.; Haber, D.A. Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* **2007**, *131*, 1097–1108.
12. Bhayani, M.K.; Calin, G.A.; Lai, S.Y. Functional relevance of miRNA sequences in human disease. *Mutat. Res.* **2012**, *731*, 14–19.
13. Khvorova, A.; Reynolds, A.; Jayasena, S.D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **2003**, *115*, 209–216.
14. Schwarz, D.S.; Hutvagner, G.; Du, T.; Xu, Z.; Aronin, N.; Zamore, P.D. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **2003**, *115*, 199–208.

15. Griffiths-Jones, S.; Hui, J.H.; Marco, A.; Ronshaugen, M. MicroRNA evolution by arm switching. *EMBO Rep* **2011**, *12*, 172–177.
16. Li, S.C.; Liao, Y.L.; Ho, M.R.; Tsai, K.W.; Lai, C.H.; Lin, W.C. miRNA arm selection and isomiR distribution in gastric cancer. *BMC Genomics Electron. Resour.* **2012**, *13*, S13.
17. Okamura, K.; Phillips, M.D.; Tyler, D.M.; Duan, H.; Chou, Y.T.; Lai, E.C. The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution. *Nat. Struct. Mol. Biol.* **2008**, *15*, 354–363.
18. Arvey, A.; Larsson, E.; Sander, C.; Leslie, C.S.; Marks, D.S. Target mRNA abundance dilutes microRNA and siRNA activity. *Mol. Syst. Biol.* **2010**, *6*, 363.
19. Larsson, E.; Sander, C.; Marks, D. mRNA turnover rate limits siRNA and microRNA efficacy. *Mol. Syst. Biol.* **2010**, *6*, 433.
20. Chatterjee, S.; Fasler, M.; Bussing, I.; Grosshans, H. Target-mediated protection of endogenous microRNAs in *C. elegans*. *Dev. Cell* **2011**, *20*, 388–396.
21. Ebert, M.S.; Neilson, J.R.; Sharp, P.A. MicroRNA sponges: Competitive inhibitors of small RNAs in mammalian cells. *Nat. Methods* **2007**, *4*, 721–726.
22. Lee, H.J.; Palkovits, M.; Young, W.S., 3rd. miR-7b, a microRNA up-regulated in the hypothalamus after chronic hyperosmolar stimulation, inhibits Fos translation. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 15669–15674.
23. Hutvagner, G.; Simard, M.J. Argonaute proteins: Key players in RNA silencing. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 22–32.
24. Meister, G.; Landthaler, M.; Patkaniowska, A.; Dorsett, Y.; Teng, G.; Tuschl, T. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* **2004**, *15*, 185–197.
25. Liu, J.; Carmell, M.A.; Rivas, F.V.; Marsden, C.G.; Thomson, J.M.; Song, J.J.; Hammond, S.M.; Joshua-Tor, L.; Hannon, G.J. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **2004**, *305*, 1437–1441.
26. Kwak, P.B.; Tomari, Y. The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat. Struct. Mol. Biol.* **2012**, *19*, 145–151.

Reprinted from *IJMS*. Cite as: Thomas, M.; Lange-Grünweller, K.; Hartmann, D.; Golde, L.; Schlereth, J.; Streng, D.; Aigner, A.; Grünweller, A.; Hartmann, R.K. Analysis of Transcriptional Regulation of the Human miR-17-92 Cluster; Evidence for Involvement of Pim-1. *Int. J. Mol. Sci.* **2013**, *14*, 12273-12296.

Article

Analysis of Transcriptional Regulation of the Human miR-17-92 Cluster; Evidence for Involvement of Pim-1

Maren Thomas [1,†], Kerstin Lange-Grünweller [1,†], Dorothee Hartmann [1], Lara Golde [1], Julia Schlereth [1], Dennis Streng [1], Achim Aigner [2], Arnold Grünweller [1,*] and Roland K. Hartmann [1,*]

[1] Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, 35032 Marburg, Germany; E-Mails: thomasm@staff.uni-marburg.de (M.T.); langegru@staff.uni-marburg.de (K.L.-G.); Dorothee_Hartmann@web.de (D.H.); laragolde@gmx.de (L.G.); julia.schlereth@staff.uni-marburg.de (J.S.); dennis.streng@pharmazie.uni-marburg.de (D.S.)

[2] Medizinische Fakultät, Rudolf-Boehm-Institut für Pharmakologie und Toxikologie, Klinische Pharmakologie, Universität Leipzig, 04107 Leipzig, Germany; E-Mail: Achim.Aigner@medizin.uni-leipzig.de

† These authors contributed equally to this work.

* Authors to whom correspondence should be addressed; E-Mails: arnold.gruenweller@staff.uni-marburg.de (A.G.); roland.hartmann@staff.uni-marburg.de (R.K.H.); Tel.: +49-6421-28-25553 (R.K.H.); Fax: +49-6421-28-25854 (R.K.H.).

Received: 29 March 2013; in revised form: 14 May 2013 / Accepted: 22 May 2013 / Published: 7 June 2013

Abstract: The human polycistronic miRNA cluster miR-17-92 is frequently overexpressed in hematopoietic malignancies and cancers. Its transcription is in part controlled by an E2F-regulated host gene promoter. An intronic A/T-rich region directly upstream of the miRNA coding region also contributes to cluster expression. Our deletion analysis of the A/T-rich region revealed a strong dependence on c-Myc binding to the functional E3 site. Yet, constructs lacking the 5'-proximal ~1.3 kb or 3'-distal ~0.1 kb of the 1.5 kb A/T-rich region still retained residual specific promoter activity, suggesting multiple transcription start sites (TSS) in this region. Furthermore, the protooncogenic kinase, Pim-1, its phosphorylation target HP1γ and c-Myc colocalize to the E3 region, as inferred from chromatin immunoprecipitation. Analysis of pri-miR-17-92 expression levels in K562 and

HeLa cells revealed that silencing of E2F3, c-Myc or Pim-1 negatively affects cluster expression, with a synergistic effect caused by c-Myc/Pim-1 double knockdown in HeLa cells. Thus, we show, for the first time, that the protooncogene Pim-1 is part of the network that regulates transcription of the human miR-17-92 cluster.

Keywords: miRNA; miR-17-92 cluster; Pim-1; miRNA promoter; c-Myc; HP1γ; RNAi

1. Introduction

MicroRNAs (miRNAs) are important post-transcriptional riboregulators of gene expression with high relevance to cancer formation and metastasis [1]. In general, miRNAs are derived from RNA polymerase II (RNAPII) primary transcripts (pri-miRNA) that are further processed to ~70 nt precursors (pre-miRNA) and after nuclear export to mature miRNAs by the activity of the two endonucleases, DROSHA/DGCR8 and DICER [2–5]. MiRNAs are incorporated into the miRNA-induced silencing complex (miRISC) and act as repressors of translation by imperfect base-pairing to their target sites in mRNAs [3]. The majority of miRNAs is encoded in intronic regions, either individually or as "polycistronic" miRNA clusters that are cotranscribed [3,6].

Several deregulated miRNAs or miRNA clusters are involved in tumorigenesis, accounting for their designation as tumor-suppressing or as oncogenic miRNAs [7]. Such miRNAs can downregulate targets involved in the regulation of apoptosis or cell cycle progression [1]. One well-characterized polycistronic cluster is the miR-17-92 cluster, also known as OncomiR-1 or C13*orf*25. This cluster encodes six miRNAs belonging to four different seed families: (i) the miR-17 family with miR-17 and miR-20a, (ii) the miR-18 family with miR-18a, (iii) the miR-19 family with miR-19a and miR-19b-1 and (iv) the miR-92 family [8–11]. The human miR-17-92 cluster is encoded in the chromosomal region, 13q31.3, and is amplified in several solid tumors, as well as in some hematopoietic malignancies [8,12]. Because of numerous known targets of its individual miRNAs, the miR-17-92 locus exerts pleiotropic functions during development, proliferation, apoptosis and angiogenesis in different cell systems [13–15]. In mice, deletion of the cluster prevents normal B-cell development as a consequence of premature cell death [14]. In a mouse B-cell lymphoma model, simultaneous overexpression of c-Myc and the miR-17-92 cluster accelerated lymphomagenesis [9]. This oncogenic effect could later be assigned primarily to miR-19a/b, which dampens expression of the tumor suppressor PTEN, thereby repressing apoptosis [13,15].

Analyses of transcriptional regulation of oncogenic miRNAs and miRNA clusters are of great importance for strategies aiming at cancer prevention. Unfortunately, most miRNA promoters have not been characterized or identified yet [16]. In the case of the miR-17-92 cluster, expression is thought to be promoted from a host gene promoter region upstream of exon 1, with transcription starting at a consensus initiator sequence downstream of a non-consensus TATA box [17,18]. Additionally, this core promoter region contains functional E2F transcription factor binding sites. E2F1-3 were shown to activate C13*orf*25 expression from this promoter and chromatin immunoprecipitation assays (ChIP) identified E2F3 to be the main E2F variant associated with the host gene promoter [17,18]. No E2F binding was detected in the region between the host gene promoter and the miR-17-92 cluster [18].

Furthermore, nucleosome mapping combined with chromatin signatures for transcriptionally active promoters [19–21] indicated that transcriptional activity of the miR-17-92 cluster also originates from the intronic A/T-rich region directly upstream of the miRNA coding sequences [16]. This is in line with the finding that cluster expression is activated by c-Myc binding to a conserved E-box element (E3) ~1.5 kb upstream of the miRNA coding sequence [9,10,20]. Indeed, luciferase reporter assays confirmed that both the host gene promoter and the intronic region confer transcriptional activity [16,21].

Here, we subjected the intronic A/T-rich region to deletion analysis using luciferase reporter constructs. Transcription was found to strongly depend on c-Myc binding to the E3 site, but even shorter fragments (<0.3 kb) of sequences directly preceding the miR-17-92 coding sequence still promoted residual, but substantial and specific transcriptional activity. Interestingly, we identified the protooncogene Pim-1 and one of its phosphorylation targets, HP1γ [22], to be associated with the chromatin region containing the E3 site, suggesting that the human C13*orf*25 locus belongs to the set of genes that are regulated by c-Myc and Pim-1 [23,24]. SiRNA-mediated Pim-1 knockdown indeed resulted in reduced pri-miR-17-92 levels, as did a knockdown of c-Myc or E2F3. In Hela cells, a double knockdown of c-Myc/Pim-1 decreased the pri-miR-17-92 levels more than single knockdowns, consistent with a synergism of c-Myc and Pim-1 at the intronic C13*orf*25 promoter.

2. Results and Discussion

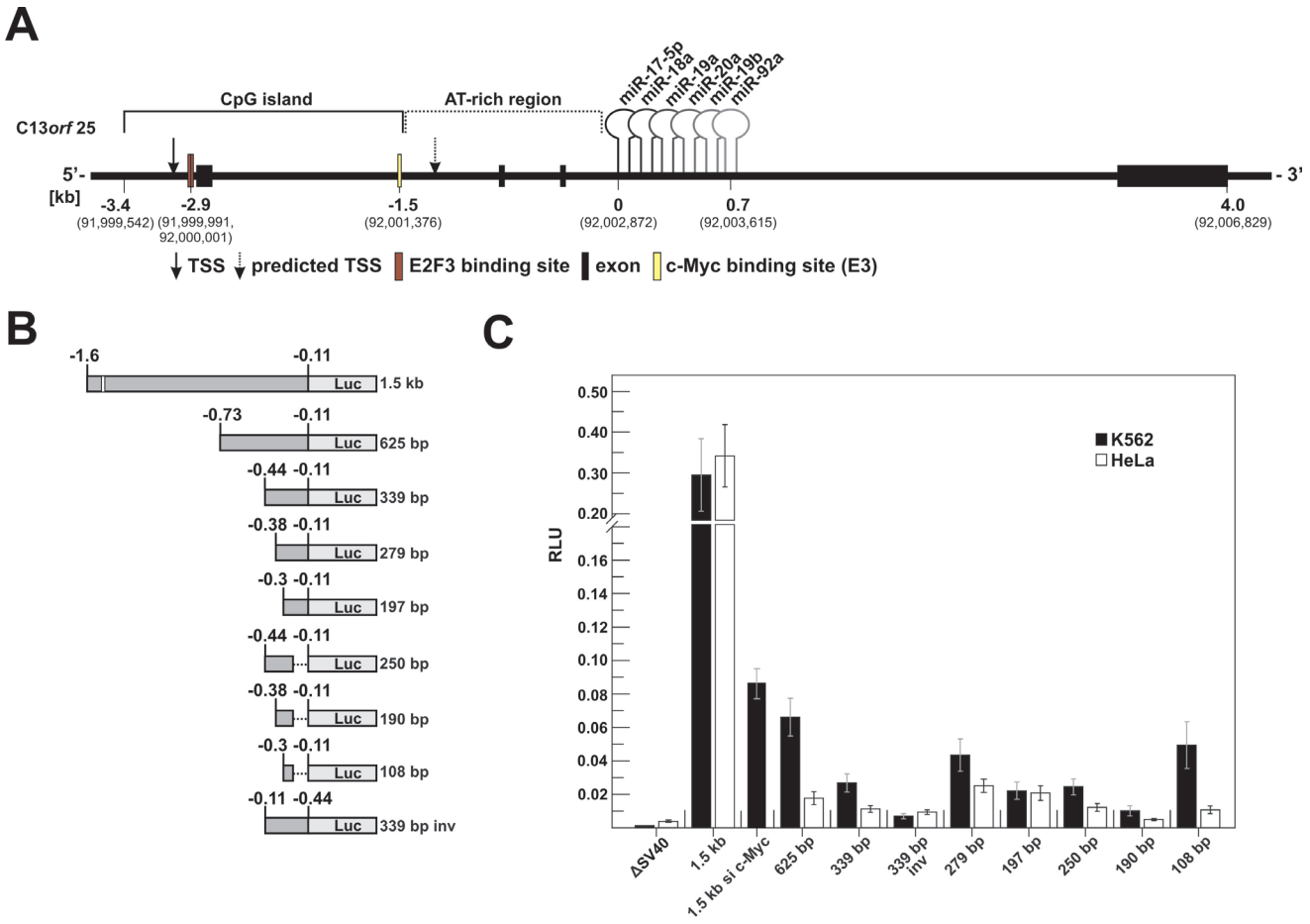
2.1. Results

2.1.1. c-Myc-Dependent Intronic Transcriptional Activity within the Human miR-17-92 Locus

The 3.4 kb upstream genomic region of the miR-17-92 coding sequence can be subdivided into a G/C-rich and an A/T-rich part. The former is a CpG island (~1.9 kb, 78% GC content; see http://genome.ucsc.edu [25], GRCh37/hg19 assembly) that has its 5'-boundary ~0.4 kb upstream of the TSS of the host gene promoter [20] and its 3'-boundary ~1.4 kb upstream of the miR-17-5p coding sequence, representing the 5'-terminal miRNA of the cluster. The A/T-rich region (~64% A/T content) following the CpG island begins immediately downstream of a functional and highly conserved c-Myc binding site (5'-CATGTG, E-box E3), which is localized ~1.5 kb upstream of the miR-17-5p coding sequence [10] (Figures 1A and S1).

We have analyzed the intronic region of C13*orf*25, including the preceding functional c-Myc box E3 [20] and truncated segments of the A/T-rich region (Figure 1B) for transcriptional activation. For this purpose, luciferase reporter constructs were transfected into K562 (a human myelogenous erythroleukemia cell line from a CML patient) and HeLa cells (an epithelial human cell line from a cervical carcinoma). We selected these two cell lines as a starting point to study transcription of the miR-17-92 cluster in the context of different cellular expression levels (Figure S2).

Figure 1. (**A**) Genomic organization of C13*orf*25. The locus consists of four exons and three introns; the six miRNAs of the miR-17-92 cluster are encoded in intron 3. Sequences upstream of the cluster can be subdivided into a G/C-rich CpG island and an A/T-rich downstream part. The host gene promoter thought to be activated by E2F3 is located in the CpG island about 3.4 kb upstream of the miR-17-5p coding sequence. The functional c-Myc site (E3) is located ~1.5 kb upstream of miR-17-5p. Sequence numbering is based on the NCBI reference sequence NG_032702.1 and the GRCh37/hg19 assembly [25]. Note that previous related studies referred to the numbering system of the previous hg18 assembly [16,17,20,21]. The numbering of the hg18 and hg19 assemblies is correlated as follows: nt 92,002,872 (0 kb in Figure 1A) of hg19 is nt 90,800,873 of hg18; (**B**) Schematic representation of the different C13*orf*25 portions fused to the luciferase structural gene. The functional E3 box for c-Myc binding is indicated in the 1.5 kb construct (white vertical line); (**C**) Promoter activities of the different luciferase reporter constructs in K562 and HeLa cells. Obtained luciferase activities were measured as relative light units (RLU) and normalized to the pGL3 control plasmid carrying the SV40 promoter (Promega). A reporter construct lacking the SV40 promoter, as well as a construct harboring the 339 bp fragment of the A/T-rich intronic region in inverted orientation were used as controls. RLU values of the individual constructs were derived from 5 to 16 experiments (+/− S.E.M.).



The ~1.5 kb reporter construct, comprising c-Myc box E3 (Figures 1A and S1) and the A/T-rich region (lacking the 113 bp preceding the mature miR-17-5p coding sequence for reasons of PCR feasibility), showed substantial transcriptional activity, amounting to 30%–35% in both cell lines relative to the pGL3 control plasmid harboring an SV40 promoter (Figure 1C). This is in line with results of a similar study of the mouse miR-17-92 locus [21]. Furthermore, Ozsolak *et al.* [16] predicted an intronic TSS to be localized ~0.2 kb downstream of the E3 site. Indeed, truncating the 1.5 kb fragment to 625 bp, which deletes the E3 site, strongly reduced reporter activity by ~4.5-fold in K562 and by almost 20-fold in HeLa cells compared to the activity of the ~1.5 kb construct (Figure 1C). To substantiate this finding, we tested the ~1.5 kb construct in K562 cells under conditions of a siRNA-mediated knockdown of c-Myc. This reduced reporter expression to a similar extent as the truncation to 625 bp, supporting the notion that c-Myc binding to the E3 site plays a key role in activating transcription from this intronic region (Figure 1C). SiRNA-mediated c-Myc knockdown in HeLa cells also suggests a ~four-fold decrease in transcription originating from the ~1.5 kb reporter construct (data not shown), again consistent with the crucial role of c-Myc binding to the E3 site. As the 625 bp fragment still conferred basal promoter activity, we further shortened this region to ~340 bp, ~280 bp and ~200 bp. Additionally, we included short fragments with their 3'-boundary ~290 bp upstream of the mature miR-17-5p coding sequence (250, 190 and 108 bp in Figure 1B). We also inversed the orientation of the ~340 bp fragment in front of the luciferase gene (Figure 1C, 339 bp inverse (inv)) to include a control fragment with comparable A/T content. This inversed fragment conferred reporter activity 5.3-fold (K562) or 2.4-fold (HeLa) higher than that of the pGL3 control vector lacking the SV40 promoter (ΔSV40, Figure 1C).

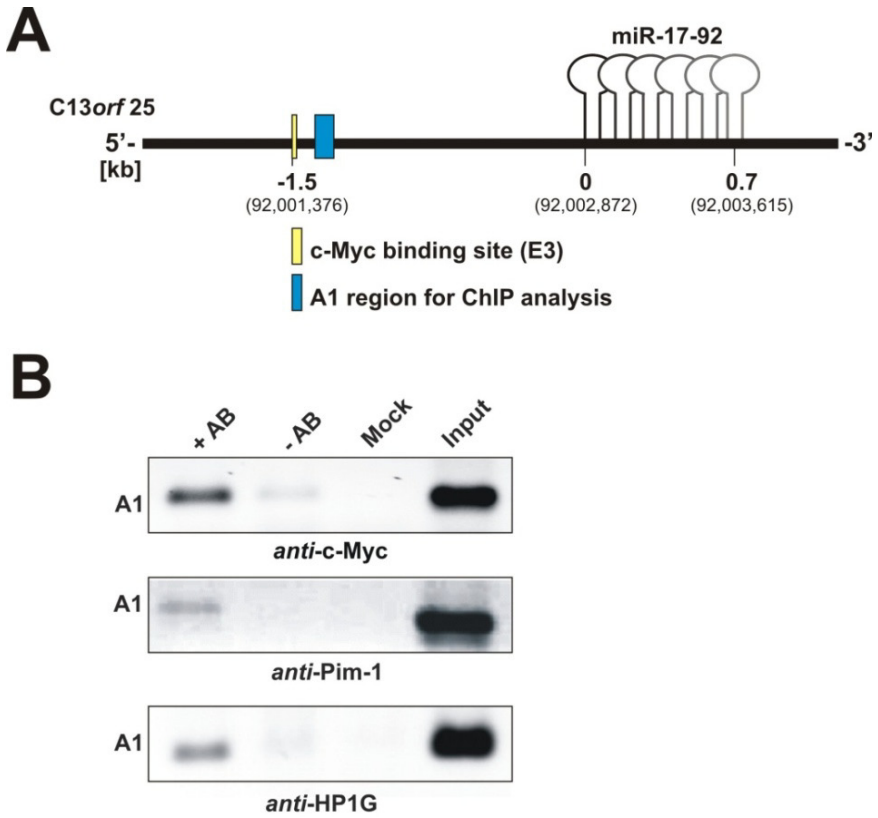
All the fragments≤ 340 bp conferred residual promoter activities, some clearly to a higher extent than the inverted 339 bp fragment in both cell lines (see the 279 and 197 bp fragments, Figure 1C). This indicates that parts of the intronic A/T-rich region promote specific transcriptional activity, the extent partly differing between the two cell lines (Figure 1C). Notably, despite using a variety of web-based promoter prediction tools (see Suppl. Material), no correlation between fragment activity and promoter elements predicted in this region was identified. In K562 cells, the smaller fragments, including the 625 bp fragment, showed an overall trend towards stronger expression relative to HeLa cells.

2.1.2. Pim-1 and HP1γ Are Associated with the Intronic c-Myc Binding Site

We next asked if other factors beyond c-Myc may be involved in human miR-17-92 cluster expression from the A/T-rich region. Transcriptional regulation by c-Myc is associated with Pim-1-dependent H3S10 phosphorylation in about 20% of all genes regulated by c-Myc [24]. Moreover, Pim-1 and c-Myc act synergistically in severe forms of B-cell lymphomas and Pim-1, as well as the miR-17-92 cluster are overexpressed in K562 cells [26]. We performed ChIP assays to test whether Pim-1 localizes to the internal promoter region of the miR-17-92 cluster. For this analysis, we amplified a ~90 bp DNA fragment (segment A1 in Figure 2A) 0.1 kb downstream of the functional c-Myc E3 site. The same DNA segment has been analyzed in a previous study on c-Myc [10]. Our ChIP analysis revealed that not only c-Myc, as expected, but also Pim-1 localizes to this genomic region (Figure 2B, left lanes in upper and middle panels). Indeed, this is consistent with the finding

that Pim-1-catalyzed H3S10 phosphorylation is required for c-Myc-dependent transcriptional activation [24]. We further analyzed another known phosphorylation target of Pim-1, the heterochromatin protein-1 gamma (HP1γ) [22], for its association with the E3 region. HP1γ localized to this genomic area, as well (Figure 2B, lower panel). Moreover, we were able to identify an association of HP1γ along the miRNA coding region, which is indicative of active transcription (see Figure S3 and Discussion section).

Figure 2. (**A**) Schematic representation of the intronic A/T-rich region preceding the miR-17-92 coding sequence. The region A1 (blue box) defines the genomic sequence 0.1 kb downstream of the functional c-Myc binding site (E3; yellow box) that was amplified in ChIP analyses; (**B**) ChIP analysis of the intronic region A1 in K562 cells, using antibodies specific for c-Myc, Pim-1 and HP1γ. +AB: with antibody; −AB: without antibody; Mock: buffer only without cell lysate; Input: supernatant of the −AB -sample after immunoprecipitation and centrifugation (for details, see Supplementary Materials).

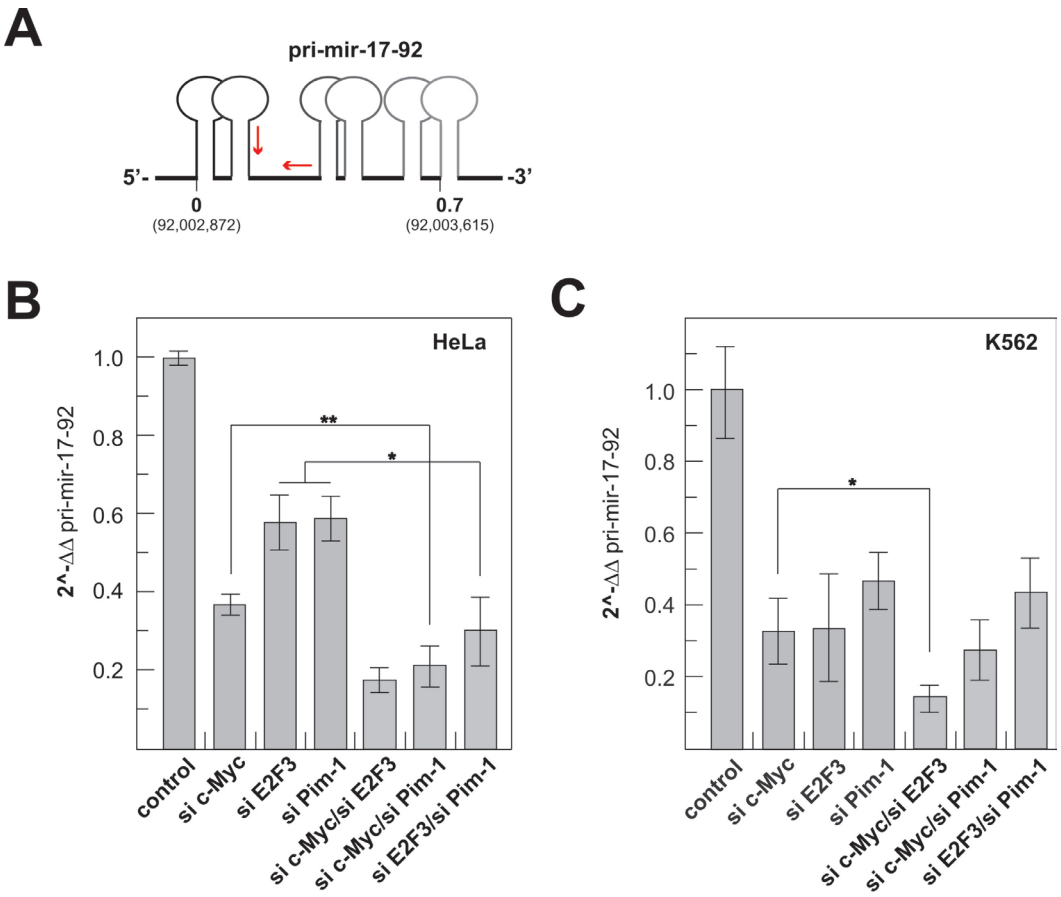


2.1.3. Transcriptional Activity of the Human miR-17-92 Cluster Depends on c-Myc and Pim-1

To further substantiate the role of Pim-1 in miR-17-92 cluster expression, we quantified the cellular pri-miR-17-92 levels by qRT-PCR (see Figure 3A for primer positions) after siRNA-mediated Pim-1 knockdown relative to a c-Myc knockdown in K562 and HeLa cells. Since combined ChIP and reporter gene assays suggested that the transcription factor E2F3 is a major activator of transcription initiated at the host gene promoter [17,18], we included E2F3 in our knockdown experiments as a possible measure for the contribution of the host gene promoter to miR-17-92 expression. We also quantified the levels of c-Myc, E2F3 and Pim-1 mRNAs after knockdown by qRT-PCR to evaluate

knockdown efficiencies (Supplementary Table S1). For Pim-1, we have shown good correlation between mRNA and protein levels [26], suggesting that reduced mRNA levels will also entail decreased protein levels. A corresponding parallel analysis of protein levels was inconclusive, owing to a non-interpretable pattern obtained with the used E2F3 antibody [18]. In the study presented here, only experiments with a knockdown efficiency >50% were included (Supplementary Table S1). Single knockdowns of either c-Myc or E2F3 decreased the pri-miR-17-92 levels in HeLa cells to ~35 and 60%, respectively, relative to the control siRNA (Figure 3B). Notably, a 40% reduction of pri-miR-17-92 levels was also observed upon Pim-1 knockdown. Similar results were obtained in K562 cells, with decreases in pri-miR-17-92 levels to ~30%, 30% and 45%, respectively (Figure 3C). However, double knockdowns had additive suppression effects on pri-miR-17-92 levels in the case of c-Myc/E2F3 (HeLa and K562), c-Myc/Pim-1 (Hela) and E2F3/Pim-1 (HeLa). To shed more light on the role of Pim-1, we further analyzed luciferase activity of the 1.5 kb construct harboring the functional c-Myc E3 site in K562 and HeLa cells upon Pim-1 knockdown. We did not observe a substantial decrease in reporter expression after Pim-1 knockdown in K562 cells (data not shown), but a three-fold reduction (Supplementary Figure S4) in HeLa cells (see Discussion).

Figure 3. (**A**) Illustration of the primers (red arrows) used for the qRT-PCR quantification of pri-miR-17-92 transcript levels; (**B**,**C**) qRT-PCR-based quantitation of pri-miR-17-92 transcript levels in HeLa (**B**) or K562 cells (**C**) after siRNA-mediated knockdown of c-Myc, E2F3 or Pim-1 or after combined knockdown of c-Myc/E2F3, c-Myc/Pim-1 or E2F3/Pim-1. 2^-ΔΔpri-17-92 values were normalized against 5S rRNA and an internal control siRNA (siVR1), representing mean values from at least three independent experiments (+/− S.E.M.). Statistical analyses were done using the software, R.



2.2. Discussion

The transcription of the oncogenic miR-17-92 cluster is thought to originate from two different TSSs: one is localized in close proximity to the host gene promoter element [17] (Supplementary Figure S1), and the other TSS was predicted to map to the region ~200 bp downstream of the functional c-Myc site E3 (Figures 1A and S1). The latter prediction was based on nucleosome mapping and chromatin signatures for active promoters. The derived algorithm identified 175 human promoters proximal to miRNA coding sequences and was reported to correctly predict transcription initiation regions to a resolution of 150 bp with high sensitivity and specificity. The majority of predictions were also consistent with known “expressed sequence tag” (EST) TSSs or cDNA 5'-ends [16].

Beyond previous studies [16,20,21], we investigated the intronic A/T-rich region preceding the human miR-17-92 cluster in more detail and compared it to siRNA-mediated knockdown of c-Myc. Similar effects were obtained, substantiating the notion that c-Myc and the c-Myc E3 site play a crucial role in activating transcription from the intronic promoter region. However, the 625 bp and even some of the further truncated fragments (~280 and ~200 bp) of the A/T-rich region conferred residual specific promoter activity in both cell types (Figure 1), indicating that parts of the A/T-rich region, downstream of the c-Myc E3 site, contribute to cluster expression. This E3 box-independent transcriptional activity was more pronounced for K562 relative to HeLa cells, which correlates with the particularly high cluster expression in K562 cells (Supplementary Figure S2). As a possible explanation, transcriptional activity of the ~1.5 kb fragment may be dominated by the recruitment of c-Myc to the E3 site region, while differential activity mediated by the smaller fragments in K562 *vs.* HeLa cells may report that their residual transcriptional activation is mechanistically different from

that of the ~1.5 kb fragment. This could mean that regulatory factors of the transcription machinery are differentially expressed in the two cell lines.

ChIP assays revealed that not only c-Myc, but also the protooncogene Pim-1 and its phosphorylation target, HP1γ, associate with the chromatin region harboring the c-Myc E3 site (Figure 2B). Importantly, Pim-1-catalyzed phosphorylation of H3S10 at c-Myc target genes is necessary to regulate key genes required for c-Myc-dependent oncogenic transformation [27].

In mammals, three paralogs of HP1 (α, β and γ) regulate heterochromatin formation, gene silencing or gene activation [28,29]. HP1α and β proteins are mainly recruited to heterochromatin regions harboring H3K9me2,3 modifications, whereas HP1γ is found in association with euchromatin [30] and active genes [29]. Furthermore, HP1c, the *Drosophila* homolog of HP1γ, associates with transcriptionally active chromatin containing H3K4me3 and H3K36me3 histone marks [28]. HP1γ can further be recruited to inducible promoters, where it replaces HP1β, thereby inducing a switch from the repressive to the active transcriptional state. This replacement with HP1γ requires H3 phospho-acetylation [31]. In this context, a transient phosphorylation of H3S10 (via Aurora B kinase) was shown to be necessary for the dissociation of HP1 proteins from chromatin during the M phase of the cell cycle [32]. In the induced state, HP1γ can also be localized within coding regions of protein genes, together with elongating RNA polymerase II [31].

Our data, showing that HP1γ colocalizes with Pim-1 and c-Myc (Figure 2), is in line with the aforementioned activating role of HP1γ during transcription. We extended our ChIP assays to the miRNA coding region of C13*orf*25 to analyze HP1γ association with this part of the cluster. Indeed, ChIP analysis along the miRNA coding sequence identified HP1γ at all four analyzed subregions (A2–A5, Supplementary Figure S3). To our knowledge, this is the first indication that HP1γ is involved in activating the transcription of miRNAs.

The association of Pim-1 with the intronic chromosomal region near the c-Myc E3 site led us to the assumption that Pim-1 plays an important role in the transcriptional activation of the miR-17-92 cluster. This was tested by RNAi also, including E2F3 as an assumed indicator of host gene promoter activity. The strong negative effects of individual knockdowns of c-Myc, Pim-1 and E2F3 on pri-miR-17-92 levels indicate that all three proteins are important for cluster expression by affecting transcription from the host gene promoter (E2F3) or the intronic promoter region (c-Myc, Pim-1). This raises the question about the mechanistic role of Pim-1 in cluster expression from the intronic promoter. In contrast to HeLa cells Supplementary (Figure S4), a Pim-1 knockdown in K562 cells failed to significantly decrease reporter activity from the ~1.5 kb fragment. Among other possibilities, Pim-1 may be recruited to the functional c-Myc E3 site in the context of the cellular chromatin structure in K562 cells, but not in the context of the plasmid-encoded reporter gene. Alternatively, Pim-1 recruitment to the E3 site occurs, as shown by the ChIP assays, but is not a crucial prerequisite for transcriptional activation in K562 cells. On the other hand, the three-fold decrease in ~1.5 kb reporter activity observed in HeLa cells upon Pim-1 depletion adds evidence in support of a crucial role for Pim-1 in miR-17-92 cluster expression, but simultaneously points to cell type-specific differences. For future investigations, other cell lines will be tested, particularly ones that express c-Myc, but not Pim-1. Clearly, decreases in pri-miR-17-92 levels upon c-Myc, E2F3 and/or Pim-1 knockdown (Figure 3) may include indirect effects, e.g., originating from inhibition of cell proliferation (Pim-1), changes in the kinetics of pri-miR-17-92 processing, global changes in

transcriptional networks (E2F3, c-Myc) or mutual transactivation (E2F3 and c-Myc) [33–35]. Further complication may arise from the fact that miR-17-5p and miR-20a of the cluster are negative regulators of E2F1-3 mRNAs [10,18].

As c-Myc, HP1γ and histone H3 are known phosphorylation targets of Pim-1, future studies may address the influence of Pim-1 on the phosphorylation status of these proteins at the E3 site, utilizing antibodies that are highly specific for the phosphorylated *versus* unphosphorylated state.

The siRNA-mediated c-Myc knockdown, decreasing c-Myc mRNA levels on average by 65% (HeLa) or 81% (K562; see Supplementary Table S1), resulted in a 60%–70% reduction in pri-miR-17-92 levels in HeLa and K562 cells (Figure 3). This effect may report a rough estimate of the proportion of cluster transcripts normally initiated in the intronic promoter region in these two cell lines, for the following reasons: the C13*orf*25 region contains four c-Myc binding sites (boxes E1-4) and two additional ones with lower c-Myc occupancy (relative to E1) upstream of the host gene promoter [20]. Box E1, immediately downstream of host gene promoter's TSS, was shown by deletion analysis to be inhibitory, which correlates with c-Myc forming heterodimers with MXI or MNT at this site to repress transcription [20]. Thus, host gene promoter activity may even somewhat increase under conditions of a c-Myc knockdown, although such an effect could, in turn, be neutralized by reduced c-Myc-mediated transactivation of E2F [35]. ChIP-Seq data for K562 and HeLa-S3 cells revealed the by far highest c-Myc occupancy at site E3 (little at E2 and E4), where c-Myc forms heterodimers with MAX to activate transcription [20]. A straightforward interpretation of the additive effect of a c-Myc/E2F3 double knockdown in Hela and K562 cells is that this combination negatively affected transcription from the host gene and intronic promoter regions.

A major finding of our study is the recruitment of Pim-1 to the intronic c-Myc E3 site (Figure 2) and the strong negative effect of a Pim-1 knockdown on cluster expression (Figure 3B,C). Interestingly, Pim-1 knockdown efficiencies are comparable in K562 (73%) and HeLa (71%) cells, whereas the effect of the knockdown on cluster expression is stronger in K562 cells (55% reduction compared to 40% in HeLa cells) with the higher Pim-1 expression level. This might be due to cell type-dependent indirect effects of Pim-1 on the regulation of the miR-17-92 cluster. Moreover, double knockdown experiments in HeLa cells revealed a synergistic effect relative to individual c-Myc and Pim-1 knockdowns (Figure 3B), which was not seen for K562 cells. The siRNA-mediated reduction of c-Myc and Pim-1 mRNAs were on average 86% and 77% in HeLa and 86% and 52%, respectively, in K562 cells (Supplementary Table S1). The somewhat weaker suppression of Pim-1 in the c-Myc/Pim-1 double knockdown context (cf. with single knockdowns, Table S1) in K562 *versus* HeLa cells may have contributed to the absence of a clear additive effect upon c-Myc/Pim-1 double knockdown in K562 cells.

3. Experimental Section

3.1. Oligonucleotides

Small interfering RNAs (siRNAs) were purchased from Dharmacon (Boulder, CO, USA):

VR1 siRNA [36] was used as an unrelated negative control, with the following sequences of sense and antisense strand.

VR1 siRNA sense 5'-GCG CAU CUU CUA CUU CAA CdTdT and antisense 5'-GUU GAA GUA GAA GAU GCG CdTdT.

The sequences of all other siRNAs used in this study are:

Pim-1 siRNA sense 5'-GAU AUG GUG UGU GGA GAU AdTdT and antisense 5'-UAU CUC CAC ACA CCA UAU CdTdT; Pim-1 siRNA 2 sense→5'-GGA ACA ACA UUU ACA ACU CdTdT and antisense 5'-GAG UUG UAA AUG UUG UUC CdTdT; c-Myc siRNA sense 5'-CAG GAA CUA UGA CCU CGA CUA dTdT and antisense 5'-UAG UCG AGG UCA UAG UUC CUG dTdT; E2F3 siRNA sense 5'-ACA GCA AUC UUC CUU AAU AdTdT and antisense 5'-UAU UAA GGA AGA UUG CUG UdTdT.

3.2. Antibodies

Antibodies against c-Myc (sc-40) and Pim-1 (sc-13513), as well as the secondary antibody (sc-2005: goat anti-mouse IgG HRP conjugated) were purchased from Santa Cruz Biotechnology (Heidelberg, Germany) except for the Phospho HP1γ (Ser83) antibody (2600S), which was obtained from Cell Signaling Technology (Danvers, MA, USA).

3.3. Plasmid Construction and Seed Mutagenesis

For the construction of promoter-luciferase fusions, the SV40 promoter of plasmid "pGL3 control" (Promega, Mannheim, Germany) was removed via digestion with *Bgl*II and *Hind*III (Thermo Fisher Scientific, Schwerte, Germany) and replaced with fragments derived from the intronic A/T-rich region of C13*orf*25 (reference nucleotide sequence NG_032702.1). Promoter fragments were amplified from genomic DNA of the human cell line K562 using primers specified in the Supplementary Material. PCR products were purified using the Wizard® SV Gel and PCR Clean-Up System (Promega, Mannheim, Germany) and digested with *Bgl*II and *Hind*III for insertion into pGL3. All constructs were cloned in *E. coli* DH5α cells and verified by DNA sequencing. The pGL3 vector lacking the SV40 promoter, as well as the pGL3 construct carrying the C13*orf*25-derived 339 bp fragment in inverse orientation (pGL3 339 bp inv), were used as negative controls.

3.4. Transfection Procedures and Luciferase Reporter Assays

Assays are described in detail in the Supplementary Materials.

3.5. Chromatin Immunoprecipitation

Chromatin immunoprecipitation (ChIP) was performed according to a protocol from the Or Gozani lab at Stanford University (http://www.stanford.edu/group/gozani) [37], with some modifications. For details, see the Supplementary Material.

4. Conclusions

We report here that miR-17-92 cluster expression from the intronic A/T-rich promoter region, although critically depending on c-Myc binding, includes some specific contribution of sequences within ~0.7 kb upstream of the mature miR-17-5p coding sequence. Our reporter expression data

suggest multiple TSSs within this A/T-rich region, although the transcription initiation region predicted by Ozsolak *et al.* [16], ~0.2 kb downstream of the c-Myc E3 box, may well be the major one. E3 site-independent transcription initiation within ~0.7 kb upstream of the mature miR-17-5p coding sequence was more pronounced in K562 *versus* HeLa cells (Figure 1), indicating cell type-specific differences in cluster expression from the intronic promoter region. By RNAi and ChIP, we establish that Pim-1 acts in concert with c-Myc at the E3 site to activate transcription from the intronic promoter region.

Acknowledgments

We are grateful to Lisa Schemberger, Nicole Bürger and Moana Klein for technical assistance, Marcus Lechner for statistical analyses and Markus Gößringer for fruitful discussions. The work was supported by grants of the Fritz Thyssen Stiftung (reference no. 10.06.1.186 to A.G. and R.K.H.) and the German Cancer Aid (Deutsche Krebshilfe, grants 106992 and 109260 to A.G., R.K.H. and A.A.).

Conflict of Interest

The authors declare no conflict of interest.

Appendix

Cell Culture

Cells (K562 and HeLa) were cultivated under standard conditions in a humidified atmosphere (37 °C, 5% CO_2) supplemented with RPMI 1640 (K562) or IMDM (HeLa) containing 10% FCS (PAA, Cölbe, Germany).

Chromatin Immunoprecipitation (ChIP)

2×10^7 K562 cells in 13 mL RPMI medium were crosslinked with 1% formaldehyde for 10 min at 37 °C. Reactions were stopped by adjusting to 0.125 M glycine, and cells were collected by centrifugation at 400*g* for 5 min at room temperature. For cell lysis, cells were resuspended in 750 µL RIPA-buffer (10 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% deoxycholate, 1% NP40, 0.1% SDS, 0.2 mM PMSF) supplemented with the complete Mini Protease Inhibitor Mix from Roche (Mannheim, Germany) according to the manufacturer's instructions. Lysed cells were sonicated in a Branson Sonifier 250 (duty cycle 50%, output control 2, for 3.5 min, on ice water) (Heinemann, Schwäbisch Gmünd, Germany) and centrifuged at 16,000*g* for 10 min at 4 °C. The supernatant was pre-cleared with 10 µL of blocked *Staphylococcus aureus* cells (Pansorbin® Cells, Calbiochem/Merck, Darmstadt, Germany) for 15 min at 4 °C on a rotor wheel. After a second centrifugation step (16,000*g*, 5 min, room temperature), the supernatant was split into two samples (each ~350 µL, representing + and− specific antibody (AB)), which were adjusted to buffer D (16.7 mM Tris-HCl pH 8.1, 167 mM NaCl, 1.2 mM EDTA, 1.1% Triton-X 100, 0.01% SDS) and a total volume of 500 µL. 1 µg (1 to 5 µL) of the respective antibody was added to "+AB" samples, whereas the same volume buffer D was added to "−AB" samples. At this point, also the "Mock" control was prepared, consisting of 500 µL buffer D.

"+AB", "–AB " and "Mock" samples were then incubated for at least 3 h at 4 °C. The following antibodies were used: monoclonal anti-c-Myc (sc-40) and anti-Pim-1 (sc-13513) (Santa Cruz Biotechnology (Heidelberg, Germany), anti-Phospho HP1γ (Ser83) polyclonal antibody (2600S) (Cell Signaling Technology, Danvers, MA, USA). In the case of mouse monoclonal antibodies (c-Myc and Pim-1), samples were additionally incubated for 1 h with a second monoclonal goat anti-mouse IgG antibody (sc-2005, Santa Cruz Biotechnology, Heidelberg, Germany). Immunoprecipitation was initiated by adding 10 µL of Pansorbin® cells (see above) to the "+AB", "–AB" and "Mock" samples, followed by incubation for 15 min at room temperature. Samples were centrifuged (16,000× *g*, 3 min, room temperature); the supernatant of the "–AB" sample was saved, later serving as the input control. Pellets were washed twice with dialysis buffer (50 mM Tris-HCl pH 8.0, 2 mM EDTA) and four times with IP-wash buffer (100 mM Tris-HCl pH 9.0, 500 mM LiCl, 1% NP40, 1% deoxycholate). Antibody-bound material was eluted from Pansorbin® cells by adding 150 µL elution buffer (50 mM $NaHCO_3$, 1% SDS), vortexing for 3 s, and centrifugation (16,000 *g*, 3 min, room temperature). The supernatant was collected and the procedure was repeated. Reverse crosslinking and RNA digestion was performed in 280 µL buffer [0.3 M NaCl, and 1 µL RNase A (10 mg/mL)] for 5 h at 67 °C. Chromatin was precipitated with ethanol, followed by a Proteinase K digest (Thermo Fisher Scientific, Bremen, Germany). DNA was purified by phenol/chloroform extraction and ethanol precipitation in the presence of 0.3 M NaOAc, pH 5.2. PCR amplification using Taq DNA polymerase and the co-immunoprecipitated DNA as template was done under the following conditions: 2 min at 95 °C in the absence of enzyme, followed by 35 amplification cycles of 45 s at 95 °C/45 s at 60 °C/45 s at 72 °C using the following primers:

A1 forward 5'-AAA GGC AGG CTC GTC GTT G
A1 reverse 5'-CGG GAT AAA GAG TTG TTT CTC CAA
A2 forward 5'-ACA TGG ACT AAA TTG CCT TTA AAT G
A2 reverse 5'-AAT CTT CAG TTT TAC AAG GTG ATG
A3 forward 5'-ACT GCA GTG AAG GCA CTT GT
A3 reverse 5'-TGC CAG AAG GAG CAC TTA GG
A4 forward 5'-CCA ATA ATT CAA GCC AAG CAA
A4 reverse 5'-AAA TAG CAG GCC ACC ATC AG
A5 forward 5'-GCC CAA TCA AAC TGT CCT GT
A5 reverse 5'-CGG GAC AAG TGC AAT ACC AT

Transfection of Reporter Constructs

For transfection of the suspension cell line K562, cells were washed in medium without serum, followed by electroporation in a BioRad Gene Pulser XCell (München, Germany) with a single pulse in a 4 mm cuvette, using 5 µg of the respective pGL3 derivative plasmid per million cells. 48 h after transfection, cells were washed in PBS, lysed and prepared for luciferase reporter assay measurements.

Transfection of siRNAs

Transfection of the suspension cell line K562 was performed by electroporation. After a washing step in medium without serum, 10^6 cells were mixed with 1 µg of siRNA (VR1 siRNA, Pim-1 siRNA, c-Myc siRNA or E2F3 siRNA). For the double knockdown experiments, 1 µg of each siRNA,

respectively, was used. Cells were electroporated at 330 V with a single pulse in a BioRad Gene Pulser XCell (München, Germany) using a 4 mm cuvette. After electroporation, K562 cells were resuspended in medium containing 10% FCS and cultivated in a humidified atmosphere at 37 °C for 24 h. Cells were washed with PBS and prepared for total RNA extraction.

Transfection of HeLa cells was done using the transfection agent Lipofectamine™ 2000 (Life Technologies Invitrogen, Darmstadt, Germany). One day before transfection, 8×10^4 cells were seeded into 24-well plates and cultivated under standard conditions. SiRNA complexes (VR1 siRNA, Pim-1 siRNA, c-Myc siRNA or E2F3 siRNA) were prepared according to the manufacturer's protocol. To perform single knockdown experiments, 40 pmol of the respective siRNA were used. In case of the double knockdown experiments, 20 pmol of each siRNA were mixed in Opti-MEM® 1 (Life Technologies Invitrogen, Darmstadt, Germany). 4 to 6 h after transfection, the medium was replaced with IMDM containing 10% FCS. Cells were cultivated 48 h under standard conditions until preparation for total RNA extraction.

RNA Preparation and Quantitative Real-Time PCR

For total RNA isolation, transfected cells (K562 or HeLa) were lysed (vortexing or mixing by pipetting up and down) in 750 µL lysis solution (0.8 M guanidinium-thiocyanate, 0.4 M ammonium-thiocyanate, 0.1 M sodium acetate pH 5.0, 5% glycerin, 38% phenol pH 4.5–5.0 (Roth®-Aqua-Phenol, Roth, Karlsruhe, Germany), 1 pellet 8-hydroxychinolin). Then, 200 µL of chloroform was added and phases were separated by centrifugation. The aqueous phase was mixed with 2 volumes of isopropanol, followed by incubation for 15 min at room temperature and centrifugation. The air-dried RNA pellet was dissolved in RNase-free water and incubated for 30 min at 37 °C with 1 U DNase I per µg RNA in 100 µL 1× DNase I buffer (DNase I, Thermo Fisher Scientific, Schwerte, Germany) according to the manufacturer's instructions. Then, another identical aliquot of DNase I was added, followed by incubation at 37 °C for another 30 min. Samples were extracted with an equal volume of Roth®-Aqua-Phenol (see above), followed by extraction of the aqueous phase with chloroform and isopropanol precipitation as above. RNA pellets were finally washed with 75% ethanol, air-dried and redissolved in 10 µL RNase-free water. 0.5 to 1 µg of total RNA were reverse-transcribed with RevertAid H Minus RT Polymerase (Thermo Fisher Scientific) according to the manufacturer's protocol. For determination of KD efficiencies a random hexamer primer was used to generate cDNA samples. In case of calculating pri-mir-17-92 levels the gene-specific reverse primer specified below was used for cDNA synthesis. Quantitative RT-PCR was performed in duplicate in a BioRad iQ™5 (BioRad, München, Germany) with the Absolute qPCR SYBR Green Capillary Mix (Thermo Scientific AbGene, Hamburg, Germany); cDNAs were diluted 1:5 or 1:10 and 4 µL of the reaction mixture used for determining RNA transcription levels. Quantitative PCR assays for miRNA detection were conducted as follows: Thermo-Start™ DNA polymerase was activated for 15 min at 95 °C followed by 55 amplification cycles of 10 s at 95 °C/20 s at 60 °C/12 s at 72 °C. Subsequently, melting curves of the PCR products were generated: samples were cooled from 95 to 65 °C (20 °C per s), kept at 65 °C for 20 s, followed by heating steps of 1 °C per cycle up to 95 °C and kept for 10 s at 95 °C.

Quantitative RT-PCR assays for mRNA detection were changed as follows: Thermo-Start™ DNA polymerase was activated for 15 min at 95 °C followed by 55 amplification cycles with a denaturation step for 10 s, primer annealing for 10 s at 55 °C and amplification at 72 °C for 10 s. Subsequently, a melting curve was generated for the PCR products; samples cooled from 95 to 65 °C (20 °C per s), kept at 65 °C for 20 s, followed by heating steps of 1 °C per cycle up to 95 °C and kept for 10 s at 95 °C.

The mRNA and pri-miR-17-92 levels were calculated from the crossing points by the 2^-$\Delta\Delta C_T$ method [38] using β-Actin mRNA or 5S rRNA as internal controls. Knockdown efficiency was quantitated by qRT-PCR and only those experiments showing more than 65% reduction in protein levels were used for quantification of pri-miR-17-92 levels. All primers for qPCR measurements were purchased from Metabion (Martinsried, Germany) and designed using the software tool Universal ProbeLibrary (Roche Applied Biosystems, Mannheim, Germany). A list of all primer sequences according to the human sequences is shown underneath:

Pri-miR-17-92: forward primer 5'-CAT CTA CTG CCC TAA GTG CTC CTT and reverse primer 5'-GCT TGG CTT GAA TTA TTG GAT GA;

5S rRNA: 5'-TCT CGT CTG ATC TCG GAA GC and 5'-AGC CTA CAG CAC CCG GTA TT;

c-Myc mRNA: forward primer 5'-CCT TGC AGC TGC TTA GAC and reverse primer 5'-GAG TCG TAG TCG AGG TCA T;
E2F3 mRNA: forward primer 5'-GAG ACT GAA ACA CAC AGT CC and reverse primer 5'-CCT GAG TTG GTT GAA GCC;
Pim-1 mRNA: forward primer 5'-ATC AGG GGC CAG GTT TTC T and reverse primer 5'-GGG CCA AGC ACC ATC TAA T;
Actin mRNA: forward primer 5'-CCA ACC GCG AGA AGA TGA and reverse primer 5'-CCA GAG GCG TAC AGG GAT AG.

Luciferase Reporter Assays

Luciferase reporter assays were performed using the Promega Luciferase Assay System (Promega, Mannheim, Germany). After aspirating the medium, cells were washed with PBS and lysed in 100 µL 1.2× reporter lysis buffer. In a 96-well plate, 10 µL of the respective cell lysate were mixed with 10 µL of luciferase substrate. Chemiluminescence was measured immediately in a Safire$^{2™}$ micro-plate reader (Tecan, Crailshaim, Germany).

Plasmid Construction and Seed Mutagenesis

All primers, which were used for plasmid construction and seed mutagenesis, are listed below (restriction sites in italics):
pGL3 1.5 kb: 5'-ATA T*AG ATC T*TG CCG CCG GGA AAC GGG TT and reverse primer R1 5'-ATA T*AA GCT T*CC ATA CAA ATT CAG CAT AAT CCC TAA TGG;
pGL3 625 bp: 5'-ATA TAG ATC TCT TTA GAC AAT GTA CCT TTT CTG and reverse primer R1 (see above);
pGL3 339 bp: 5'-ATA T*AG ATC T*GT GGA AGC CAG AAG AGG AGG A and reverse primer R1 (see above);
pGL3 279 bp: 5'-ATA TAG ATC TGG TAC ACA TGG ACT AAA TTG CC and reverse primer R1 (see above);
pGL3 197 bp: 5'-ATA T*AG ATC T*CT CTA TGT GTC AAT CCA TTT GGG AG and reverse primer R1 (see above);
pGL3 250 bp: 5'-ATA T*AG ATC T*GT GGA AGC CAG AAG AGG AGG A and reverse primer R3 5'- ATA T*AA GCT T*GC CTT AAG AAT TCT TTA CAG AAG GC;
pGL3 190 bp: 5'-ATA T*AG ATC T*GG TAC ACA TGG ACT AAA TTG CC and reverse primer R3 (see above);
pGL3 108 bp: 5'-ATA T*AG ATC T*CT CTA TGT GTC AAT CCA TTT GGG AG and reverse primer R3 (see above);
pGL3 339 inv: 5'-ATA T*AA GCT T*GT GGA AGC CAG AAG AGG AGG A and 5'-ATA T*AG ATC T*CC ATA CAA ATT CAG CAT AAT CCC TAA TGG.

Bioinformatical Promoter Analysis

For promoter analyses of the intronic A/T-rich region, a sequence of 1490 bp (human), beginning at the functional c-Myc site (E3) and ending at the 5'-end of the first mature miRNA sequence, miR-17-5p (for details, see Supplementary Figure S1), was analyzed with several web-based promoter prediction tools. The following tools were used to predict promoter elements or putative TSSs:

Neural Network promoter prediction: (http://www.fruitfly.org/seq_tools/promoter.html) [39];
McPromoter006: (http://tools.genome.duke.edu/generegulation/McPromoter/) [40];
Promoter 2.0 Prediction Server: (http://www.cbs.dtu.dk/services/Promoter/) [41];
PromPredict: (http://nucleix.mbu.iisc.ernet.in/prompredict/prompredict.html) [42].

Putative TSSs were predicted and calculated using the software available at http://rulai.cshl.org/tools/genefinder/CPROMOTER/human.htm [43].

All promoter prediction tools predicted several different promoter elements in the 1.5 kb region upstream of miR-17-5p and only the web-based tool Promoter 2.0 failed. The calculated promoter predictions of the indicated tools did not match in any sequence region, giving the assumption that the A/T-rich 1.5 kb intronic region in front of the miR-17-92 coding sequence has only weak promoter activity itself. Due to the fact that this intronic region has an overall high A/T-content, nearly all software programs were able to detect putative promoter regions.

Statistical Analysis

Statistical analyses were done using the software R [44]. *p*-Values were calculated with the Welch Two Sample Test.

Supplementary Figures

Figure S1. Relevant sequence region of C13*orf*25, including the CpG island harboring the host gene promoter, the A/T-rich region, and the miR-17-92 cluster; sequence and position of the last exon 4 of C13*orf*25 is indicated at the end. Shown sequences are based on the NCBI reference sequence NG_032702.1 and the GRCh37/hg19 assembly [25]. The boundaries of the CpG island, important previously identified regulatory elements, mature miRNA coding sequences and relevant primer sequences are highlighted in the sequence and annotated at the margins.

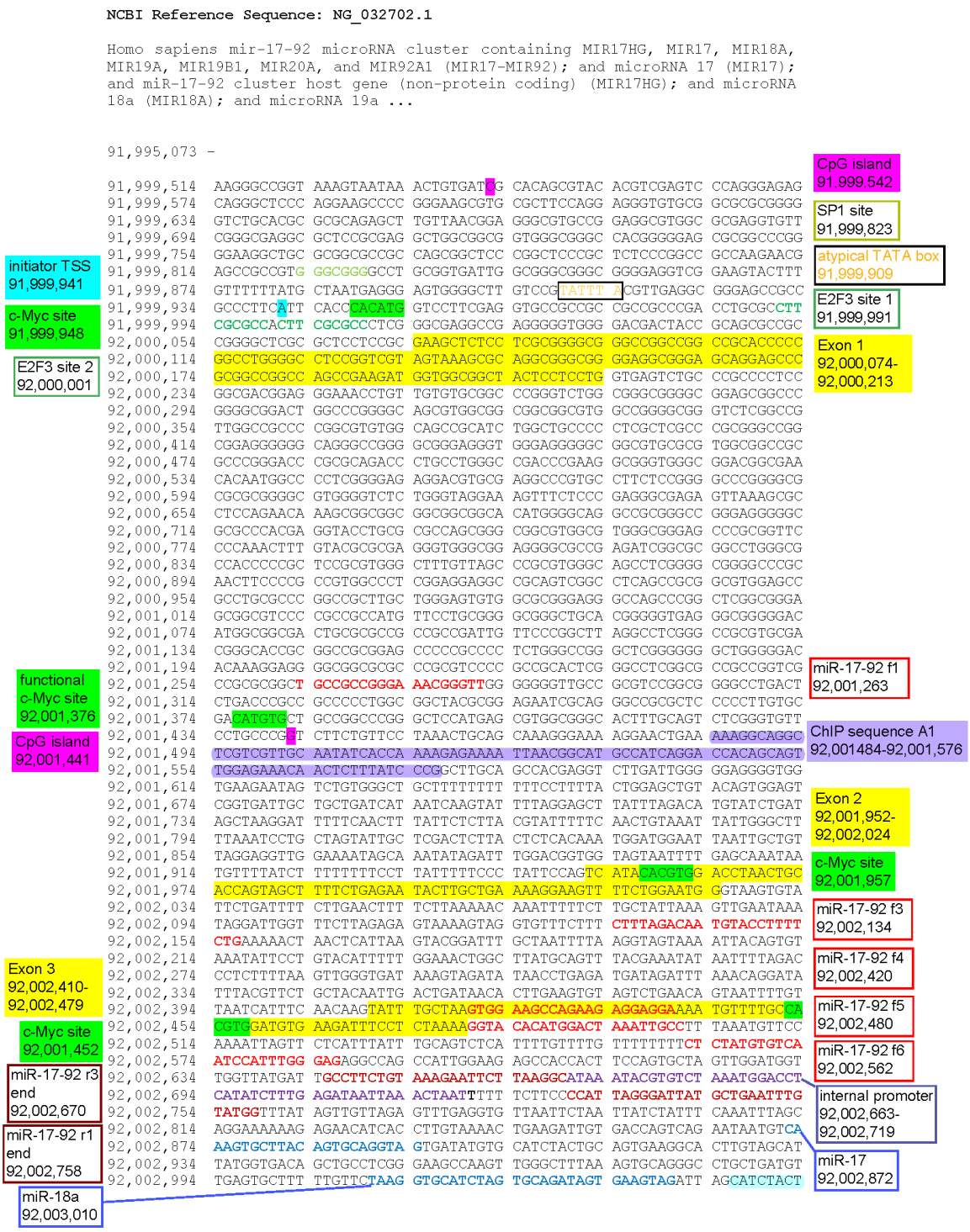


Figure S1. *Cont.*

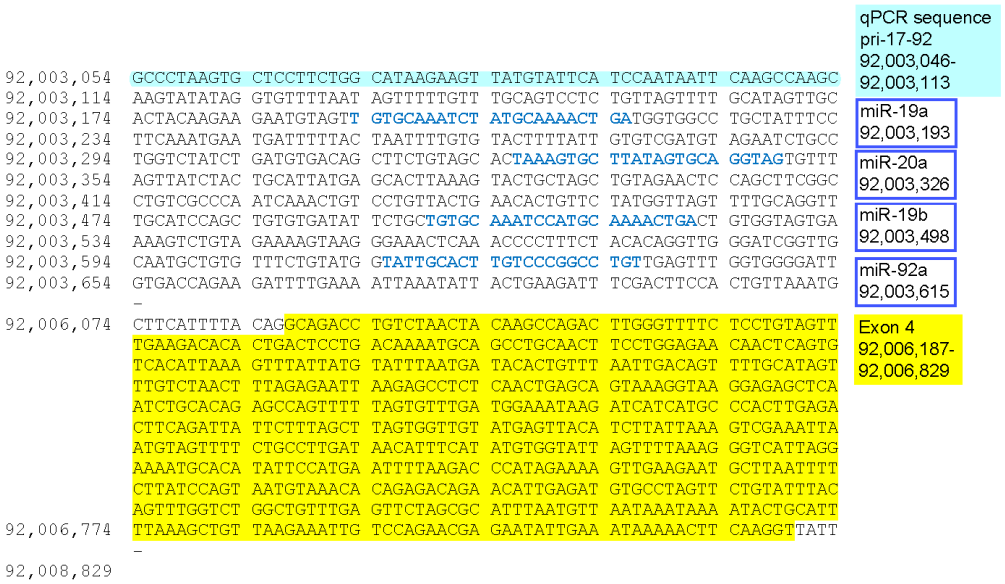


Figure S2. Quantitative RT-PCR of the pri-mir-17-92 transcription levels in the human cell lines K562, HeLa and HUH7 (hepato cellular carcinoma cells). 2^-ΔΔpri-mir-17-92 values are normalized against 5S rRNA and obtained from at least 3 independent experiments (+/−S.E.M.). The amount of pri-mir-17-92 transcript in K562 cells was set to 1.

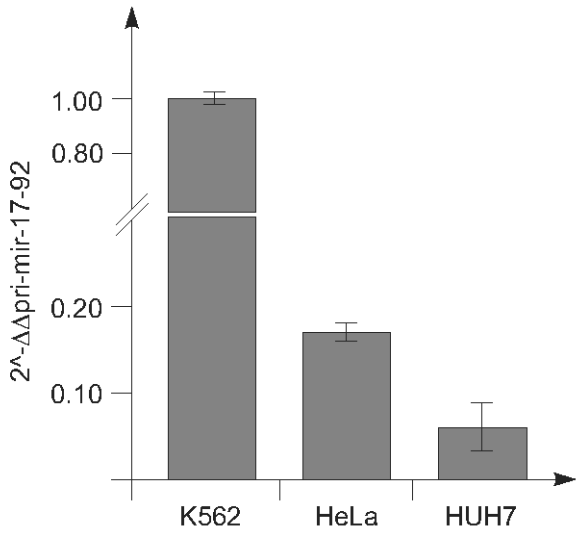


Figure S3. (**A**) Schematic representation of the intronic A/T-rich region preceding the miR-17-92 coding sequence. The region A1 defines the genomic sequence 0.1 kb downstream of the functional c-Myc binding site (E3 box) that was amplified in ChIP analyses. A2 covers a segment immediately upstream of the miRNA-coding region; A3–A5 are located along the coding sequence of the human miR-17-92 cluster. The length (bp) of each amplicon is indicated at the top; (**B**) ChIP analysis of the regions A1 to A5 in K562 cells, using an antibody specific for HP1γ or RNA polymerase II (only A2 analyzed). +AB, with antibody; −AB, without antibody; Mock, buffer only without cell lysate; Input, supernatant of the "−AB" sample after immunoprecipitation and centrifugation (for details, see Supplementary Material); (**C**) ChIP analysis of the A2 region in HeLa cells using the antibody specific for HP1γ.

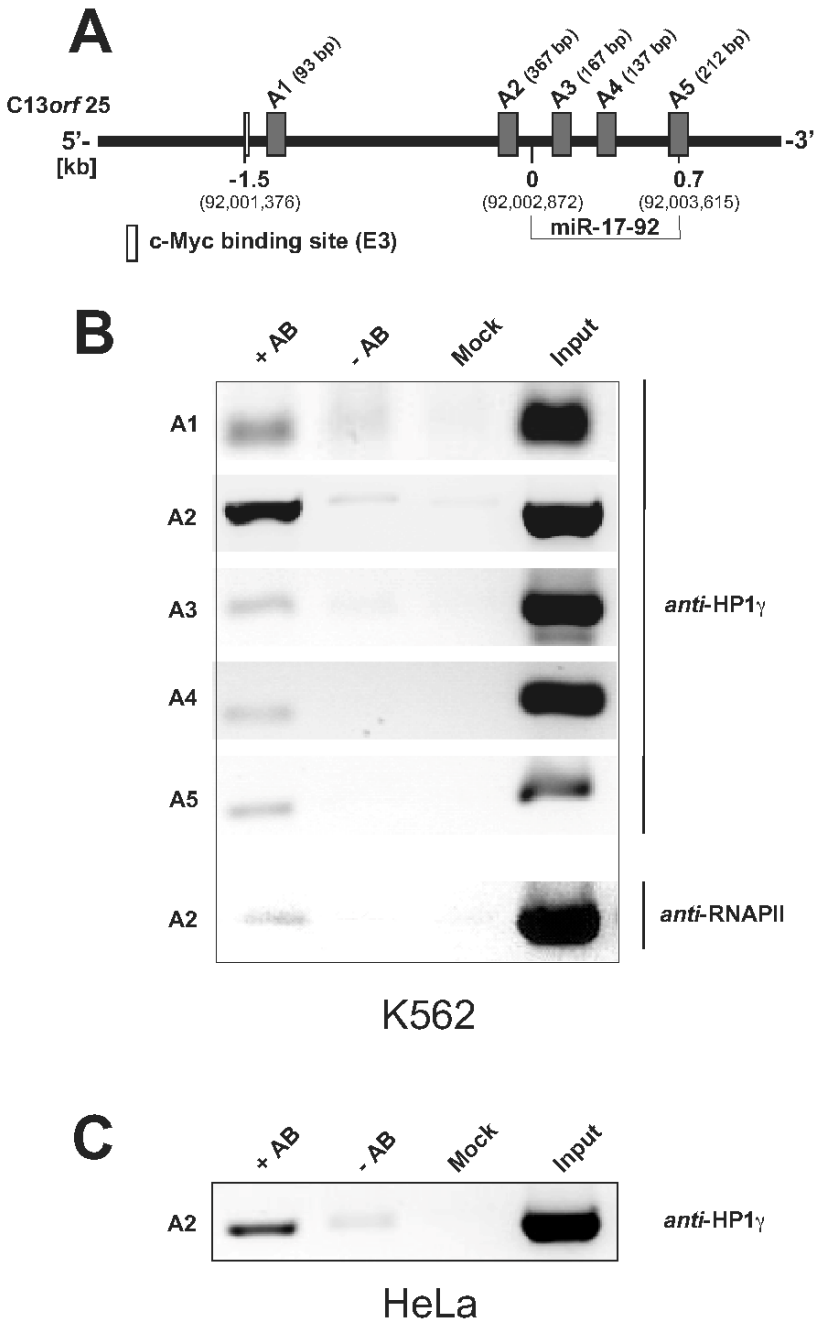
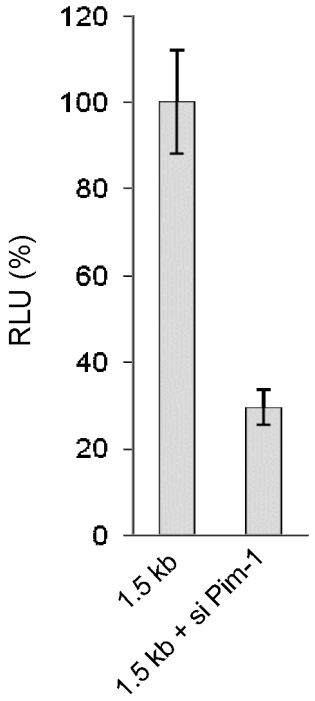


Figure S4. Effect of an siRNA-mediated Pim-1 knockdown on promoter activity of the ~1.5 kb reporter construct in HeLa cells. RLU values were derived from 5 independent triplicate experiments (+/−S.D.). RLU values for the control (left bar, transfected with the reporter plasmid but in the absence of a siRNA) were set to 100%. Lipofectamine transfection of HeLa cells was done as described under Supplementary methods, with the following alterations: 2×10^5 cells were used, and 40 pmol (0.6 µg) siRNA plus 0.5 µg of the reporter plasmid were combined in 50 µL Opti-MEM® I medium and mixed with 1.5 µL Lipofectamine™ 2000 in 50 µL Opti-MEM® I medium. The resulting mixture (~100 µL) was incubated for 20 min at room temperature to allow complex formation before addition to the cells. For the control (left bar), the siRNA was omitted.



Supplementary Table

Table S1. Quantification of the pri-miR-17-92 levels and the c-Myc, E2F3 and Pim-1 mRNA levels after siRNA-dependent knockdown by qRT-PCR. Expression levels were calculated from the crossing points by the 2^-ΔΔC_T method [38] using β-Actin mRNA or 5S rRNA as internal controls. To determine the c-Myc, E2F3 and Pim-1 knockdown efficiencies, expression levels were normalized to the levels obtained by transfection of K562- or HeLa cells with an unrelated siRNA directed against the vanilloid receptor (siVR1).

K562	c-Myc KD		E2F3 KD		Pim-1 KD	
	2^-ΔΔC_T pri-mir-17-92	KD efficiency	2^-ΔΔC_T pri-mir-17-92	KD efficiency	2^-ΔΔC_T pri-mir-17-92	KD efficiency
	0.13	77	0.21	95	1.23	72
	0.02	87	0.16	98	0.16	69
	0.56	59	0.60	19	1.39	55
	0.75	74			0.76	59
	0.33	75			0.22	49
	0.14	99			2.38	78
	0.20	97			1.84	87
	0.55	80			0.80	88
					0.27	85
					0.40	89
					0.52	67
					0.69	65
					0.37	87
Mean	**0.33**	**81**	**0.34**	**89**	**0.47**	**73**
SEM	0.09	5.0	0.15	8.0	0.08	5.0

K562	c-Myc / E2F3 KD		c-Myc / Pim-1 KD		Pim-1 / E2F3 KD	
	2^-ΔΔC_T pri-mir-17-92	KD efficiency	2^-ΔΔC_T pri-mir-17-92	KD efficiency	2^-ΔΔC_T pri-mir-17-92	KD efficiency
	0.22	82 / 75	0.45	72 / 55	0.69	88 / 86
	0.23	86 / 72	0.19	90 / 54	0.56	68 / 73
	0.15	81 / 48	0.19	95 / 48	0.19	95 / 99
	0.08	99 / 99			0.52	87 / 81
	0.04	99 / 99			0.23	91 / 93
Mean	**0.15**	**89 / 79**	**0.28**	**86 / 52**	**0.44**	**86 / 86**
SEM	0.04	4.0 / 9.6	0.08	7.0 / 2.2	0.09	4.7 / 4.5

HeLa	c-Myc KD		E2F3 KD		Pim-1 KD	
	2^-ΔΔC_T pri-mir-17-92	KD efficiency	2^-ΔΔC_T pri-mir-17-92	KD efficiency	2^-ΔΔC_T pri-mir-17-92	KD efficiency
	0.37	70	0.72	51	0.50	47
	0.29	82	0.66	78	0.34	72
	0.32	61	0.40	76	0.71	63
	0.41	99	0.53	73	0.58	80
	0.48	55	0.64	73	0.54	87
	0.37	64			0.78	72
					0.70	76
Mean	**0.37**	**65**	**0.59**	**70**	**0.59**	**71**
SEM	0.03	6.7	0.07	6.3	0.06	4.9

HeLa	c-Myc / E2F3 KD		c-Myc / Pim-1 KD		Pim-1 / E2F3 KD	
	2^-ΔΔC_T pri-mir-17-92	KD efficiency	2^-ΔΔC_T pri-mir-17-92	KD efficiency	2^-ΔΔC_T pri-mir-17-92	KD efficiency
	0.24	79 / 64	0.24	79 / 84	0.49	81 / 96
	0.13	91 / 81	0.25	77 / 61	0.33	91 / 85
	0.11	93 / 86	0.15	86 / 58	0.10	97 / 94
	0.12	90 / 91	0.12	95 / 91	0.26	92 / 94
	0.23	87 / 74	0.41	84 / 73		
	0.26	83 / 66	0.10	97 / 95		
Mean	**0.18**	**87 / 77**	**0.21**	**86 /77**	**0.30**	**90 / 92**
SEM	0.03	2.2 / 4.4	0.053	3.3 / 6.0	0.08	3.3 / 2.5

References

1. Iorio, M.V.; Croce, C.M. MicroRNAs in cancer: Small molecules with a huge impact. *J. Clin. Oncol.* **2009**, *27*, 5848–5856.
2. Krol, J.; Loedige, I.; Filipowicz, W. The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* **2010**, *11*, 597–610.
3. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297.
4. Denli, A.M.; Tops, B.B.J.; Plasterk, R.H.; Ketting, R.F.; Hannon, G.J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **2004**, *432*, 231–235.
5. Yi, R.; Qin, Y.; Macara, I.G.; Cullen, B.R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **2003**, *17*, 3011–3016.
6. Lee, Y.; Jeon, K.; Lee, J.-T.; Kim, S.; Kim, V.N. MicroRNA maturation: Stepwise processing and subcellular localization. *EMBO J.* **2002**, *21*, 4663–4670.
7. Esquela-Kerscher, A.; Slack, F.J. Oncomirs—microRNAs with a role in cancer. *Nat. Rev. Cancer* **2006**, *6*, 259–269.
8. Ota, A.; Tagawa, H.; Karnan, S.; Tsuzuki, S.; Karpas, A.; Kira, S.; Yoshida, Y.; Seto, M. Identification and characterization of a novel gene, C13orf25, as a target for 13q31–q32 amplification in malignant lymphoma. *Cancer Res.* **2004**, *64*, 3087–3095.
9. He, L.; Thomson, J.M.; Hemann, M.T.; Hernando-Monge, E.; Mu, D.; Goodson, S.; Powers, S.; Cordon-Cardo, C.; Lowe, S.W.; Hannon, G.J.; *et al*. A microRNA polycistron as a potential human oncogene. *Nature* **2005**, *435*, 828–833.
10. O'Donnell, K.A.; Wentzel, E.A.; Zeller, K.I.; Dang, C.V.; Mendell, J.T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **2005**, *435*, 839–843.
11. Tanzer, A.; Stadler, P.F. Molecular evolution of a microRNA cluster. *J. Mol. Biol.* **2004**, *339*, 327–335.
12. Lu, J.; Getz, G.; Miska, E.A.; Alvarez-Saavedra, E.; Lamb, J.; Peck, D.; Sweet-Cordero, A.; Ebert, B.L.; Mak, R.H.; Ferrando, A.A.; *et al*. MicroRNA expression profiles classify human cancers. *Nature* **2005**, *435*, 834–838.
13. Olive, V.; Bennett, M.J.; Walker, J.C.; Ma, C.; Jiang, I.; Cordon-Cardo, C.; Li, Q.-J.; Lowe, S.W.; Hannon, G.J.; He, L. miR-19 is a key oncogenic component of mir-17–92. *Genes Dev.* **2009**, *23*, 2839–2849.
14. Ventura, A.; Young, A.G.; Winslow, M.M.; Lintault, L.; Meissner, A.; Erkeland, S.J.; Newman, J.; Bronson, R.T.; Crowley, D.; Stone, J.R.; *et al*. Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. *Cell* **2008**, *132*, 875–886.
15. Mu, P.; Han, Y.-C.; Betel, D.; Yao, E.; Squatrito, M.; Ogrodowski, P.; de Stanchina, E.; D'Andrea, A.; Sander, C.; Ventura, A. Genetic dissection of the miR-17–92 cluster of microRNAs in Myc-induced B-cell lymphomas. *Genes Dev.* **2009**, *23*, 2806–2811.
16. Ozsolak, F.; Poling, L.L.; Wang, Z.; Liu, H.; Liu, X.S.; Roeder, R.G.; Zhang, X.; Song, J.S.; Fisher, D.E. Chromatin structure analyses identify miRNA promoters. *Genes Dev.* **2008**, *22*, 3172–3183.
17. Woods, K.; Thomson, J.M.; Hammond, S.M. Direct regulation of an oncogenic micro-RNA cluster by E2F transcription factors. *J. Biol. Chem.* **2007**, *282*, 2130–2134.

18. Sylvestre, Y.; de Guire, V.; Querido, E.; Mukhopadhyay, U.K.; Bourdeau, V.; Major, F.; Ferbeyre, G.; Chartrand, P. An E2F/miR-20a autoregulatory feedback loop. *J. Biol. Chem.* **2007**, *282*, 2135–2143.
19. Guenther, M.G.; Levine, S.S.; Boyer, L.A.; Jaenisch, R.; Young, R.A. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **2007**, *130*, 77–88.
20. Ji, M.; Rao, E.; Ramachandrareddy, H.; Shen, Y.; Jiang, C.; Chen, J.; Hu, Y.; Rizzino, A.; Chan, W.C.; Fu, K.; *et al*. The miR-17–92 microRNA cluster is regulated by multiple mechanisms in B-cell malignancies. *Am. J. Pathol.* **2011**, *179*, 1645–1656.
21. Pospisil, V.; Vargova, K.; Kokavec, J.; Rybarova, J.; Savvulidi, F.; Jonasova, A.; Necas, E.; Zavadil, J.; Laslo, P.; Stopka, T. Epigenetic silencing of the oncogenic miR-17–92 cluster during PU.1-directed macrophage differentiation. *EMBO J.* **2011**, *30*, 4450–4464.
22. Koike, N.; Maita, H.; Taira, T.; Ariga, H.; Iguchi-Ariga, S.M.M. Identification of heterochromatin protein 1 (HP1) as a phosphorylation target by Pim-1 kinase and the effect of phosphorylation on the transcriptional repression function of HP1. *FEBS Lett.* **2000**, *467*, 17–21.
23. Nawijn, M.C.; Alendar, A.; Berns, A. For better or for worse: The role of Pim oncogenes in tumorigenesis. *Nat. Rev. Cancer* **2011**, *11*, 23–34.
24. Zippo, A.; de Robertis, A.; Serafini, R.; Oliviero, S. PIM1-dependent phosphorylation of histone H3 at serine 10 is required for MYC-dependent transcriptional activation and oncogenic transformation. *Nat. Cell Biol.* **2007**, *9*, 932–944.
25. UCSC Genome Bioinformatics. Available online: http://genome.ucsc.edu (assessed on 1 February 2009).
26. Thomas, M.; Lange-Grünweller, K.; Weirauch, U.; Gutsch, D.; Aigner, A.; Grünweller, A.; Hartmann, R.K. The proto-oncogene Pim-1 is a target of miR-33a. *Oncogene* **2012**, *31*, 918–928.
27. Venturini, L.; Battmer, K.; Castoldi, M.; Schultheis, B.; Hochhaus, A.; Muckenthaler, M.U.; Ganser, A.; Eder, M.; Scherr, M. Expression of the miR-17–92 polycistron in chronic myeloid leukemia (CML) CD34+ cells. *Blood* **2007**, *109*, 4399–4405.
28. Kwon, S.H.; Workman, J.L. HP1c casts light on dark matter. *Cell Cycle* **2011**, *15*, 625–630.
29. Vakoc, C.R.; Mandat, S.A.; Olenchock, B.A.; Blobel, G.A. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol. Cell* **2005**, *19*, 381–391.
30. Minc, E.; Allory, Y.; Worman, H.J.; Courvalin, J.; Buendia, B. Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells. *Chromosoma* **1999**, *108*, 220–234.
31. Mateescu, B.; Bourachot, B.; Rachez, C.; Ogryzko, V.; Muchardt, C. Regulation of an inducible promoter by an HP1beta-HP1gamma switch. *EMBO Rep.* **2008**, *9*, 267–272.
32. Fischle, W.; Tseng, B.S.; Dormann, H.L.; Ueberheide, B.M.; Garcia, B.A.; Shabanowitz, J.; Hunt, D.F.; Funabiki, H.; Allis, C.D. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **2005**, *438*, 1116–1122.
33. Adams, M.R.; Sears, R.; Nuckolls, F.; Leone, G.; Nevins, J.R. Complex transcriptional regulatory mechanisms control expression of the E2F3 locus. *Mol. Cell Biol.* **2000**, *20*, 3633–3639.
34. Thalmeier, K.; Synovzik, H.; Mertz, R.; Winnacker, E.L.; Lipp, M. Nuclear factor E2F mediates basic transcription and trans-activation by E1a of the human MYC promoter. *Genes Dev.* **1989**, *3*, 527–536.

35. Wong, J.V.; Dong, P.; Nevins, J.R.; Mathey-Prevot, B.; You, L. Network calisthenics: Control of E2F dynamics in cell cycle entry. *Cell Cycle* **2011**, *15*, 3086–3094.
36. Grünweller, A.; Wyszko, E.; Bieber, B.; Jahnel, R.; Erdmann, V.A.; Kurreck, J. Comparison of different antisense strategies in mammalian cells using locked nucleic acids, 2'-*O*-methyl RNA, phosphorothioates and small interfering RNA. *Nucleic Acids Res.* **2003**, *31*, 3185–3193.
37. Gozani Lab. Available online: http://www.stanford.edu/group/gozani (assessed on 1 June 2005).
38. Schmittgen, T.D.; Livak, K.J. Analyzing real-time PCR data by the comparative CT method. *Nat. Prot.* **2008**, *3*, 1101–1108.
39. Reese, M.G. Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Comput. Chem.* **2001**, *26*, 51–56.
40. Ohler, U. Identification of core promoter modules in *Drosophila* and their application in accurate transcription start site prediction. *Nucleic Acids Res.* **2006**, *34*, 5943–5950.
41. Knudsen, S. Promoter2.0: For the recognition of PolII promoter sequences. *Bioinformatics* **1999**, *15*, 356–361.
42. Rangannan, V.; Bansal, M. Relative stability of DNA as a generic criterion for promoter prediction: whole genome annotation of microbial genomes with varying nucleotide base composition. *Mol. BioSys.* **2009**, *5*, 1758–1769.
43. Zhang, M.Q. Identification of human gene core promoters *in silico*. *Genome Res.* **1998**, *8*, 319–326.
44. The R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2012.

Reprinted from *IJMS*. Cite as: Nagata, Y.; Shimizu, E.; Hibio, N.; Ui-Tei, K. Fluctuation of Global Gene Expression by Endogenous miRNA Response to the Introduction of an Exogenous miRNA. *Int. J. Mol. Sci.* **2013**, *14*, 11171-11189.

Article

Fluctuation of Global Gene Expression by Endogenous miRNA Response to the Introduction of an Exogenous miRNA

Yoshiro Nagata [1,†], Eigo Shimizu [2,†], Naoki Hibio [1] and Kumiko Ui-Tei [1,2,*]

[1] Department of Computational Biology, Graduate School of Frontier Sciences, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa-shi, Chiba-ken 277-8561, Japan;
E-Mails: a5205075@gmail.com (Y.N.); lapcat.nh@gmail.com (N.H.)

[2] Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan; E-Mail: eigo0shimizu@gmail.com

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: ktei@bi.s.u-tokyo.ac.jp; Tel.: +81-3-5841-3043; Fax: +81-3-5841-3044.

Received: 21 March 2013; in revised form: 15 April 2013 / Accepted: 6 May 2013 / Published: 27 May 2013

Abstract: Most of the intracellular endogenous microRNAs (endo-miRNAs) are considered to be saturated in Argonaute (Ago) proteins in the RNA-induced silencing complexes (RISCs). When exogenous miRNAs (exo-miRNAs) are introduced into cells, endo-miRNAs in the RISC may be replaced with exo-miRNAs or exo-miRNAs, and endo-miRNAs might also compete for the position in the newly synthesized RISC with each other. This would lead to the fluctuation of global gene expression not only by repression of exo-miRNA target gene expression, but also by the increase of the endo-miRNA target gene expression. In the present study, we quantified the changes in the expression levels of target genes of exo-miRNA and endo-miRNA in the cells transfected with fifteen different exo-miRNAs by microarray experiments. Different exo-miRNAs increased ratios of expression levels of target genes of a given endo-miRNA to different extents, suggesting that the replacement efficiencies might differ according to the exo-miRNA types. However, the increased ratios in the expression levels of each endo-miRNA target genes by the transfection of any particular exo-miRNA were mostly equivalent, suggesting that the endo-miRNAs present in the RISC might be replaced with excessive exo-miRNAs at similar levels, probably because they exist in single-stranded

forms in the RISC.

Keywords: exogenous microRNA; endogenous microRNA; microarray; seed-matched target

1. Introduction

MicroRNAs (miRNAs) are an abundant class of non-coding RNAs, about ~22 nucleotides long, that are key posttranscriptional regulators of gene expression in various organisms, including animals, plants and protozoa [1–3]. Most endogenous miRNAs (endo-miRNAs) actively silence target genes mainly by contiguous and perfect Watson-Crick base-pairing between the miRNA 5'-proximal seed region (positions 2–8) and its complementary sequences in 3' untranslated regions (3' UTRs) of target genes [4–6]. However, transfection of the small interfering RNA (siRNA)/miRNA expression construct into cells relieves repression of the target genes of endo-miRNAs dose-dependently at low concentrations and reaches the saturation level at high concentration [7–10].

These interfering effects are shown to be caused by competition for RNA silencing components. One of the key components is the nuclear karyopherin Exportin-5 [11–13], which binds to siRNA/miRNA precursors to transport them from the nucleus to the cytoplasm in the presence of Ran-GTP. The overloading of Exportin-5 by the excessive production of hairpin-structured siRNAs/miRNAs transcribed from their expression constructs could result in a decrease of cellular miRNA function [7,14]. However, the silencing activities of siRNA/miRNA duplexes are not affected by Exportin-5, since they do not need to be transported from the nucleus to the cytoplasm for their function [8,10,11]. Nonetheless, siRNA/miRNA duplexes have been shown to cause upregulation of their non-target genes by competition with endo-miRNAs. They compete for another saturable component of RNA silencing machinery, the RNA-induced silencing complex (RISC) [8,10,15]. The displacement of endo-miRNAs from the RISC by the introduction of synthetic siRNAs/miRNA duplexes are shown to be observed as the increase of endo-miRNA target gene expression [10].

RISC is the cytoplasmic effector machine of the miRNA silencing pathway. RISC assembly is mediated by the RISC loading complex, which is a multi-protein complex composed of the core protein, Argonaute (Ago), the RNase Dicer and the double-stranded RNA-binding protein, TRBP (TAR RNA binding protein) [16–19]. Initially, miRNA duplexes and siRNAs are loaded into Ago protein contained in the RISC loading complex [20,21]. To form the active RISC that performs gene silencing, the small RNA duplex needs to be separated or unwound into the single stranded form guiding it to its target mRNAs, within Ago protein. Then Dicer and its interactor TRBP dissociate from the RISC. Generally, miRNA forms an imperfect duplex composed of a miRNA strand and an opposite-strand miRNA. Evolutionary pressure has selected one particular strand of the duplex as the main regulator, which is preferentially loaded onto RISC, with the opposite strand being less functional [22,23]. The strand choice is considered to be not random and is partly determined by the intrinsic sequence and/or mismatched base-pairing of the miRNA duplex. The major determinants of RISC loading have been shown to be the thermodynamic properties: the strand with the less stable 5' end is more often loaded onto active RISC [24–26]. Furthermore, central mismatches of miRNAs have also been shown to promote RISC loading [27]. In the cells, RISC is considered to be saturated with

endo-miRNAs. Because these experiments were carried out by the transfection of exogenous miRNAs (exo-miRNAs), exo-miRNAs are anticipated to exclude the endo-miRNAs from RISC or compete with endo-miRNAs for RISC with each other.

In this study, to investigate the mechanism to perturb the endo-miRNA function caused by the introduction of exo-miRNA, we performed microarray profiling to quantify changes in the expression levels of endo-miRNA target genes following the transfection of fifteen different synthetic exo-miRNA duplexes. Such competition among cellular miRNAs is considered to have a role in normal biological and disease-related cellular processes. The exo-miRNAs, in addition to silencing their own target genes, clearly increased the expression of endo-miRNA target genes. The increased levels of endo-miRNA target gene expression were varied according to the types of the transfected exo-miRNA duplexes, suggesting each exo-miRNA duplex is presumed to have distinct characteristics, such as structures and sequences, which affect its incorporation into the RISC. However, the increased levels were mostly equivalent according to the types of endo-miRNAs after transfection of any particular exo-miRNA duplex. These results might propose the possibility that the endo-miRNAs present in the RISC in single-stranded form, but not in duplex form, nor in target-pairing form, were replaced with exo-miRNA duplexes, because double-stranded endo-miRNAs with different characteristics might not be replaced with exo-miRNAs in the similar levels. A part of exo-miRNAs transfected might compete with endo-miRNAs for newly synthesized Ago proteins, but such effects might be negligible, because the transfected exo-miRNAs are sufficiently abundant. Furthermore, the endo-miRNA target genes with a large number of target sites in their 3' UTRs are revealed to be strongly repressed by endo-miRNAs in normal conditions. So, the genes with short 3' UTRs, which have a small number of miRNA target sites, were weakly repressed by endo-miRNAs in the normal condition, then efficiently repressed when exo-miRNAs were transfected.

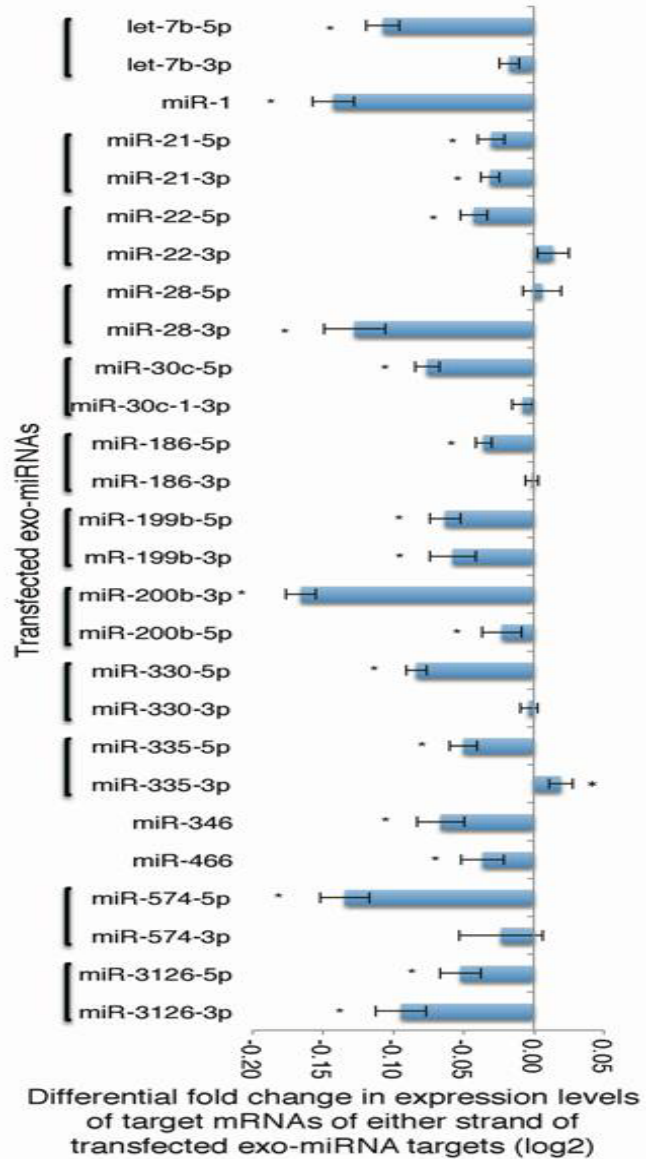
2. Results and Discussion

2.1. Microarray Profiling of the Reduced Expression of Exo-miRNA Target Genes

To quantify global changes in the expression levels of endo-miRNA target genes resulting from the transfection of various exo-miRNA duplexes, microarray experiments were carried out using fifteen different exo-miRNA duplexes (let-7b, miR-1, miR-21, miR-22, miR-28, miR-30c-1, miR-186, miR-199b, miR-200b, miR-330, miR-335, miR-346, miR-466, miR-574 and miR-3126), which were chemically synthesized to form the duplex structures the same as those shown in miRBase [28]. At first, 24 h following the transfection of each exo-miRNA duplexes into human HeLa cells, the effects on own target genes of either of both miRNA strand were examined. We chose this time point, because it is reported that RNA silencing is generally maximal ~24 h post-transfection and that protein silencing varies depending on the target, but is generally maximal ~48–72 h post-transfection [29]. Thus, the result at 24 h post-transfection may largely reflect the direct effects without downstream effects of exo-miRNAs, since major downstream effects should be observed after 48 h post-transfection. It has been reported that miRNA target genes mainly contain sequences complementary to miRNA seed regions at positions 2–8 from the 5' terminus [4–6], then the expression patterns of these genes were analyzed. The expression patterns of each miRNA target genes

containing seed-complementary sequences in HeLa cells transfected with fifteen exo-miRNAs are shown as MA (M = intensity ratio, A = average intensity) plots (Figure S1A-1~O-1 and S1A-3~O-3) and cumulative distributions (Figure S1A-2~O-2 and S1A-4~O-4). The changes in expression were calculated as the difference values subtracting the average fold changes (log_2) for exo-miRNA seed-matched target genes (blue line in Figure S1A-2~O-2 and S1A-4~O-4) from the average value for the genes without seed-matched sequence (black line in Figure S1A-2~O-2 and S1A-4~O-4). The differential fold change values were between 0.02 and −0.17, as summarized in Figure 1, indicating that all of the exo-miRNAs used in this study could decrease their own target genes at the different levels, due to the miRNA species. However, these decreases in exo-miRNA target gene expression, as a result of competition, with the corresponding endo-miRNAs might be underestimated in this experiment.

Figure 1. Microarray analysis of exo-miRNA target gene expression. HeLa cells transfected with each of exo-miRNA duplexes were subjected to microarray profiling. Differential fold changes in gene expression caused by the transfection of either strand of fifteen different exo-miRNAs. The results of the opposite strands of miR-1, miR-346 and miR-466 were not shown, because these miRNAs were not registered in miRBase. The MA plots and cumulative distribution patterns are shown in Figure S1.



2.2. Microarray Profiling of the Increased Expression of Endo-miRNA Target Genes Resulting from the Introduction of Exo-miRNAs

Having found reduced expression of exo-miRNA target genes, we next analyzed the expression of endo-miRNA target genes. As let-7b, miR-21, miR-27a, miR-17, miR-26a, miR-24, miR-30a, miR-92a, miR-19a, miR-15a, miR-22, miR-29a, miR-125a, miR-93, miR-191, miR-103a, miR-143, miR-100, miR-23a and miR-186 are reported to be the top 20 most highly expressed miRNA families in HeLa cells [30], changes in the target gene expression of these representative miRNAs were calculated (Figures 2 and 3). In this analysis, "endo-miRNA target genes" were defined as genes with sequences complementary to endo-miRNA seed sequences (positions 2–8) in their 3' UTRs, but not to both strands of exo-miRNA seed sequences. The "endo-miRNA non-target genes" were genes with no sequences complementary to endo-miRNA or both strands of exo-miRNA seed sequences. The MA plots and cumulative distributions were shown in Figures S2–S21. The difference values were calculated by subtracting the average fold changes ($\log_2$) for endo-miRNA target genes (blue line in Figures S2–S21 A~O-2) from the average value for endo-miRNA non-target genes (black line in Figures S2–S21 A-2~O-2). The expression of most of the endo-miRNA target genes was increased by the transfection of exo-miRNAs (Figures 2 and 3, Figures S2–S21).

Next, we compared changes in the expression of endo-miRNA target genes according to the types of exo-miRNAs (Figure 2). Expression changes were determined by subtracting the average $\log_2$ fold change for endo-miRNA target genes containing more than one target site(s) from the average value for endo-miRNA non-target genes and shown as a differential fold change ($\log_2$). The increases in target gene expression of endogenous top 20 miRNAs varied, but showed similar tendencies according to the kind of exo-miRNA transfected (Figure 2). Of the fifteen transfected exo-miRNAs, the exogenously transfected miR-466 duplex increased the expression levels of target genes of at least 16 endo-miRNAs (let-7b-5p, miR-21-5p, miR-27a-3p, miR-17-5p, miR-26a-5p, miR-24-3p, miR-30a-5p, miR-92a-3p, miR-19a-3p, miR-15a-5p, miR-22a-3p, miR-29a-3p, miR-93a-5p, miR-143-3p, miR-23a-3p and miR-186-5p) significantly (Figure 2A–T), and their averaged expression level was the highest (Figure 2U). The exo-miRNA duplexes, miR-21, miR-22, miR-28, miR-30c-1, miR-200b, miR-346, miR574 and miR-3126, produced significant, but modest levels of increases in the endo-miRNA target genes (Figure 2U), while exo-miRNA duplexes, let-7b, miR-1, miR-28, miR-199b and miR-335, showed low values on average (Figure 2U).

Considering another dimension of the results, changes in the target gene expression of top 20 endo-miRNAs were examined. The results clearly show that the increased ratios of target genes of any endo-miRNAs were roughly equivalent when a particular exo-miRNA was transfected (Figure 3). The average expression of seed-matched target genes of these 20 endo-miRNAs for each of 15 transfected exo-miRNA are shown in Figure 3. These results indicated that the increased fold changes of expression levels of endo-miRNA target genes were essentially determined by the kind of exo-miRNA and not by the type of endo-miRNA, although the increase of target genes of a few types of endo-miRNAs, such as miR-1 (Figure 3B), miR-28 (Figure 3E), miR-199b (Figure 3H) and miR-335 (Figure 3K), by the transfection of a given miRNA were not equivalent, and averaged values after transfection of these exo-miRNAs became low due to combining the up- and down-regulated targets (Figure 2U). The similar results, indicating the equivalent expression levels of endo-miRNA targets by most of exo-miRNAs, except for miR-1, miR-28, miR-199b and miR-335, were also observed when the increased levels were calculated for top 383 endo-miRNAs (Figure S22).

Figure 2. Mean fold changes of the expression levels of endo-miRNA target transcripts by the transfection of different exo-miRNAs. HeLa cells were transfected with each of fifteen exo-miRNA duplexes, $\log_2$ mean differential fold changes of the expression levels of seed-matched target genes of each of the top 20 endo-miRNAs were calculated. Data were shown with respect to endo-miRNAs, let-7b-5p (**A**); miR-21 (**B**); mR-27a-3p (**C**); miR-17-5p (**D**); miR-26a-5p (**E**); miR-24-3p (**F**); miR-30a-5p (**G**); miR-92a-5p (**H**); miR-19a-5p (**I**); miR-15a-5p (**J**); miR-22-3p (**K**); miR-29a-3p (**L**); miR-125a-5p (**M**); miR-93-5p (**N**); miR-191-5p (**O**); miR-103a-3p (**P**); miR-143-3p (**Q**); miR-100-5p (**R**); miR-23a-3p (**S**); and mR-186-5p (**T**); and their averaged values were shown in (**U**). Note that each exo-miRNA increased the endo-miRNA target genes in different degrees. The individual data are shown in Figures S2–S21. Data represent the mean ± SE (* $p < 0.05$).

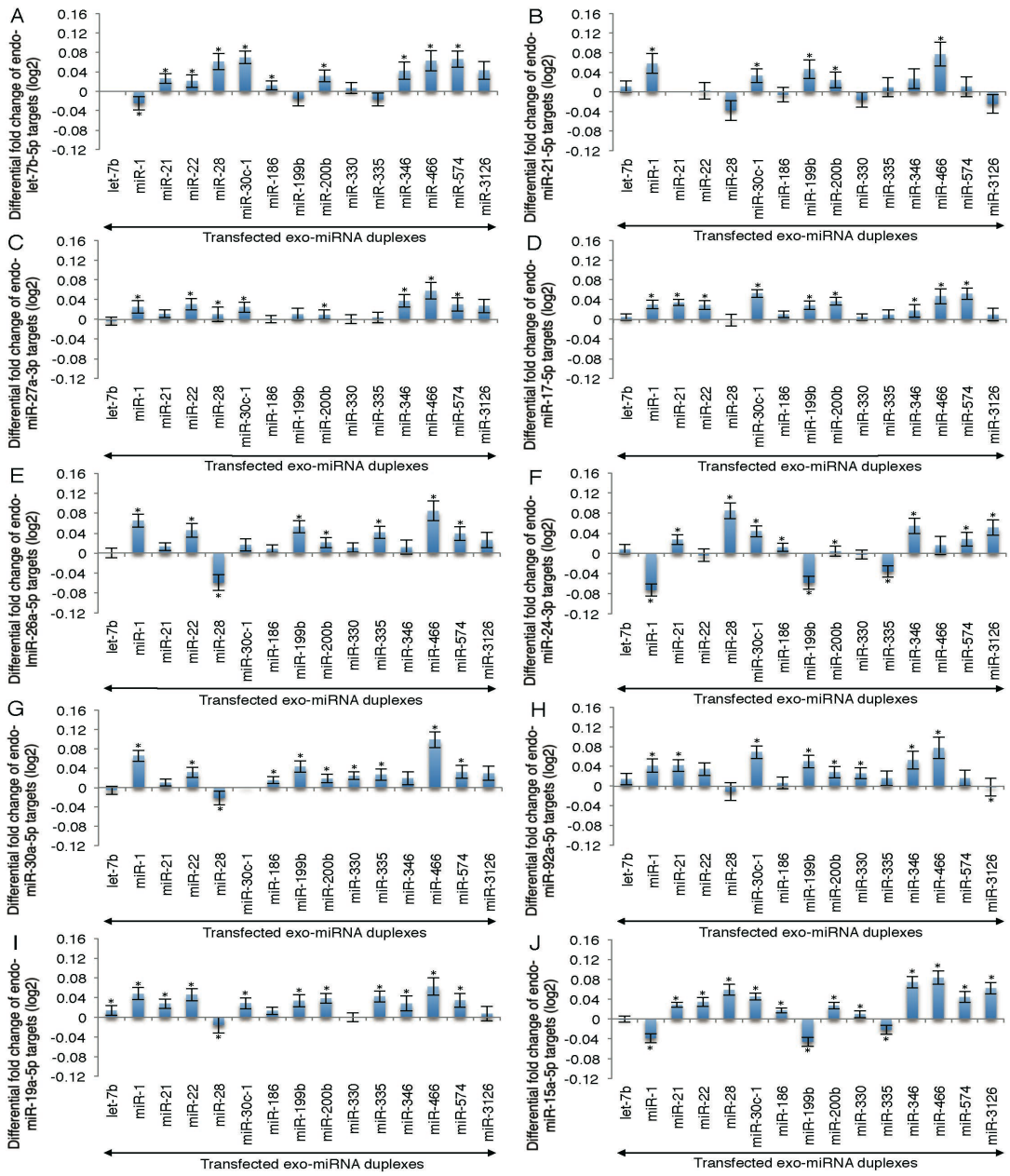


Figure 2. *Cont.*

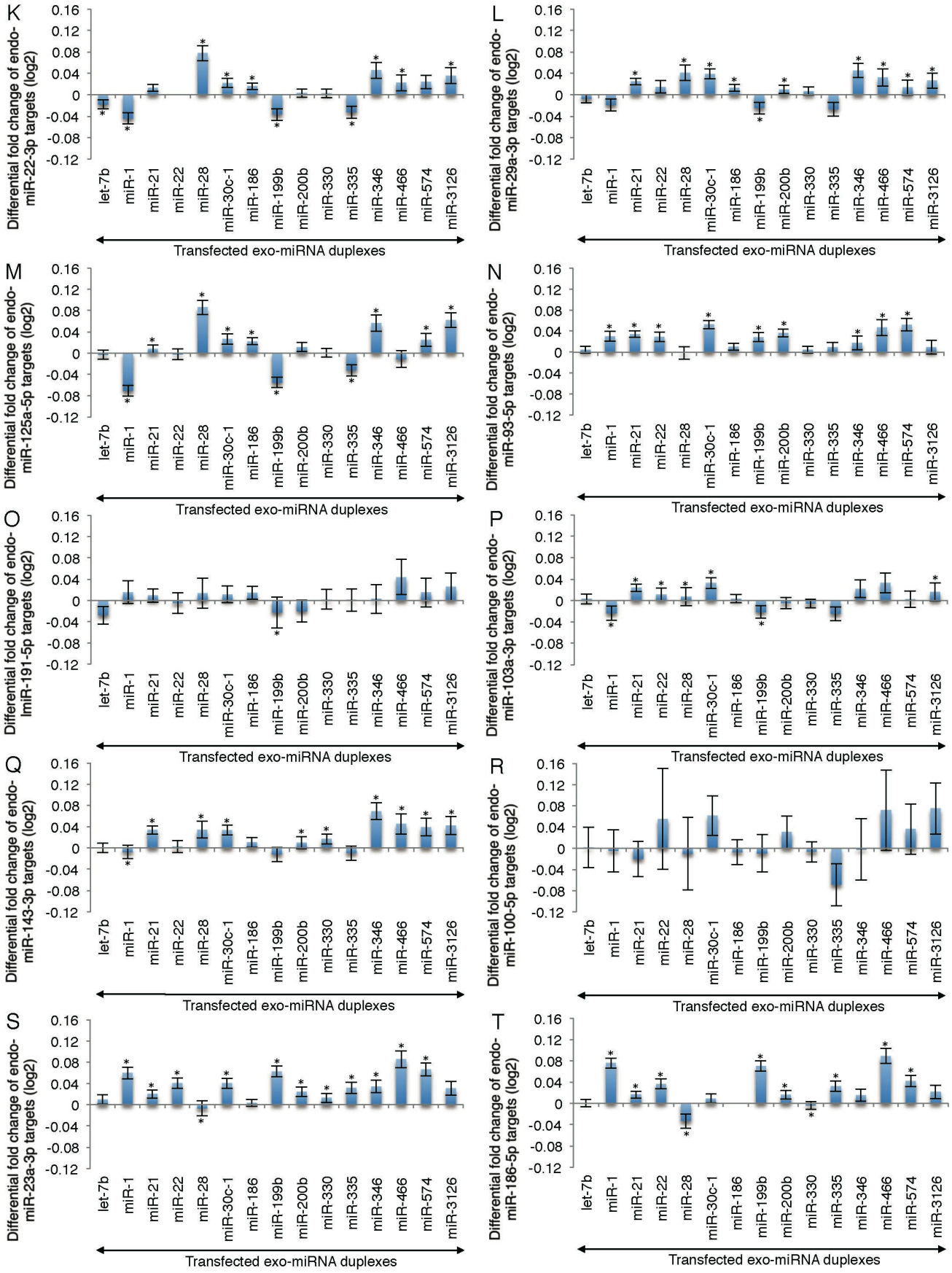

K
Differential fold change of endo-miR-22-3p targets (log2)
L
Differential fold change of endo-miR-29a-3p targets (log2)
M
Differential fold change of endo-miR-125a-5p targets (log2)
N
Differential fold change of endo-miR-93-5p targets (log2)
O
Differential fold change of endo-lmiR-191-5p targets (log2)
P
Differential fold change of endo-miR-103a-3p targets (log2)
Q
Differential fold change of endo-miR-143-3p targets (log2)
R
Differential fold change of endo-miR-100-5p targets (log2)
S
Differential fold change of endo-miR-23a-3p targets (log2)
T
Differential fold change of endo-miR-186-5p targets (log2)
0.16
0.12
0.08
0.04
0
-0.04
-0.08
-0.12
let-7b
miR-1
miR-21
miR-22
miR-28
miR-30c-1
miR-186
miR-199b
miR-200b
miR-330
miR-335
miR-346
miR-466
miR-574
miR-3126
Transfected exo-miRNA duplexes

Figure 2. *Cont.*

Our results were obtained from our own microarray experiments using one human cell line; it should reduce variables compared to the previous studies. To confirm the reliability of the microarray data, we analyzed 28 transcripts by quantitative RT-PCR (Figure S23). The expression levels estimated by quantitative RT-PCR were essentially identical to those obtained in the microarray analysis, with an estimated correlation coefficient of 0.83.

Figure 3. Mean fold changes of the expression levels of endo-miRNA target transcripts following the transfection of exo-miRNAs. HeLa cells were transfected with each of fifteen exo-miRNA duplexes, $\log_2$ mean differential fold changes of the expression levels of seed-matched target genes of each of the top 20 endo-miRNAs were calculated. Data were shown with respect to transfected exo-miRNAs, miR-7b (**A**); miR-1 (**B**); miR-21 (**C**); miR-22 (**D**); miR-28 (**E**); miR-30c-1 (**F**); miR-186 (**G**); miR-199b (**H**); miR-200b (**I**); miR-330 (**J**); miR-335 (**K**); miR-346 (**L**); miR-466 (**M**); miR-574 (**N**); miR-3126 (**O**); and their averaged values were shown in (**P**). Target genes of different endo-miRNAs are increased by exo-miRNAs in approximately comparable levels. The individual data are shown in Figures S2–S21. Data represent the mean ± SE (* $p < 0.05$).

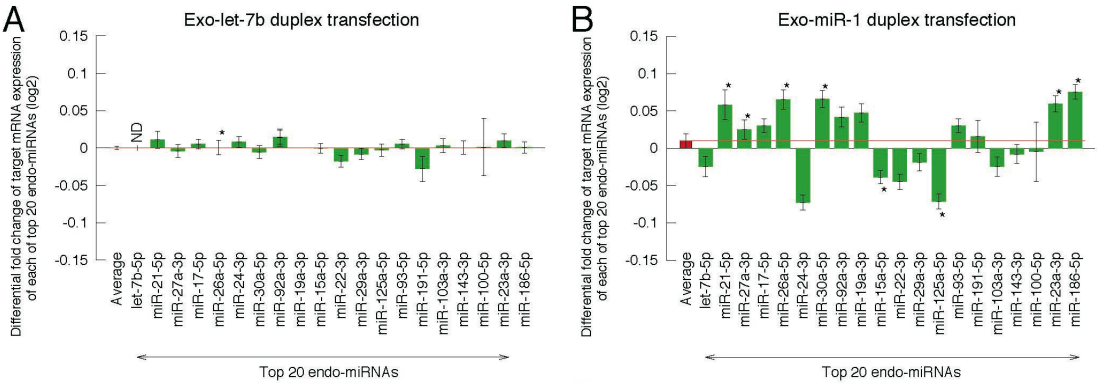


Figure 3. *Cont.*

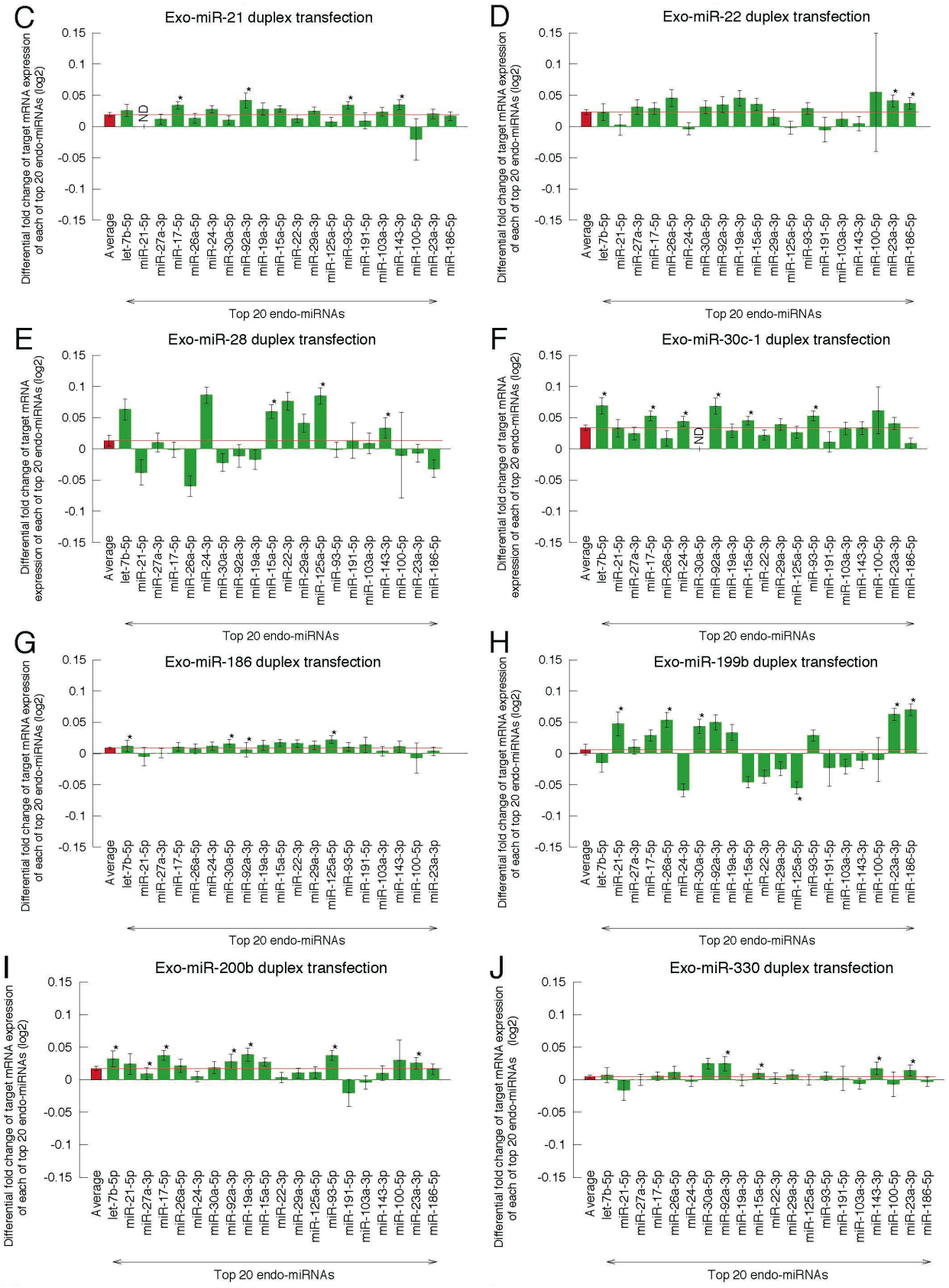
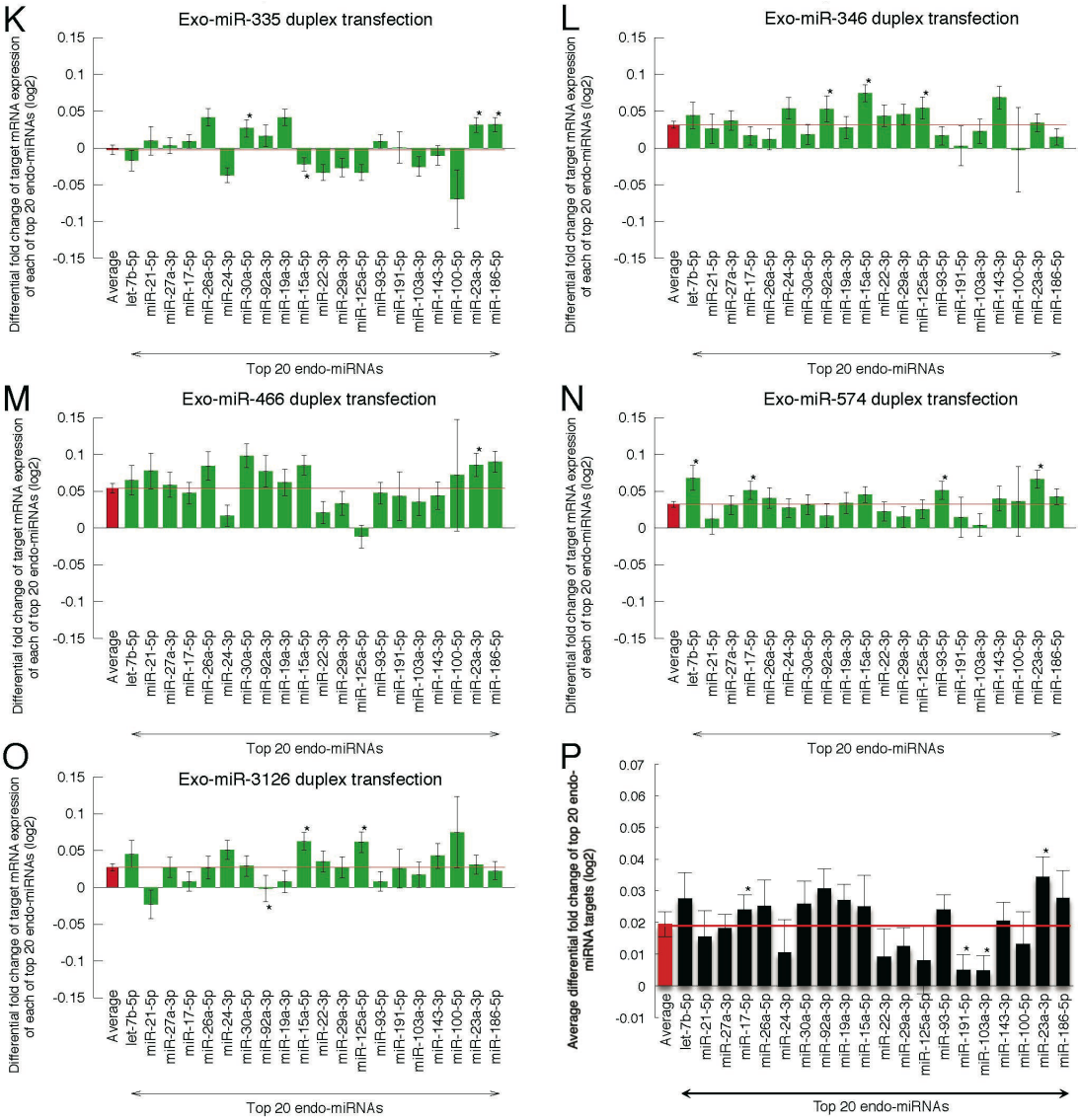


Figure 3. *Cont.*



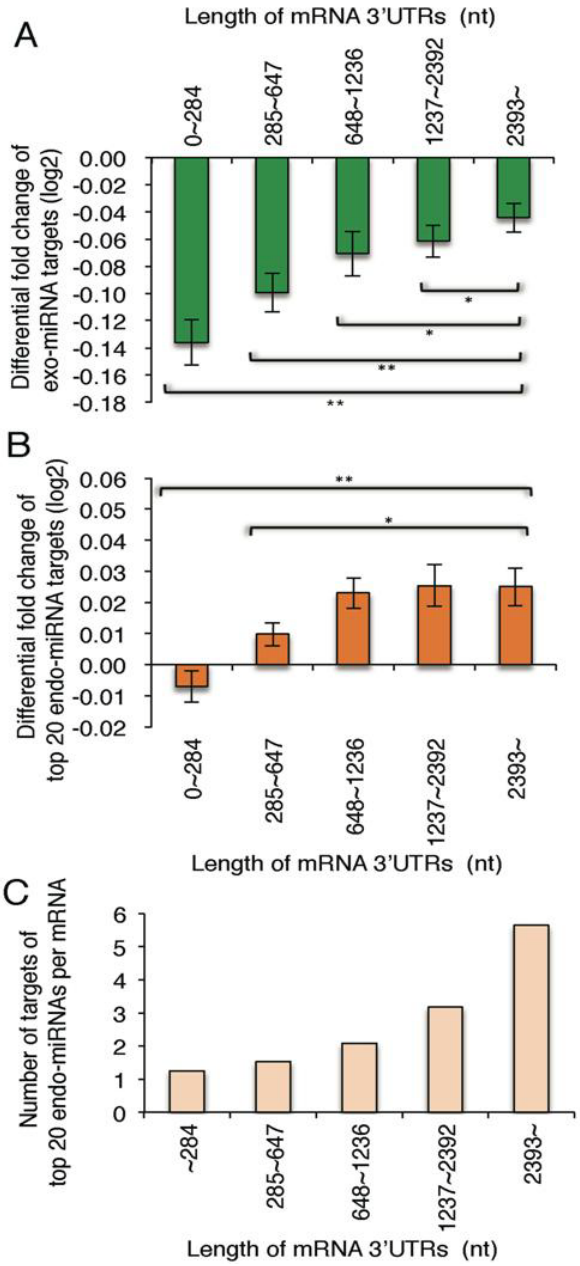
2.3. Reporter Analysis of Exo-miRNA and Endo-miRNA Target Expression in Cells Transfected with Exo-miRNA

A part of the target genes with 3' UTRs complementary to endo-miRNA seed sequences is not necessarily downregulated by endo-miRNAs in normal conditions. The accessibilities of endo-miRNAs to such genes are speculated to be interfered by some causes, such as binding of specific RNA binding proteins or RNA secondary structures. Thus, the changes in the expression levels of seed-complementary target genes of endo-miRNAs shown in Figures 2 and 3 might be modest or low (generally ~20 to 10%). To confirm the downregulation of exo-miRNA target gene expression and upregulation of endo-miRNA target gene expression resulting from transfection of exo-miRNA, we carried out reporter assays using luciferase expression constructs carrying sequences perfectly complementary to exo-miRNAs or endo-miRNAs in the 3' UTR of the *Renilla* luciferase gene in psiCHECK-1 (Figure 4A). Even when the changes in the expression levels of seed-complementary endogenous target genes of exo-miRNAs or endo-miRNAs are small, the expression levels of luciferase reporter

with perfect complementary sequences of miRNAs are expected to show remarkable effects, although the effect is not biologically relevant. Exo-miR-200b or exo-miR-330 was transfected into HeLa cells with each reporter construct and a firefly luciferase expression construct (pGL3-Cont) (internal control). A double-stranded DNA (miDNA), which mimics the miRNA structure, was used as a miRNA control. At 24 h post-transfection, the relative luciferase activity was calculated (Figure 4). Both exo-miRNAs silenced their own targets (Figure 4B,E). In contrast, the luciferase activities of psiCHECK-1 constructs containing the endo-miRNA target sequences apparently increased (Figure 4C,D,F,G). These results suggest that exo-miRNAs interfered with endo-miRNA silencing activity by competing for the RNA silencing machinery downstream of Exportin-5 probably replacing with endo-miRNAs in the RISC, thereby leading to increased expression of endo-miRNA target genes.

Figure 4. Luciferase reporter analyses of changes in the expression of exo-miRNA and endo-miRNA targets. (**A**) Schematic structure of a luciferase reporter containing a miRNA completely matched target sequence in the 3' UTR and mRNA transcribed from the reporter. The exo-miRNAs, miR-200b (**B**–**D**) and miR-330 (**E**–**G**), respectively, were transfected into HeLa cells at 50 nM. At 24 h post-transfection, luciferase activities were measured. The luciferase activities derived from the constructs carrying exo-miRNA target sequences were decreased by the transfection of respective exo-miRNA (**B**,**E**). In contrast, the luciferase activities from the constructs carrying endo-let-7b-5p (**C**), endo-miR-21-5p (**D**,**G**) and endo-miR-7b-5p (**F**) target sequences were increased, indicating that exo-miRNAs repressed endo-miRNA silencing activities.

Figure 5. Mean fold changes of target genes of exo-miRNAs and endo-miRNAs according to the 3' UTR lengths. Differential fold changes (log_2) of expression levels of exo-miRNA targets (**A**); and endo-miRNA targets (**B**); and the number of top 20 endo-miRNA target sites (**C**) according to the lengths of 3' UTRs. The averaged numbers of exo-miRNA target sites are 53, 137, 210, 250 and 149 (**A**), and those of endo-miRNA target sites are 53, 134, 185, 184 and 84 (**B**) in the 3' UTRs of 0–284, 285–647, 648–1236, 1237–2392 and 2393~nucleotides (nts). Each fraction contains the same number of mRNAs registered in the RefSeq database.



2.4. Exo-miRNA Targets with Short 3' UTRs Are Efficiently Downregulated and Endo-miRNA Targets with Long 3'UTRs Are Efficiently Upregulated by the Introduction of Exo-miRNA Duplex

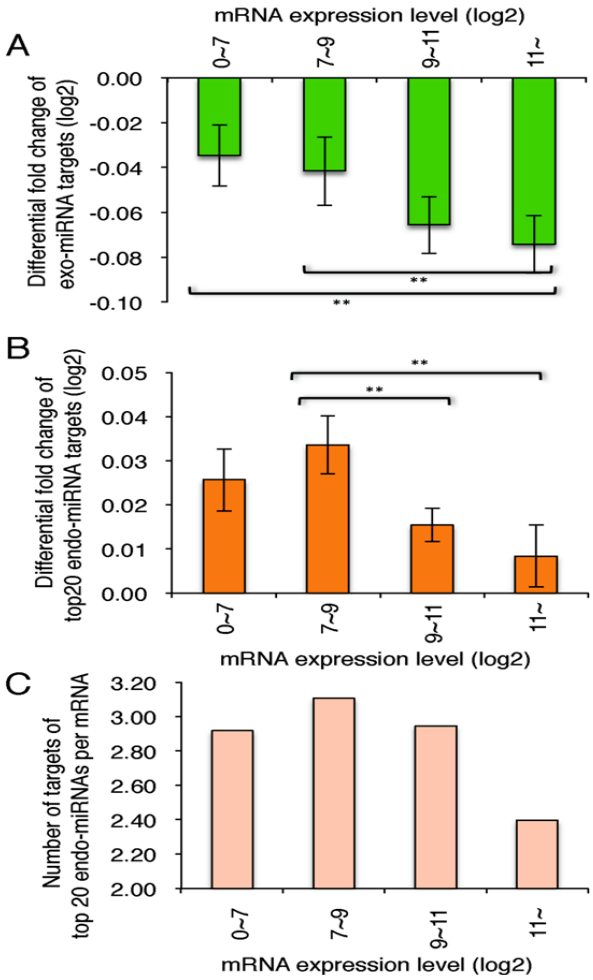
Among exo-miRNA target genes, the genes with short 3' UTRs were more efficiently downregulated than those with long 3' UTRs by the transfection of exo-miRNAs (Figure 5A),

consistent with the previous report [31]. In this analysis, we used the fold change values of exo-miRNA targets with 3' UTRs complementary to seed sequence of either one strand of exo-miRNA duplex: let-7b-5p, miR-1, miR-21-5p, miR-22-5p, miR-28-5p, miR-30c-5p, miR-186-5p, miR-199b-5p, miR-200b-3p, miR-330-5p, miR-335-5p, miR-346, miR-466, miR-574-5p and miR-3126-5p. In contrast, the top 20 endo-miRNA targets, which have no exo-miRNA target sites, were found to be more strongly upregulated, according to the increase of the 3' UTR lengths (Figure 5B). It was presumed that a large number of endo-miRNA target sites are situated in the long 3' UTRs, but small in the short 3' UTRs; then, the genes with long 3' UTRs might be strongly repressed by endo-miRNAs in the normal condition, but those with short 3' UTRs might be repressed weakly. So, we calculated the number of top 20 endo-miRNA target sites in the various length of mRNA 3' UTRs. As expected, a large number of target sites of top 20 endo-miRNAs were found in the long 3' UTRs, but a small number of them were in the short 3' UTRs (Figure 5C), indicating that the genes with long 3' UTRs should be strongly repressed by endo-miRNAs compared to those with short 3' UTRs in the normal condition. However, once exo-miRNAs are transfected into the cells, a part of the endo-miRNAs loaded on the RISCs should be replaced with exo-miRNAs. As a result, the constant downregulation of endo-miRNA targets by endo-miRNAs might be cancelled, and the expression of these genes are upregulated. Thus, the changes of expression levels of the exo-miRNA target genes might be elaborately regulated by the endo-miRNAs pre-situated in the RISCs.

2.5. Highly Expressed Exo-miRNA Targets Are Efficiently Downregulated and Endo-miRNA Targets with Low Expression Levels Are Efficiently Upregulated by the Transfection of the Exo-miRNA Duplex

Among exo-miRNA targets, the genes with high expression levels in the normal conditions were more efficiently downregulated by the transfection of exo-miRNAs compared to those with low expression levels (Figure 6A). The results showed good agreement with the previous study [31]. In contrast, the endo-miRNA targets with low expression levels in the normal conditions were uncovered to be intensively upregulated by the exo-miRNA transfection, and those with high expression levels were not upregulated efficiently (Figure 6B). So, we investigated the number of endo-miRNA target sites in gene with different expression levels. As a result, a large number of target sites were found in the mRNAs with low expression levels, but a small number of the sites were detected in the mRNAs with high expression levels (Figure 6C), suggesting that many genes with low expression levels might be downregulated by endo-miRNAs, but those with high expression levels are not repressed by endo-miRNAs in the normal condition. Thus, the changes of expression levels of the exo-miRNA target genes might be also suitably regulated by the endo-miRNAs pre-situated in the RISCs.

Figure 6. Mean fold changes of target genes of exo-miRNAs and endo-miRNAs, according to the expression levels. Differential fold changes (log_2) of target gene expression levels of exo-miRNAs (**A**) and endo-miRNAs (**B**) and the number of the top 20 endo-miRNA target sites (**C**), according to the expression levels of exo-miRNA targets. The averaged numbers of exo-miRNA target sites are 130, 217, 237 and 215 (**A**), and those of endo-miRNA target sites are 92, 164, 176 and 202 (**B**) in the 3' UTRs of genes with differential fold changes of 0~7, 7~9, 9~11 and 11~.



3. Experimental Section

3.1. Cell Culture and miRNA Synthesis

Human HeLa cells were cultured and used in reporter assays and microarray analyses. Cells were cultured at 37 °C in Dulbecco's modified Eagle's Medium (Invitrogen, Carlsbad, NM, USA), supplemented with 10% heat-inactivated fetal bovine serum (Sigma, St. Louis, MO, USA). They were plated on 24-well culture plates (1×10^5 cells/mL/well) 24 h prior to transfection. Transfection was carried out using Lipofectamine 2000 (Invitrogen, Carlsbad, NM, USA). Each RNA strand of the miRNA duplex was chemically synthesized (Sigma, St. Louis, MO, USA) and annealed to form the duplex structures the same as those shown in miRBase [28]. The sequences of the synthetic miRNAs (let-7b, miR-1, miR-21, miR-22, miR-28, miR-30c-1, miR-186, miR-199b, miR-200b, miR-330, miR-335, miR-346, miR-466, miR-574 and miR-3126) are listed in Table S1.

3.2. Construction of Luciferase Reporters

All of the reporter plasmids constructed were derivatives of psiCHECK-1 (Promega, Fitchburg, WI, USA). Oligonucleotides with target sequences completely matched to each miRNA strand (cm-target) were chemically synthesized with cohesive *Xho*I/*Eco*RI ends (Table S2). They were then inserted into the corresponding restriction sites of psiCHECK-1 to generate miRNA cm-targets (miR-200b-3p target, miR-7b-5p target, miR-21-5p target and miR-330-5p target). Each of the inserted targets was expressed as part of the 3' UTR region of *Renilla* luciferase mRNA in transfected cells.

HeLa cells growing in 24-well plates were transfected simultaneously with miRNA target (100 ng), pGL3-Control (Promega, 0.5 μg) and miRNA (50 nM). The cells were harvested 24 h post-transfection and the relative luciferase activity (*Renilla* luc activity/firefly luc activity) was determined using a Dual-Luciferase Reporter Assay System (Promega, Fitchburg, WI, USA). The pGL3-Control encoding firefly luciferase served as a control for the calculation of relative luciferase activity for miRNAs.

3.3. Microarray Analysis

HeLa cells (1×10^5 cells/mL) were transfected with 50 nM of each of 15 miRNA duplexes. At 24 h post-transfection, total RNA was purified using an RNeasy Kit (Qiagen, Hilden, Germany). The steps were repeated four times and RNA quality assessed using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA) and a Bioanalyzer (Agilent, Santa Clara, CA, USA). RNAs recovered independently were mixed equally for cDNA synthesis using an Agilent One Color Spike Mix Kit (Agilent, Santa Clara, CA, USA). Cy3-labeled cRNA was synthesized using a Quick Amp Labeling Kit (Agilent, Santa Clara, CA, USA) and was hybridized to an Agilent Whole Human Genome Microarray (4 × 44 K multi-pack format), according to the manufacturer's protocol. RNA from mock-transfected cells treated with transfection reagent in the absence of miRNA was used as a control. Transcript expression values were calculated using Microarray Suite 5.0 (MAS5: Affymetrix, Santa Clara, CA, USA) [32] with quantile normalization [33]. To identify transcripts whose expression was upregulated or downregulated, the cumulative distribution of expression changes for transcripts containing the site was compared with that for transcripts with no canonical site. NCBI's Reference Sequence (RefSeq) was used to identify mRNAs with sequences complementary to the seed regions of the transfected miRNAs. Data are presented as an MA plot (M = intensity ratio, A = average intensity) and a cumulative frequency distribution. Changes in expression are shown as fold changes ($\log_2$).

3.4. Quantitative RT-PCR

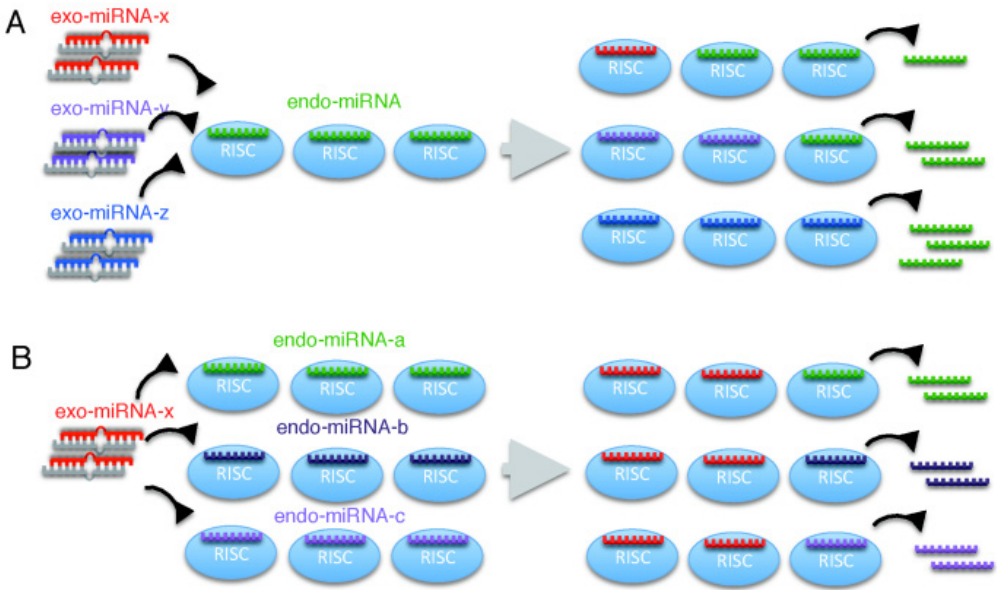
Total RNA was reverse-transcribed using a Transcriptor High Fidelity cDNA Synthesis kit (Roche, Basel, Switzerland). The resultant cDNA samples were incubated with FastStart Universal SYBR Green Master (Roche, Basel, Switzerland) at 95 °C for 10 min, followed by PCR amplification. PCR product levels were monitored using an ABI PRISM 7000 sequence detection system and analyzed with ABI PRISM 7000 SDS software (Applied Biosystems). The expression of each target gene was first normalized to that of β-actin and then to the mock-transfection control. The primer sets used are listed in Table S3.

4. Conclusions

In this study, we quantified the changes in expression levels of endo-miRNA target genes resulting from the transfection of exo-miRNA duplexes. The expression levels of endo-miRNA target genes with seed-complementary sequences were increased by the transfection of exo-miRNA duplexes, while exo-miRNA target gene expression was reduced (Figures 1–3 and Figures S1–S3). These results suggest that exo-miRNA duplex transfected into the cells may compete with endo-miRNAs for the RISC, which may be saturated with endo-miRNAs under normal conditions in HeLa cells.

In miRNA-mediated gene silencing, the structure of the RNA-protein complex is known to be altered [23]. The miRNA duplex in RISC loading complex is unwound, yielding single-stranded RNA, which is loaded onto the RISC and recognizes target mRNAs through base-pairing in the seed region [4–6]. Our results indicate that the expression of the target genes of a given endo-miRNA differed according to the exo-miRNA duplex that was transfected (Figures 2 and 7A), whereas with a given exo-miRNA duplex, the fold changes in the target gene expression of the endo-miRNAs examined were mostly equivalent, except for a limited types of exo-miRNAs (miR-1, miR-28, miR-199b and miR-335), despite differences in the structures and sequences of the endo-miRNA duplexes (Figure 3 and Figure 7B). One of the possible explanations of these results is that the RISC exchange reaction might be occurred associated with single-stranded endo-miRNAs and not endo-miRNA duplexes or target-paired endo-miRNAs (Figure 7), because double-stranded endo-miRNAs in the RISC might not be replaced with exo-miRNAs at similar levels, due to their different structures and sequences. Most miRNA-RISCs might be in the form of single-stranded miRNAs in RISCs, so as to be readily replaced by double-stranded miRNAs.

Figure 7. Predicted model for the RNA-induced silencing complexes (RISCs) replacement. Different types of exo-miRNAs transfected into cells may be replaced with single-stranded endo-miRNAs loaded on the RISC with different efficiencies (**A**); however, different types of endo-miRNAs may be replaced with a given exo-miRNAs with similar efficiencies (**B**).



Furthermore, it was apparently revealed that endo-miRNAs constantly repress the expression of endogenous mRNAs with endo-miRNA target sites, and their repression is probably relieved by the replacement of endo-miRNAs on the RISC by the exo-miRNAs transfected (Figures 5 and 6). Thus, global gene expression by endogenous miRNAs might be fluctuated by the transfection of exo-miRNAs or the increase of expression levels of endo-miRNAs. Competition similar to that shown here between exo-miRNA duplexes and endo-miRNAs may also occur among newly transcribed endo-miRNAs and may provide a mechanism for orchestrating cellular programs.

Acknowledgments

We thank Yuki Naito for helpful discussions about the bioinformatics analysis and Kenji Nishi for technical advises on microarray experiments. This work was supported by a Grant-in-Aid for Scientific Research (grant numbers 21115004, 21310123), the Cell Innovation Project from the Ministry of Education, Culture, Sports, Science and Technology of Japan and the Core Research Project for Private University matching fund subsidy to K.U.-T.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Rana, T.M. Illuminating the silence: Understanding the structure and function of small RNAs. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 23–36.
2. Filipowicz, W.; Bhattacharyya, S.N.; Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? *Nat. Rev. Genet.* **2008**, *9*, 102–114.
3. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297.
4. Grimson, A.; Farh, K.K.; Johnston, W.K.; Garrett-Engele, P.; Lim, L.P.; Bartel, D.P. MicroRNA targeting specificity in mammals: Determinants beyond seed pairing. *Mol. Cell* **2007**, *27*, 91–105.
5. Lewis, B.P.; Burge, C.B.; Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **2005**, *120*, 15–20.
6. Lim, L.P.; Lau, N.C.; Garrett-Engele, P.; Grimson, A.; Schelter, J.M.; Castle, J.; Bartel, D.P.; Linsley, P.S.; Johnson, J.M. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **2005**, *433*, 769–773.
7. Grimm, D.; Streetz, K.L.; Jopling, C.L.; Storm, T.A.; Pandey, K.; Davis, C.R.; Marion, P.; Salazar, F.; Kay, M.A. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* **2006**, *441*, 537–541.
8. Castanotto, D.; Sakurai, K.; Lingeman, R.; Li, H.; Shively, L.; Aagaard, L.; Soifer, H.; Gatignol, A.; Riggs, A.; Rossi, J.J. Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC. *Nucleic Acids Res.* **2007**, *35*, 5154–5164.
9. Stewart, C.K.; Li, J.; Golovan, S.P. Adverse effects induced by short hairpin RNA expression in porcine fetal fibroblasts. *Biochem. Biophys. Res. Commun.* **2008**, *370*, 113–117.

10. Khan, A.A.; Betel, D.; Miller, M.L.; Sander, C.; Leslie, C.S.; Marks, D.S. Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs. *Nat. Biotechnol.* **2009**, *27*, 549–555.
11. Yi, R.; Qin, Y.; Macara, J.G.; Cullen, B.R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **2003**, *17*, 3011–3016.
12. Lund, E.; Guttinger, S.; Calado, A.; Dahlberg, J.E.; Kutay, U. Nuclear export of microRNA precursors. *Science* **2004**, *303*, 95–98.
13. Bohnsack, M.T.; Czaplinski, K.; Gorlich, D. Exportin-5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* **2004**, *10*, 185–191.
14. Kim, V.N. MicroRNA precursors in motion; exportin-5 mediates their nuclear export. *Trends Cell Biol.* **2004**, *14*, 156–159.
15. Vankoningsloo, S.; de Longueville, F.; Evrard, S.; Rahier, P.; Houbion, A.; Fattaccioli, A.; Gastellier, M.; Remacle, J.; Raes, M.; Ranard, P.; *et al*. Gene expression silencing with "specific" small interfering RNA goes beyond specificity–a study of key parameters to take into account in the onset of small interfering RNA off-target effects. *FEBS J.* **2008**, *275*, 2738–2753.
16. Gregory, R.I.; Chendrimada, T.P.; Cooch, N.; Shienkhattar, R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **2005**, *123*, 631–540.
17. Haase, A.D.; Jaskiewicz, L.; Zhang, H.; Lainé, S.; Sack, R.; Gatignol, A.; Filipowicz, S. TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functionas in RNA silencing. *EMBO J.* **2005**, *6*, 961–967.
18. Macrae, I.J.; Ma, E.; Zhou, M.; Robinson, C.V.; Doudna, J.A. *In vitro* reconstitution of the human RISC-loading complex. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 512–517.
19. Winter, J.; Jung, S.; Keller, S.; Gregory, R.; Diederichs, S. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat. Cell Biol.* **2009**, *11*, 228–234.
20. Matranga, C.; Tomari, Y.; Shin, C.; Bartal, D.P.; Zamore, P.D. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **2005**, *123*, 607–620.
21. Leuschner, P.J.; Ameres, S.L.; Kueng, S.; Martinez, J. Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Rep.* **2006**, *7*, 314–320.
22. Tolia, N.H.; Joshua-Tor, L. Slicer and the Argonaute. *Nat. Chem. Biol.* **2006**, *3*, 36–43.
23. Czech, B.; Hannon, G.J. Small RNA sorting: Matchmaking for Argonautes. *Nat. Rev. Genetics* **2011**, *12*, 19–31.
24. Khvorova, A.; Reynolds, A.; Jayasena, S.D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **2003**, *115*, 209–215.
25. Schwarz, D.S.; Hutvágner, G.; Du, T.; Xu, A.; Aronin, N.; Zamore, P.D. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **2003**, *115*, 199–205.
26. Ui-Tei, K.; Naito, Y.; Takahashi, F.; Haraguchi, T.; Ohki-Hamazaki, H.; Juni, A.; Ueda, R.; Saigo, K. Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. *Nucleic Acids Res.* **2004**, *32*, 936–948.
27. Yoda, M.; Kawamata, T.; Paroo, Z.; Ye, X.; Iwasaki, S.; Liu, Q.; Tomari, Y. ATP-dependent human RISC assembly pathways. *Nat. Struct. Mol. Biol.* **2010**, *17*, 17–23.

28. Griffiths-Jones, S.; Saini, H.K.; van Dongen, S.; Enright, A.J. MiRBase tools for microRNA genomics. *Nucleic Acids Res.* **2008**, *36*, D154–D158.
29. Jackson, A.L.; Burchard, J.; Schelter, J.; Chau, B.N.; Cleary, M.; Lim, L.; Linsley, P.S. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA* **2006**, *12*, 1179–1187.
30. Mayr, C.; Bartel, D.P. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **2009**, *138*, 673–684.
31. Saito, T.; Sætrom, P. Target gene expression levels and competition between transfected and endogenous microRNAs are strong confounding factors in microRNA high-throughput experiments. *Silence* **2012**, *3*, 3.
32. Hubbell, E.; Liu, W.M.; Mei, R. Robust estimation for expression analysis. *Bioinformatics* **2002**, *18*, 1585–1592.
33. Bolstad, B.M.; Irizarry, R.A.; Gautier, L.; Wu, Z. Preprocessing High-Density Oligonucleotide Arrays. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*; Gentleman, R., Carey, V.J., Huber, W., Irizarry, R.A., Dudoit, S., Eds.; Springer: New York, NY, USA, 2005; pp. 13–32.

Reprinted from *IJMS*. Cite as: Tomaselli, S.; Bonamassa, B.; Alisi, A.; Nobili, V.; Locatelli, F.; Gallo, A. ADAR Enzyme and miRNA Story: A Nucleotide that Can Make the Difference. *Int. J. Mol. Sci.* **2013**, *14*, 22796-22816.

Review

ADAR Enzyme and miRNA Story: A Nucleotide that Can Make the Difference

Sara Tomaselli [1,†], Barbara Bonamassa [1,†], Anna Alisi [2,*], Valerio Nobili [2], Franco Locatelli [1,3] and Angela Gallo [1,*]

[1] Laboratory of RNA Editing, Onco-haematology Department, Bambino Gesù Children's Hospital, IRCCS, Piazza S. Onofrio 4, Rome 00165, Italy; E-Mails: sara.tomaselli@opbg.net (S.T.); barbara.bonamassa@opbg.net (B.B.); franco.locatelli@opbg.net (F.L.)

[2] Hepato-Metabolic Disease Unit and Liver Research Unit, Bambino Gesù Children's Hospital, IRCCS, Piazza S. Onofrio 4, Rome 00165, Italy; E-Mail: valerio.nobili@opbg.net

[3] Department of Pediatric Science, Università di Pavia, Strada Nuova 65, Pavia 27100, Italy

† These authors contributed equally to this work.

* Authors to whom correspondence should be addressed; E-Mails: anna.alisi@opbg.net (A.A.); angela.gallo@opbg.net (A.G.); Tel.: +39-066-859-2658 (A.G.); Fax: +39-066-859-2904 (A.G.).

Received: 8 October 2013; in revised form: 4 November 2013 / Accepted: 5 November 2013 / Published: 19 November 2013

Abstract: Adenosine deaminase acting on RNA (ADAR) enzymes convert adenosine (A) to inosine (I) in double-stranded (ds) RNAs. Since Inosine is read as Guanosine, the biological consequence of ADAR enzyme activity is an A/G conversion within RNA molecules. A-to-I editing events can occur on both coding and non-coding RNAs, including microRNAs (miRNAs), which are small regulatory RNAs of ~20–23 nucleotides that regulate several cell processes by annealing to target mRNAs and inhibiting their translation. Both miRNA precursors and mature miRNAs undergo A-to-I RNA editing, affecting the miRNA maturation process and activity. ADARs can also edit 3' UTR of mRNAs, further increasing the interplay between mRNA targets and miRNAs. In this review, we provide a general overview of the ADAR enzymes and their mechanisms of action as well as miRNA processing and function. We then review the more recent findings about the impact of ADAR-mediated activity on the miRNA pathway in terms of biogenesis, target recognition, and gene expression regulation.

Keywords: microRNA; Adenosine deaminase acting on RNA (ADAR); A-to-I RNA editing; double-stranded RNA (dsRNA); non-coding sequence

1. Introduction

Protein-coding genes account for approximately 1% of the mammalian genome and 70%–90% of the rest can be transcribed but not translated [1]. Therefore, a large part of the human transcriptome consists of non-coding RNA sequences, (*i.e.*, UTRs, introns of protein-coding genes and non-coding RNAs, such as transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), microRNAs (miRNAs), small interfering RNAs (siRNAs), piwi-interacting RNAs (piRNAs), long non-coding RNAs (lncRNAs), small nuclear RNAs (snRNAs), and small nucleolar RNAs (snoRNAs)). Both protein-coding and non-coding RNAs undergo several post-transcriptional modifications, which (partially) account for the complexity of both the transcriptome and proteome that characterizes the high level of gene regulation in higher eukaryotes [2]. Among these post-transcriptional mechanisms, RNA editing is an ubiquitous and crucial modification event that alters RNA molecules by nucleotide modification bypassing the genomic information [3,4]. There are different types of RNA editing [3], but the best characterized and frequent editing event in higher eukaryotes involves the conversion of adenosine (A) to inosine (I) in double-stranded RNA (dsRNA) regions through the action of the Adenosine Deaminase Acting on RNA (ADAR) enzymes [4–6].

Computational analysis combined with next generation sequencing (NGS) has recently been used to identify A-to-I RNA editing sites [7–10]. Reverse transcriptase recognizes Inosine as Guanosine. Therefore, an A-to-I RNA editing site can be identified when a cDNA sequence and the corresponding genomic DNA (gDNA) sequence are aligned. Surprisingly, several editing sites were found in non-coding regions of the human transcriptome (~15,000 sites, mapped in ~2000 different genes) and most of them are clustered within inversely oriented repetitive Alu elements (~90%). On the basis of this analysis, it is predicted that >85% of pre-mRNAs are possibly edited, with the vast majority being targeted in introns (~90%) and UTRs [10].

As Inosine is interpreted as Guanosine by splicing and translational machineries, A-to-I editing can change the informational content of the RNA coding molecules by altering splicing and translation processes. Moreover, Inosine has different base-pairing properties compared to Adenosine and differs from Guanosine by the loss of the N_2 amino group (due to the ADAR deamination event), which accounts for the less strong interaction with Cytosine (two H-bound instead of three). Thus, A-to-I RNA editing has the potential to alter RNA structure by introducing bulges/mismatches or creating different base pairs (for examples, A-U base pairs can change into I:U mismatches in dsRNAs). The final picture is that ADARs can alter splicing, translation, and the dsRNA structure. It was originally thought that the main function of ADAR enzymes was their re-coding capacity. However, A-to-I editing most frequently targets non-coding sequences [9,11,12] and, recently, numerous interactions between ADARs and miRNA/siRNA pathways [13] have been discovered, which suggests a role of ADARs and A-to-I editing in RNA-mediated regulation of gene expression. In this review, we first provide a general overview of the ADAR enzymes and their mechanisms of action. We then focus on

the miRNA pathway and the effects of ADAR-mediated modifications on the biogenesis and functions of miRNAs.

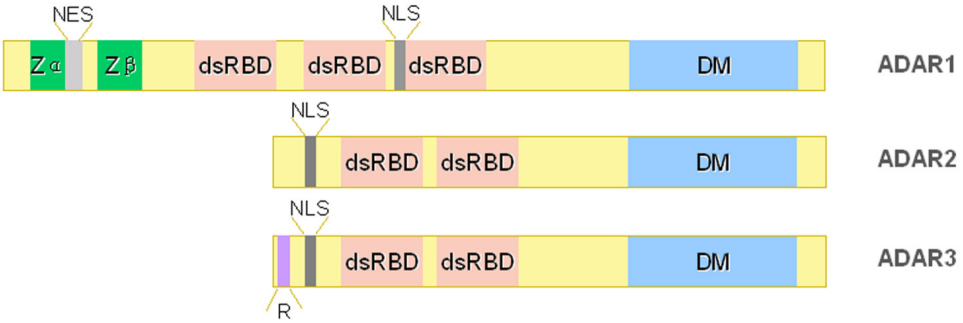
2. ADAR Family

ADAR-mediated A-to-I RNA editing converts A to I by hydrolytic deamination of adenine bases. Three ADARs (ADAR1, ADAR2, and ADAR3) are present in vertebrates (Figure 1). ADARs contain a highly conserved catalytic deaminase domain (DM) at their *C*-terminal. Crystallography structure of the DM showed that the surface of this domain contains a positively-charged cleft for the binding of negatively-charged dsRNA and that it catalyses the hydrolytic deamination of Adenosine via a catalytic zinc ion [14]. Moreover, an inositol hexakisphosphate (IP6) was found buried within the enzyme core that contributes to the protein fold [14]. A nucleotide "flip-out" mechanism is necessary to force the targeted Adenosine into the catalytic pocket in the correct orientation for the deamination reaction [15].

The second key domain of all ADAR enzymes is the dsRNA-binding domain (dsRBD) at the *N*-terminus. Each dsRBD (three for ADAR1 and two for ADAR2-3) has an α-β-β-β-α topology consisting of approximately 70 amino acids, with the two α helices packing against a three-stranded anti-parallel β-sheets. Multiple dsRBDs are thought to act synergistically, which, as a consequence, increases both the affinity and specificity for dsRNA targets [16].

At the *N*-terminus, ADAR1 carries a Z-DNA binding domains (Zα plus Zβ) that suggests its localization at highly transcribed DNA sites. Moreover, as these domains can also bind Z-RNA, ADAR1 is also able to localize to underwound dsRNAs in RNA virus [17,18].

Figure 1. Structure of ADAR family proteins: ADAR1, ADAR2, and ADAR3. The ADAR enzymes contain a *C*-terminal conserved catalytic deaminase domain (DM), two or three dsRBDs in the *N*-terminal portion. ADAR1 full-length protein also contains a *N*-terminal Zα domain with a nuclear export signal (NES) and a Zβ domain, while ADAR3 has a R-domain. A nuclear localization signal is also indicated.



2.1. ADAR1

The *ADAR1* human gene is located on the long arm of chromosome 1 (1q21.3) spanning ~30Kbp [19]. The protein was discovered to exist in two isoforms of different size, *i.e.*, the interferon (IFN-α, -β, and -γ)-inducible long (ADAR1L, 150 KDa) and the constitutive short (ADAR1S, 110 KDa) isoform, which result from the use of alternative start codons and promoters [20]. While

ADAR1S promoter is constitutively active, IFN can induce ADAR1L, suggesting a role in the cellular response to stress factors such as viral infections [21]. In addition to the finding of regulatory elements within the IFN-inducible ADAR1 promoter, recent studies revealed distinct tissue-specific expression features for different ADAR1 transcripts [22]. Both transcripts contain three dsRBDs but the *N*-terminus of ADAR1L has several domains that are absent in ADAR1S, including an arginine-glycine-enriched domain (RG domain) and a nuclear export signal (NES) within the Zα domain. Thus, ADAR1L is found both in the cytoplasm and nucleus since it also has a nuclear/nucleolar localization signal (NLS/NoLS) [23]. Consequently, the intracellular distribution of the various ADAR1 isoforms is determined by the export/import regulatory proteins available in a cell. On the contrary, ADAR1S localizes mainly to the nucleus since it carries only the NLS/NoLS signal. However, it has been shown that ADAR1S can also localize to the cytoplasm thanks to the cooperative action of all three dsRBDs, with dsRNAs able to interact with exportin-5 [24].

The small ubiquitin-like modifier 1 (SUMO-1) binds ADAR1 at lysine 418, decreasing the editing activity of the enzyme [25]. ADAR enzymes can form homo- and hetero-dimers and dimerization is essential for their editing activity [26,27]. Several studies have shown that ADAR1 (and ADAR2) can work as homodimer, whereas other investigations have demonstrated that also heterodimers can be formed, which may be necessary for the ADARs to act as active deaminases [27–29]. Other ADAR-interacting proteins include the nuclear factor 90 (NF90) proteins [30], the protein-kinase RNA-activated protein (PKR) [31], the adenovirus-associated (VAI) RNA [32], and the Vaccinia virus E3L protein [33].

2.2. ADAR2

The *ADAR2* human gene is located on the long arm of chromosome 21 (21q22.3), spanning ~153 Kbp [34]. The promoter that directs ADAR2 expression has not been functionally characterized, although a putative promoter region upstream of a newly identified exon was described for both the human and the mouse *Adar2* gene [35]. This promoter includes a TATA box sequence and the consensus binding sites for the Nuclear factor kappa-light-chain-enhancer of activated B cells (Nf-κB) and for the Specificity Protein 1 (SP1) [35]. While it has to be established whether ADAR2 possesses multiple promoters to produce multiple transcripts like ADAR1, the regulatory mechanism(s) driving the transcriptional control of ADAR2 in a tissue- and cell type-specific fashion have been partially unveiled. Indeed, it was shown that cAMP response element-binding (CREB) can indirectly induce ADAR2 expression [36]. More recently, Yang *et al.* [37] demonstrated that JNK1 serves as a crucial component in mediating glucose-responsive up-regulation of ADAR2 expression in pancreatic β-cells, suggesting that the JNK1 pathway may be functionally linked to the nutrient-sensing actions of ADAR2-mediated RNA editing in professional secretory cells.

ADAR2 *N*-terminus has an arginine-enriched domain (R-domain) (similar to that identified in the ADAR3 protein, Figure 1) that contains a NLS [35,38], while an extra NLS is located before the first dsRBD [35,39]. Consequently, ADAR2 localizes into the cell nuclei thanks to the action of importin α1, α4, and α5 [39].

ADAR2 can form homodimers and heterodimers with ADAR1 [27–29]. ADAR2 dimerization seems to be essential for editing activity, although it is not clear whether the interaction is or not dsRNA-mediated [27,40,41].

2.3. ADAR3

The *ADAR3* human gene is located on the short arm of chromosome 10 (10p15) in proximity of the telomere [42]. Although ADAR3 has conserved all the key catalytic residues of the ADAR family members, no deaminase activity has been found for this enzyme so far [43]. All the editing sites have been, thus, attributed to ADAR1 and 2 activity. ADAR3 protein carries two dsRBDs and, additionally, an R-domain that binds single-stranded RNAs (ssRNAs) [43,44], suggesting that both ss- and dsRNAs can be bound by the enzyme. ADAR3 is localized in the nucleus of the cell and interacts with the importin α1 through the R-motif [35].

Differently from the ubiquitously expressed ADAR1 and 2, ADAR3 is expressed at detectable levels only in certain post-mitotic cells in the central nervous system (CNS) [43]. Furthermore, ADAR3 remains in the monomeric form, which may explain the lack of editing activity, at least in part [28]. Thus, ADAR3 function is unknown so far, although its ability to bind both ss- and dsRNAs would suggest a regulatory activity over ADAR1 and 2. Indeed, ADAR3 can compete for dsRNA substrates preventing the binding of the other ADAR enzymes [43,45].

3. ADAR Substrates

Any dsRNAs of ≥20 bp can be an ADAR substrate [6]. ADAR substrates were originally identified by chance, comparing cDNAs to their genomic counterparts and finding editing events as a mixture of A/G instead of A only. Different editing sites have been identified over the years, particularly in transcripts coding for proteins expressed in the CNS [4,5,46], *i.e.*, those coding for subunits of the glutamate receptor super-family GluR, the serotonin 5-hydroxytryptamine 2C (5-HT2C)-receptor, the potassium voltage-gated channel (Kv1.1), and the a3 subunit of the γ-aminobutyric acid (GABAA) receptor. These editing events have a major impact on protein properties.

More recently, bio-computational studies and innovative sequencing techniques have demonstrated that A-to-I RNA editing mainly affects non-coding RNAs [6]. Importantly, the majority of editing events occur in introns and 5'–3' UTRs enriched with Alu repeat-mediated dsRNAs. Recently, a database collecting the identified (validated or not) editing RNAs has become available (http://darned.ucc.ie) [47].

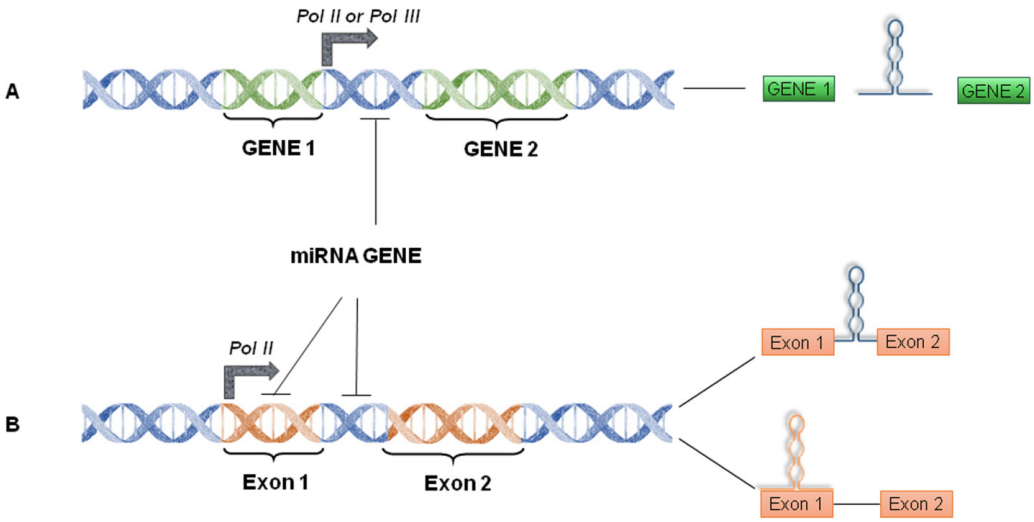
ADAR-mediated editing levels range from 2% to 100% [5,13,46], depending on cell and tissue type [48] as well as developmental stages [49]. How ADAR chooses the target adenosine is still not completely clear. ADARs show slight sequence preferences [50]. However, dsRNA length and structure seem to play an important role. For example, dsRNAs of 15–40 bp are edited selectively at very few sites, whereas those longer than 50 bp are extensively or non-selectively deaminated (with 50%–60% of adenosines being edited) [51]. Similarly, selective deamination is also observed in dsRNAs with bulges, loops, and mismatches [52]. It has been suggested that ADAR substrate specificity may also depend on editor modulators (such as snoRNAs) [53] and on the different dsRBD number and spacing of ADAR proteins that allow discrimination between dsRNA structures and stabilities. While the importance of site-specific editing (within coding sequence genes or microRNAs) has been explored and was found to affect the final protein or miRNA maturation/targeting, the role of the non-specific/promiscuous editing (within non-coding RNA portions such as introns and

5'–3' UTRs) is still poorly understood. However, recent studies would point out their involvement in modulation of gene expression, which may occur by changing the splicing enhancers/silencers recognition sites [54–57], by perturbing/inducing the binding of RBPs for RNA nuclear localization/ retention [58] or inducing inosine-specific degradation (Tudor-SN nuclease) [59].

4. miRNA World Machinery Overview

miRNAs are short (~20–23 nucleotides) ssRNAs that regulate, at post-transcriptional level, several genes playing crucial roles in various cellular processes such as cell cycle, apoptosis, differentiation and, when deregulated, neoplastic transformation [60]. Mammalian miRNA genes (in cluster or as single unit) are located either in introns/exons of protein-coding genes, in non-coding genes, or in intra-genic regions of the genome [61–63] (Figure 2). Intronic/exonic miRNAs are often transcribed by the RNA polymerase (Pol) II and co-expressed with their host gene, while intergenic miRNAs are independently transcribed by either RNA Pol II or III [64,65]. Usually, miRNA promoters located in the inter-genic or non-coding regions of the genome are regulated by transcriptional or epigenetic factors like protein-coding genes [66].

Figure 2. Schematic diagram of the miRNA genes. (**A**) Monocistronic intergenic miRNA gene; (**B**) Monocistronic exonic/intronic miRNA gene.



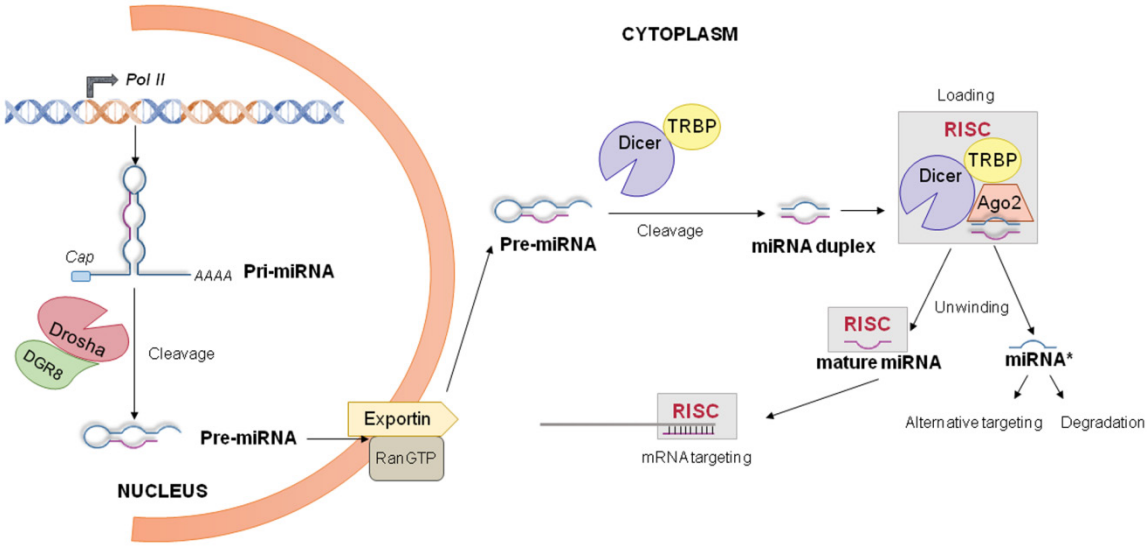
Each miRNA may regulate several mRNAs post-transcriptionally, while a single mRNA can be targeted by several miRNAs via base-pairing to the mRNA 3' UTRs [67,68].

The conventional theory assumes that the “seed sequence” (~6–8 nucleotides in length) at miRNA 5' end is crucial for target specificity and mediates its binding to 3' UTRs of target mRNAs, causing their translational repression or degradation [69]. However, recent studies suggest that miRNAs can exert their action over specific targets using alternative mechanisms, including the binding to specific proteins or to non-coding RNAs [70,71]. The biogenesis and processing of miRNAs occur in the nucleus/cytoplasm due to the action of multiple proteins. Some of these have a well-known role(s) in miRNA processing, including Drosha, exportin 5, Argonaute (Ago), and Dicer, while others have partially been explored such as ADAR1 [72].

4.1. miRNA Biogenesis and Processing into the Nucleus

The early step of miRNA biogenesis in the nucleus is the transcription of a miRNA precursor (Figure 3). Mature miRNAs are generated from long, hairpin-shaped primary transcripts (pri-miRNA) that are usually several thousand nucleotides long [66]. After transcription, pri-miRNAs undergo multiple steps of processing into the nucleus. Conventional nuclear processing of pri-miRNAs happen due to their cleavage by a large microprocessor complex (650 kDa in humans) consisting of the RNase III enzyme Drosha and the DiGeorge syndrome Critical Region gene 8 (DGCR8) protein [73,74]. Specifically, Drosha, a nuclear protein of 130–160 kDa, cuts the 5' and 3' ends of the pri-miRNA molecule with its RNase domain, giving a short hairpin of 60–70 nucleotides long (pre-miRNA) [66]. Although DGCR8-Drosha microprocessor is involved in the cropping of many miRNAs, Drosha may also form larger complexes with other proteins (e.g., RNA helicases, dsRNA binding proteins, heterogeneous nuclear ribonucleoproteins, *etc*.) to regulate the processing of specific pri-miRNAs [75]. A recent study provides evidence that certain mature miRNAs combined with Ago proteins may re-enter the nucleus and inhibit the pri-miRNA processing [76].

Figure 3. miRNA biogenesis and processing. Canonical biogenesis of pri-miRNA transcription is mediated by Pol II. Next, the microprocessor complex composed of Drosha and DGCR8 mediates the nuclear cleavage of pri-miRNA into pre-miRNA. The nuclear export of pre-miRNA is subsequently mediated by exportin-5/Ran-GTP61. Cytoplasmic pre-miRNA is processed by Dicer into a duplex microRNA. The next step is the unwinding of the duplex into a mature ~22 nucleotide miRNA and a miRNA* by the RISC complex. The mature miRNA is generally conveyed by the RISC on the targeted mRNA, whilst miRNA* can be degraded or alternatively perform a different targeting.



Following the nuclear processing, pre-miRNAs are exported to the cytoplasm by an energy-dependent mechanism involving the exportin-5/Ran-GTP61 complex. Exportin-5 binds pre-miRNA molecules and Ran-GTP61, which catalyses GTP hydrolysis and the consequent release of pre-miRNA short precursors into the cytoplasm. Interestingly, Exportin-5 also hampers pre-miRNA nuclear accumulation, protecting them from a potential nuclear digestion and retention [77,78]. In

addition to the nuclear-to-cytoplasm pre-miRNA flux, the presence of functional mature miRNAs into the nucleus suggests a retrograde transport regulated by other carriers such as Importin 8 [79].

4.2. miRNA Processing into the Cytoplasm

Once exported from the nucleus, the cytoplasmic pre-miRNA duplex is further processed by Dicer and other accessory proteins, including the transactivation response RNA binding protein (TRBP), the protein activator of the dsRNA-dependent protein kinase (PACT), and the Ago proteins (Figure 3). Together they form the RNA-induced silencing complex (RISC) [80–83].

For miRNAs displaying a high degree of complementarity along the hairpin stem, a preliminary Ago 2-dependent cleavage is required before Dicer action. This Ago 2 slicer activity generates a nicked hairpin, producing a precursor miRNA or ac-pre-miRNA that is further processed by Dicer [84]. Dicer typically cleaves pre-miRNA duplexes near the terminal loop, releasing a small RNA duplex of ~22 nucleotides [66].

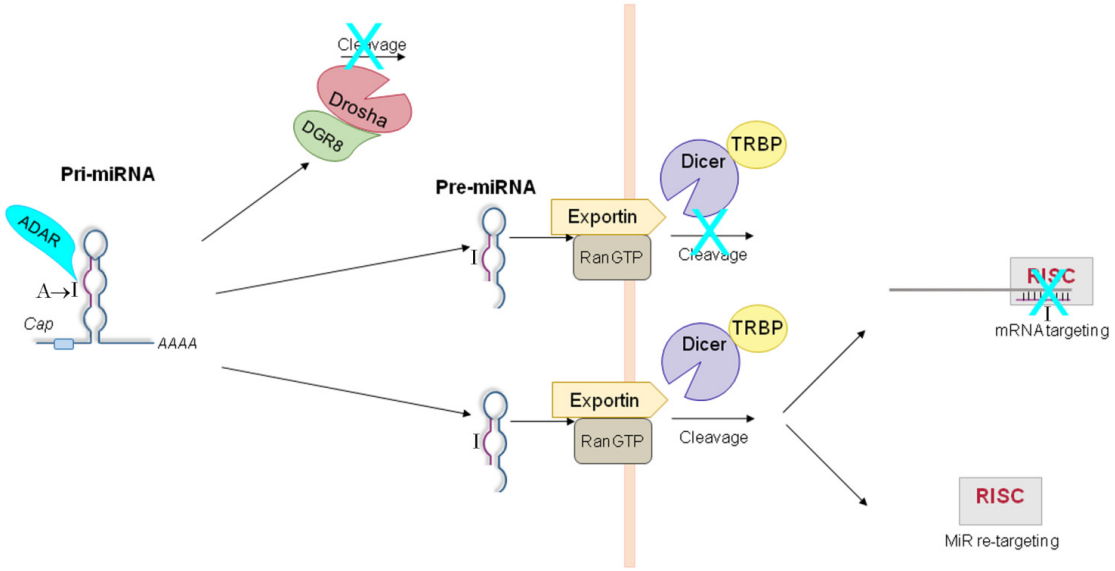
After Dicer-mediated cleavage, the small RNA duplex is loaded onto an Ago protein (Ago 1–4 in mammals) of the RISC to generate the microRNA containing ribonucleoprotein complex, *i.e.*, miRNP or miRISC. Usually one single-strand (named guide) of the duplex (which is complementary to the target mRNA) is charged on Ago 2 as a mature miRNA, while the other strand of the duplex (named passenger or miRNA*) is usually degraded. miRNA guide (or in some cases miRNA* [85,86]) is selected to associate with Ago proteins by their thermodynamic stability [87]. There are at least two other hypotheses to explain duplex unwinding into guide and passenger strand. Dicer could cleave the miRNA*, releasing the miRNA guide that is subsequently captured by Ago 2. Alternatively, the miRNA* of a loaded duplex could be cleaved by the slicer activity of Ago 2, which simultaneously retains the miRNA guide. The activated RISC can bind the target mRNA, and direct its degradation, or repress its translation [88]. However, it has been reported that in some cases, miRNAs can also up-regulate the expression of their targets [85,89].

5. ADAR-Dependent Effects on miRNA Pathway

As ADARs can bind to and edit any dsRNA, the discovery that these enzymes are able to modify dsRNA substrates that enter the miRNA-mediated gene silencing and RNA interference (RNAi) pathways, *i.e.*, miRNA and siRNA precursors [13], does not come as a surprise. It has been shown that mammalian pri-miRNAs undergo A-to-I RNA editing in adult brain [86,90–92]. Furthermore, NGS analysis has shown that ADARs can alter miRNA processing and sequence in *C. elegans*, mouse embryos, human and mouse brain [93–96]. Moreover, a more recent study showed that ADAR1 forms a complex with Dicer, promoting miRNA processing, RISC loading of miRNAs and silencing of target RNAs independently of its deaminase activity [72], as previously suggested [97].

In summary, several miRNA precursors (pri- and pre-miRs) undergo specific A-to-I RNA editing that may inhibit their maturation process and, thus, the production of mature miRNAs, affecting the loading of the edited miRNA to the RISC complex, or redirecting the edited miRNA to a new set of target mRNAs (Figure 4). Considering that A-to-I editing can also occur within the 3' UTR regions of mRNAs, the picture of miRNA-ADAR interaction becomes even more complex, underlining the high level of regulation of the miRNA world.

Figure 4. Editing-dependent effects of ADARs on miRNA pathway. miRNA precursors (pri- and pre-miRs) undergo specific A-to-I RNA editing that (i) may block their maturation process at either Drosha or Dicer step; (ii) may affect the loading of the edited miRNA to the RISC complex; (iii) may redirect the edited miRNA to a new set of target mRNAs.



5.1. ADAR-Dependent Effects on Pri-miRs

The first report of RNA editing events in a miRNA precursor dates back to almost ten years ago, when Maas and co-authors detected a low level (~5% in human brain) of A-to-I changes within the pri-miR-22 [86]. Using human cell lines (HEK293T), ectopically expressing ADAR1 or ADAR2, they found that pri-miR-22 is mainly edited by ADAR1, although the physiological role of this editing was not elucidated.

A couple of years later, Yang *et al*. confirmed that ADARs can interact with pri-miRNAs using RNA editing assays and data from *Adar1* and *Adar2* null mice [98]. Four out of the eight analyzed miRNA precursors displayed A-to-I editing *in vitro* (*i.e.*, pri-miR-142, -223, -1-1, -143), with pri-miR-142 harboring the highest editing levels. Both ADAR1S and ADAR2 are able to edit pri-miR-142 at 11 specific sites, nine of which lie within the mature miRNA sequence. Transfecting edited pri-miR-142 in HEK293 cells, the authors determined that editing at the +4 and +5 sites destroys the integrity of the stem-loop structure, inhibiting the maturation of the pri- to pre-miRs. The consequence is a reduced production of mature miR-142. Indeed, the levels of endogenous miR-142 were lower in wild-type mouse spleens than those in *Adar1* and *Adar2* null mouse spleens. However, some editing sites (such as the one at site +40) seem not to affect pri-miR-142 processing. Editing-mediated inhibition of miRNA maturation at the pri-miR step does not cause accumulation of the edited pri-miR-142, as it may be degraded by Tudor-SN, a component of the RISC complex [99], known to mediate the degradation of inosine-containing dsRNAs (IU-dsRNAs) *in vitro* [59].

The discovery that edited pri-miRs can undergo rapid degradation by Tudor-SN suggests that the amount of edited pri-miRNAs into a cell could be higher than previously hypothesized. A recent study showed that ADAR1L (the ADAR1 nucleus/cytoplasmatic shuttling isoform) and Tudor-SN co-localize in the cytoplasm within stress granules (SGs) in HeLa cells under various stress

conditions [100]. The authors speculated that ADAR1 may edit target dsRNAs in the cytoplasm and the resultant IU-dsRNA may recruit Tudor-SN to form SGs during cell stress responses. However, further experiments are needed to better define the role of ADAR1 in this context and the importance of the ADAR1-mediated SG formation.

5.2. ADAR-Dependent Effects on Pre-miRs

Editing can also influence Dicer cleavage, which is responsible for the processing of pre- into miRNAs. This has been first demonstrated for pri-miR-151 [101]. ADAR1-dependent editing at the -1 and +3 site has been reported [90,101], which reduces the efficiency of the Dicer-TRBP activity and results in the production of unedited mature miR-151 [101]. Interestingly, editing of mouse pri-miR-151 is CNS-specific, although both ADAR1 and pri-miR-151 were found expressed in many non-brain tissues.

5.3. ADAR-Dependent Effects on RISC-Loading

Epstein-Barr virus (EBV) encodes 23 miRNAs that are implicated in the attenuation of host antiviral immune response and the transition from latent to lytic replication [102,103]. Among EBV miRNAs, four primary miRNAs were found to undergo site-specific A-to-I editing events [104]. The authors focused on pri-miR-BART6, which showed high editing levels at the +20 in EBV latently infected cell lines. This editing reduces the correct loading of miR-BART6-5p into the RISC complex. Remarkably, this is the first report of pri-miRNA A-to-I editing that suppresses RISC loading [104]. Editing of pri-miR-BART6 reduces the activity of mature miR-BART6, playing a crucial role in the regulation of EBV life cycle and cell immune response.

Recently, new A-to-I editing events have been reported within another EBV miRNA, *i.e.*, pri-miR-BART3. Editing was found at four sites in EBV-infected epithelial carcinoma cells and in nasopharyngeal carcinoma samples, affecting both the biogenesis and targeting of mature miR-BART3 [105].

5.4. ADAR-Dependent Effects on Retargeting

A specific ADAR-mediated A-to-I change has been reported in Kaposi's sarcoma-associated herpesvirus (KSHV) transcripts [106,107]. This alteration modifies the seed sequence of the mature miR-K10, potentially affecting its target mRNAs [106,107]. ADAR1S heavily edits the K12 transcript in a specific site, as shown by *in vitro* editing assays [106]. Importantly, the authors observed that this editing event has a functional significance, playing a key role in the replication strategy of HHV-8 and in its tumorigenic potential. This was the first evidence that ADAR-mediated editing can also affect the target specificity of a mature miRNA.

Subsequently, Nishikura and colleagues demonstrated that edited mature miRNAs play a biological function *in vivo* [91]. The human pri-miR-376a1, previously showed to be edited [90], is situated in a cluster of 6 pri-miRNAs. The authors disclosed that five out of these six miRNAs are edited in human tissues (*i.e.*, pri-miR-367a1, -367a2, -367b, -368, -B2). Several adenosines within the miR-376 cluster members undergo A-to-I editing, with two positions showing the highest editing levels (nearly 100% in certain tissues), *i.e.*, the +4 site, which is preferentially edited by ADAR2, and the +44 site, which is

selectively edited by ADAR1. These editing events do not affect the primary transcript maturation steps. However, both +4 and +44 sites lay within the seed sequences of miR-376a* and miR-376a respectively, suggesting that the edited miRNAs could have a different target mRNA profile. In particular, the authors demonstrated that a single ADAR2-mediated base change (at the +4 site) is able to modulate the expression of phosphoribosyl pyrophosphate synthetase 1 (PRPS1), a mouse protein involved in purine metabolism and uric acid synthesis [91].

Notably, a recent work has elegantly demonstrated the existence of a tight link between miR-376a editing and human brain tumors [108]. Choudhury *et al*. found that RNA editing of miR-376 cluster is extremely reduced in human gliomas, with glioblastoma cells accumulating almost exclusively the unedited form of miR-376a*. The unedited miRNA promotes glioma cell migration and invasion, whilst the edited form inhibits these capacities *in vitro*. These effects are the consequence of a different mRNA target specificity of the edited and unedited form of the miRNA [108]. As ADAR2 is responsible for miR-376a* editing, these findings strengthen the notion that this enzyme plays a crucial role in glioma progression, as previously shown [45,109,110].

5.5. ADAR-Dependent Effects on Target 3′ UTRs

As most of the editing sites are also located in 3' UTRs of human mRNAs [111], an additional interplay between ADAR activity and miRNAs is possible. Computational screening showed that RNA editing tends to avoid miRNA binding sites, with less than 10% of editing events occurring in 3' UTR regions recognised by miRNAs [111]. However, it was also found that editing can create new miRNA target sites [111].

More recent analyses indicate that up to 20% of the editing sites in the 3' UTR of human mRNAs may alter miRNA target sites [112], making the mRNA resistant to miRNA activity. In addition, in mouse tissues, A-to-I changes seem to be highly frequent in 3' UTR regions, including miRNA target sites [113]. Wang *et al.* provided novel insights into the mechanism by which ADAR1 and its activity regulate miRNA-mediated modulation of target gene expression [114]. Indeed, multiple A-to-I RNA editing events (mediated by ADAR1) were found within the 3' UTR of *ARHGAP26*, encoding the Rho GTPase activating protein 26. Furthermore, the authors revealed that both miR-30b* and miR-573 are able to target ARHGAP26, but that editing make this transcript resistant to repression mediated by these two miRNAs.

5.6. ADAR-Mediated Editing-Independent Effects on miRNAs

In addition to A-to-I sequence changes on miRNAs, ADARs can also act through an editing-independent mechanism by binding dsRNAs [97]. Heale *et al.* found that ADAR1 and ADAR2 editing activity can result in retargeting of human miR-376a2, as shown previously for mouse miR-376 [91]. By performing *in vitro* pri-miRNA processing assays, they also pointed out that, even in the absence of editing, ADAR2 can inhibit the processing of pri-miR-376a2 at the Drosha cleavage step [97]. Therefore, the simple binding of ADAR proteins to dsRNAs may have a range of biological roles that are still to be fully discovered.

6. Large-Scale Surveys

Initial low-throughput experiments followed by NGS approaches have been performed by several groups, adding new insights on the role of ADARs in the miRNA pathway.

One of the original systematic survey proposed that 6% of all human pri-miRNAs are edited [90]. The author determined that six out of 99 pri-miRNAs undergo editing (*i.e.*, pri-miR-151, -197, -223, -376a, -379, -99a) in humans. The extent of editing ranged from ~10% to 70%, depending on sites and different tissues analyzed. Most of the editing events were located in the mature miRNA seed sequence, suggesting that RNA editing may contribute to increase miRNA diversity. This paper established that ADARs edit miRNAs but did not elucidate the functional consequences of these events.

A couple of years later, a larger scale survey of 209 human pri-miRNAs showed that ~16% of them undergo A-to-I editing in human brain, with editing levels ranging from ~10% to 100% [92]. Then, for six randomly chosen edited pri-miRNAs (*i.e.*, pri-let-7g, pri-miR-33, -133a2, -197, -203, -379) it was discovered that editing alters either the Drosha or the Dicer cleavage step. It is worth noting that the processing of two pri-miRNAs (*i.e.*, pri-miR-197 and -203) was enhanced by editing. The authors also showed that some pri-miRs are preferentially edited by ADAR1 (*i.e.*, pri-miR-99b, -151, -376b, -411, -423), while others by ADAR2 (*i.e.*, pri-let-7g, pri-miR-27a, -99a, -203, -376a, -379) [92].

Recent advances in high-throughput small RNA sequencing (smRNA-Seq) have reshaped the miRNA research landscape, including RNA editing analysis. Using a novel strategy to avoid cross-mapping artefacts, de Hoon *et al*. found that editing prevalence in human mature miRNAs is extremely low in a human monocytic leukemia cell line (THP-1) [115]. Ten potential miRNA editing sites were found. However, eight of these were due to cross-mapping, one was due to a single nucleotide polymorphism, and the remaining editing site (in the mature miR-376c) was already identified [91]. Similar results were obtained by sequencing small RNAs from mouse brain [116].

Recently, Vesely and co-workers analyzed the frequency and sequence composition of miRNA pools from transgenic *Adar* null mouse embryos by NGS [93]. *Adar2* deficiency leads to a change in the expression level of specific target mRNAs when compared to wild-type embryos. In particular, the authors detected 10 edited miRNAs, four of which had been identified previously (*i.e.*, mmu-miR-378, -376b, -381, -3099) and six were novel edited miRNAs (*i.e.*, mmu-miR-1957, -467d*, -706, -1186, -3102-5p.2, -703). Some editing events were located in the seed region, opening the possibility that editing could lead to their retargeting. However, the biological consequences of the observed editing events are difficult to interpret, especially because of the low levels detected.

Using NGS followed by bioinformatics analysis, Eisenberg and co-workers found a clear A-to-I signal in mature miRNAs of human brain [94]. Overall, 19 statistically significant modification sites (mainly due to ADAR2 activity) were detected in 18 different miRNAs, confirming previously detected editing sites as well as revealing several novel ones. Most of the detected A-to-G modifications were within the miRNA seed sequence, with editing significantly changing their binding specificity. As previously reported, a relatively low editing level was found, with few exceptions (editing percentage ranging from 0.2% to 70%) [94].

7. Stimulative Role of ADAR1

ADAR1 has been emerging as a promoter for small non-coding RNAs. Indeed, a recent study has highlighted the important role of ADAR1 in interacting with Dicer to form heterodimers [72]. Notably, the authors established that ADAR1 uses its second dsRBD to form ADAR1/Dicer heterodimers (acting as modulator of RNAi machinery) and its third dsRBD to form ADAR1/ADAR1 homodimers (acting as an RNA editing enzyme). The ADAR1/Dicer interaction increases the rate of processing from pre- to mature miRNAs, promotes the RISC loading and, consequently, the mRNA silencing efficacy [72]. It seems that neither dsRNA-binding nor deaminase activity of ADAR1 is required for these effects. As expected, the authors found that the miRNA expression is inhibited in *Adar1* null mouse embryos, as a consequence of the lack of formation of the Dicer/ADAR1 complex with a final alteration of the target genes [72].

8. Conclusions

A-to-I editing is believed to be an important way of generating protein diversity by codon alteration in mRNAs. However, editing sites in some coding targets make up only a tiny fraction of all editing events, most of which are actually located in non-coding sequences such as introns, UTRs or regulatory RNAs (miRNAs and their precursors). The biological function of editing in non-coding RNA sequences remains not completely disclosed. As far as miRNAs are concerned, the general feeling about A-to-I changes is that they regulate the levels of cellular dsRNAs, which, if not kept under control, are potent triggers of gene silencing and signaling pathway. Despite this, important questions still stand. At which extent and how diffuse is the RNA editing on mature miRNAs and their precursors? Is it a developmentally regulated or a tissue specific phenomenon? In principle, editing at any level of miRNA biogenesis may have a broad influence on expression patterns. Although the evidence is still limited, a critical examination of data reported in the literature does offer some examples of miRNA down-stream activity misregulation. One more question is whether there is any correlation between edited miRNAs and human diseases. While alterations in both substrate editing and ADAR expression/activity are often reported in different pathologies [4,5,117], the effects of edited miRNA pathways on disease onset/progression still deserves further investigation. In this context, it is worth noting that Choudhury *et al.* demonstrated that a single editing event in the miR-376a* seed sequence dramatically alters the selection of its target genes and redirects its function from inhibiting to promoting glioma cell invasion [108]. Overall, these pieces of information set the stage for further investigations, either to address the aforementioned questions and, possibly, to score against ADAR/miRNA editing-linked human diseases.

Acknowledgments

This work was supported by the PRIN 2012NA9E9Y_004 and the IG grant of AIRC (Milan, Italy) and to A.G., my First AIRC Grant to A.A., and the special project 5 × 1000 AIRC to F.L. We thank Marion Huth for the English revision.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Lee, J.T. Epigenetic regulation by long noncoding RNAs. *Science* **2012**, *338*, 1435–1439.
2. Baltimore, D. Our genome unveiled. *Nature* **2001**, *409*, 814–816.
3. Gott, J.M.; Emeson, R.B. Functions and mechanisms of RNA editing. *Annu. Rev. Genet.* **2000**, *34*, 499–531.
4. Gallo, A.; Locatelli, F. ADARs: Allies or enemies? The importance of A-to-I RNA editing in human disease: From cancer to HIV-1. *Biol. Rev. Camb. Philos. Soc.* **2011**, *87*, 95–110.
5. Keegan, L.P.; Gallo, A.; O'Connell, M.A. The many roles of an RNA editor. *Nat. Rev. Genet.* **2001**, *2*, 869–878.
6. Nishikura, K. Functions and regulation of RNA editing by ADAR deaminases. *Annu. Rev. Biochem.* **2010**, *79*, 321–349.
7. Blow, M.; Futreal, P.A.; Wooster, R.; Stratton, M.R. A survey of RNA editing in human brain. *Genome Res.* **2004**, *14*, 2379–2387.
8. Kim, D.D.; Kim, T.T.; Walsh, T.; Kobayashi, Y.; Matise, T.C.; Buyske, S.; Gabriel, A. Widespread RNA editing of embedded alu elements in the human transcriptome. *Genome Res.* **2004**, *14*, 1719–1725.
9. Levanon, E.Y.; Eisenberg, E.; Yelin, R.; Nemzer, S.; Hallegger, M.; Shemesh, R.; Fligelman, Z.Y.; Shoshan, A.; Pollock, S.R.; Sztybel, D.; *et al*. Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat. Biotechnol.* **2004**, *22*, 1001–1005.
10. Athanasiadis, A.; Rich, A.; Maas, S. Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. *PLoS Biol.* **2004**, *2*, 2144–2158.
11. Morse, D.P.; Aruscavage, P.J.; Bass, B.L. RNA hairpins in noncoding regions of human brain and Caenorhabditis elegans mRNA are edited by adenosine deaminases that act on RNA. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7906–7911.
12. Maas, S.; Rich, A.; Nishikura, K. A-to-I RNA editing: Recent news and residual mysteries. *J. Biol. Chem.* **2003**, *278*, 1391–1394.
13. Nishikura, K. Editor meets silencer: Crosstalk between RNA editing and RNA interference. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 919–931.
14. Macbeth, M.R.; Schubert, H.L.; Vandemark, A.P.; Lingam, A.T.; Hill, C.P.; Bass, B.L. Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science* **2005**, *309*, 1534–1539.
15. Kuttan, A.; Bass, B.L. Mechanistic insights into editing-site specificity of ADARs. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E3295–E3304.
16. Ryter, J.M.; Schultz, S.C. Molecular basis of double-stranded RNA-protein interactions: Structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J.* **1998**, *17*, 7505–7513.
17. Herbert, A.; Alfken, J.; Kim, Y.G.; Mian, I.S.; Nishikura, K.; Rich, A. A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 8421–8426.

18. Brown, B.A., 2nd; Lowenhaupt, K.; Wilbert, C.M.; Hanlon, E.B.; Rich, A. The zalpha domain of the editing enzyme dsRNA adenosine deaminase binds left-handed Z-RNA as well as Z-DNA. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 13532–13536.
19. Weier, H.U.; George, C.X.; Greulich, K.M.; Samuel, C.E. The interferon-inducible, double-stranded RNA-specific adenosine deaminase gene (DSRAD) maps to human chromosome 1q21.1-21.2. *Genomics* **1995**, *30*, 372–375.
20. Patterson, J.B.; Samuel, C.E. Expression and regulation by interferon of a double-stranded-RNA-specific adenosine deaminase from human cells: Evidence for two forms of the deaminase. *Mol. Cell. Biol.* **1995**, *15*, 5376–5388.
21. Patterson, J.B.; Thomis, D.C.; Hans, S.L.; Samuel, C.E. Mechanism of interferon action: Double-stranded RNA-specific adenosine deaminase from human cells is inducible by alpha and gamma interferons. *Virology* **1995**, *210*, 508–511.
22. George, C.X.; Das, S.; Samuel, C.E. Organization of the mouse RNA-specific adenosine deaminase Adar1 gene 5'-region and demonstration of STAT1-independent, STAT2-dependent transcriptional activation by interferon. *Virology* **2008**, *380*, 338–343.
23. Poulsen, H.; Nilsson, J.; Damgaard, C.K.; Egebjerg, J.; Kjems, J. CRM1 mediates the export of ADAR1 through a nuclear export signal within the Z-DNA binding domain. *Mol. Cell. Biol.* **2001**, *21*, 7862–7871.
24. Fritz, J.; Strehblow, A.; Taschner, A.; Schopoff, S.; Pasierbek, P.; Jantsch, M.F. RNA-regulated interaction of transportin-1 and exportin-5 with the double-stranded RNA-binding domain regulates nucleocytoplasmic shuttling of ADAR1. *Mol. Cell. Biol.* **2009**, *29*, 1487–1497.
25. Desterro, J.M.; Keegan, L.P.; Jaffray, E.; Hay, R.T.; O'Connell, M.A.; Carmo-Fonseca, M. SUMO-1 modification alters ADAR1 editing activity. *Mol. Biol. Cell* **2005**, *16*, 5115–5126.
26. Gallo, A.; Keegan, L.P.; Ring, G.M.; O'Connell, M.A. An ADAR that edits transcripts encoding ion channel subunits functions as a dimer. *EMBO J.* **2003**, *22*, 3421–3430.
27. Valente, L.; Nishikura, K. RNA binding-independent dimerization of adenosine deaminases acting on RNA and dominant negative effects of nonfunctional subunits on dimer functions. *J. Biol. Chem.* **2007**, *282*, 16054–16061.
28. Cho, D.S.; Yang, W.; Lee, J.T.; Shiekhattar, R.; Murray, J.M.; Nishikura, K. Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. *J. Biol. Chem.* **2003**, *278*, 17093–17102.
29. Chilibeck, K.A.; Wu, T.; Liang, C.; Schellenberg, M.J.; Gesner, E.M.; Lynch, J.M.; MacMillan, A.M. FRET analysis of *in vivo* dimerization by RNA-editing enzymes. *J. Biol. Chem.* **2006**, *281*, 16530–16535.
30. Nie, Y.; Ding, L.; Kao, P.N.; Braun, R.; Yang, J.H. ADAR1 interacts with NF90 through double-stranded RNA and regulates NF90-mediated gene expression independently of RNA editing. *Mol. Cell. Biol.* **2005**, *25*, 6956–6963.
31. Clerzius, G.; Gelinas, J.F.; Daher, A.; Bonnet, M.; Meurs, E.F.; Gatignol, A. ADAR1 interacts with PKR during human immunodeficiency virus infection of lymphocytes and contributes to viral replication. *J. Virol.* **2009**, *83*, 10119–10128.
32. Lei, M.; Liu, Y.; Samuel, C.E. Adenovirus VAI RNA antagonizes the RNA-editing activity of the ADAR adenosine deaminase. *Virology* **1998**, *245*, 188–196.

33. Liu, Y.; Wolff, K.C.; Jacobs, B.L.; Samuel, C.E. Vaccinia virus E3L interferon resistance protein inhibits the interferon-induced adenosine deaminase A-to-I editing activity. *Virology* **2001**, *289*, 378–387.
34. Mittaz, L.; Scott, H.S.; Rossier, C.; Seeburg, P.H.; Higuchi, M.; Antonarakis, S.E. Cloning of a human RNA editing deaminase (ADARB1) of glutamate receptors that maps to chromosome 21q22.3. *Genomics* **1997**, *41*, 210–217.
35. Maas, S.; Gommans, W.M. Novel exon of mammalian ADAR2 extends open reading frame. *PLoS One* **2009**, *4*, e4225.
36. Peng, P.L.; Zhong, X.; Tu, W.; Soundarapandian, M.M.; Molner, P.; Zhu, D.; Lau, L.; Liu, S.; Liu, F.; Lu, Y. ADAR2-dependent RNA editing of AMPA receptor subunit GluR2 determines vulnerability of neurons in forebrain ischemia. *Neuron* **2006**, *49*, 719–733.
37. Yang, L.; Huang, P.; Li, F.; Zhao, L.; Zhang, Y.; Li, S.; Gan, Z.; Lin, A.; Li, W.; Liu, Y. c-Jun amino-terminal kinase-1 mediates glucose-responsive upregulation of the RNA editing enzyme ADAR2 in pancreatic beta-cells. *PLoS One* **2012**, *7*, e48611.
38. Desterro, J.M.; Keegan, L.P.; Lafarga, M.; Berciano, M.T.; O'Connell, M.; Carmo-Fonseca, M. Dynamic association of RNA-editing enzymes with the nucleolus. *J. Cell Sci.* **2003**, *116*, 1805–1818.
39. Maas, S.; Gommans, W.M. Identification of a selective nuclear import signal in adenosine deaminases acting on RNA. *Nucleic Acids Res.* **2009**, *37*, 5822–5829.
40. Jaikaran, D.C.; Collins, C.H.; MacMillan, A.M. Adenosine to inosine editing by ADAR2 requires formation of a ternary complex on the GluR-B R/G site. *J. Biol. Chem.* **2002**, *277*, 37624–37629.
41. Poulsen, H.; Jorgensen, R.; Heding, A.; Nielsen, F.C.; Bonven, B.; Egebjerg, J. Dimerization of ADAR2 is mediated by the double-stranded RNA binding domain. *RNA* **2006**, *12*, 1350–1360.
42. Mittaz, L.; Antonarakis, S.E.; Higuchi, M.; Scott, H.S. Localization of a novel human RNA-editing deaminase (hRED2 or ADARB2) to chromosome 10p15. *Hum. Genet.* **1997**, *100*, 398–400.
43. Chen, C.X.; Cho, D.S.; Wang, Q.; Lai, F.; Carter, K.C.; Nishikura, K. A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *RNA* **2000**, *6*, 755–767.
44. Melcher, T.; Maas, S.; Herb, A.; Sprengel, R.; Higuchi, M.; Seeburg, P.H. RED2, a brain-specific member of the RNA-specific adenosine deaminase family. *J. Biol. Chem.* **1996**, *271*, 31795–31798.
45. Cenci, C.; Barzotti, R.; Galeano, F.; Corbelli, S.; Rota, R.; Massimi, L.; Di Rocco, C.; O'Connell, M.A.; Gallo, A. Down-regulation of RNA editing in pediatric astrocytomas: ADAR2 editing activity inhibits cell migration and proliferation. *J. Biol. Chem.* **2008**, *283*, 7251–7160.
46. Bass, B.L. RNA editing by adenosine deaminases that act on RNA. *Annu. Rev. Biochem.* **2002**, *71*, 817–846.
47. Kiran, A.; Baranov, P.V. DARNED: A DAtabase of RNa EDiting in humans. *Bioinformatics* **2010**, *26*, 1772–1776.
48. Galeano, F.; Leroy, A.; Rossetti, C.; Gromova, I.; Gautier, P.; Keegan, L.P.; Massimi, L.; Di Rocco, C.; O'Connell, M.A.; Gallo, A. Human BLCAP transcript: New editing events in normal and cancerous tissues. *Int. J. Cancer* **2010**, *127*, 127–137.

49. Wahlstedt, H.; Daniel, C.; Enstero, M.; Ohman, M. Large-scale mRNA sequencing determines global regulation of RNA editing during brain development. *Genome Res.* **2009**, *19*, 978–986.
50. Lehmann, K.A.; Bass, B.L. Double-stranded RNA adenosine deaminases ADAR1 and ADAR2 have overlapping specificities. *Biochemistry* **2000**, *39*, 12875–12884.
51. Nishikura, K.; Yoo, C.; Kim, U.; Murray, J.M.; Estes, P.A.; Cash, F.E.; Liebhaber, S.A. Substrate specificity of the dsRNA unwinding/modifying activity. *EMBO J.* **1991**, *10*, 3523–3532.
52. Lehmann, K.A.; Bass, B.L. The importance of internal loops within RNA substrates of ADAR1. *J. Mol. Biol.* **1999**, *291*, 1–13.
53. Vitali, P.; Basyuk, E.; Le Meur, E.; Bertrand, E.; Muscatelli, F.; Cavaille, J.; Huttenhofer, A. ADAR2-mediated editing of RNA substrates in the nucleolus is inhibited by C/D small nucleolar RNAs. *J. Cell Biol.* **2005**, *169*, 745–753.
54. Lev-Maor, G.; Sorek, R.; Levanon, E.Y.; Paz, N.; Eisenberg, E.; Ast, G. RNA-editing-mediated exon evolution. *Genome Biol.* **2007**, *8*, R29.
55. Lev-Maor, G.; Sorek, R.; Shomron, N.; Ast, G. The birth of an alternatively spliced exon: 3' splice-site selection in Alu exons. *Science* **2003**, *300*, 1288–1291.
56. Lev-Maor, G.; Ram, O.; Kim, E.; Sela, N.; Goren, A.; Levanon, E.Y.; Ast, G. Intronic Alus influence alternative splicing. *PLoS Genet.* **2008**, *4*, e1000204.
57. Sakurai, M.; Yano, T.; Kawabata, H.; Ueda, H.; Suzuki, T. Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome. *Nat. Chem. Biol.* **2010**, *6*, 733–740.
58. Zhang, Z.; Carmichael, G.G. The fate of dsRNA in the nucleus: A p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* **2001**, *106*, 465–475.
59. Scadden, A.D. The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage. *Nat. Struct. Mol. Biol.* **2005**, *12*, 489–496.
60. He, L.; He, X.; Lim, L.P.; de Stanchina, E.; Xuan, Z.; Liang, Y.; Xue, W.; Zender, L.; Magnus, J.; Ridzon, D.; *et al*. A microRNA component of the p53 tumour suppressor network. *Nature* **2007**, *447*, 1130–1134.
61. Lagos-Quintana, M.; Rauhut, R.; Lendeckel, W.; Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **2001**, *294*, 853–858.
62. Lau, N.C.; Lim, L.P.; Weinstein, E.G.; Bartel, D.P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **2001**, *294*, 858–862.
63. Lee, R.C.; Ambros, V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **2001**, *294*, 862–864.
64. Rodriguez, A.; Griffiths-Jones, S.; Ashurst, J.L.; Bradley, A. Identification of mammalian microRNA host genes and transcription units. *Genome Res.* **2004**, *14*, 1902–1910.
65. Morlando, M.; Ballarino, M.; Gromak, N.; Pagano, F.; Bozzoni, I.; Proudfoot, N.J. Primary microRNA transcripts are processed co-transcriptionally. *Nat. Struct. Mol. Biol.* **2008**, *15*, 902–909.
66. Kim, V.N.; Han, J.; Siomi, M.C. Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell. Biol.* **2009**, *10*, 126–139.
67. Lewis, B.P.; Burge, C.B.; Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **2005**, *120*, 15–20.
68. Filipowicz, W.; Bhattacharyya, S.N.; Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? *Nat. Rev. Genet.* **2008**, *9*, 102–114.

69. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297.
70. Chen, X.; Liang, H.; Zhang, C.Y.; Zen, K. miRNA regulates noncoding RNA: A noncanonical function model. *Trends Biochem. Sci.* **2011**, *37*, 457–459.
71. Fabbri, M.; Paone, A.; Calore, F.; Galli, R.; Gaudio, E.; Santhanam, R.; Lovat, F.; Fadda, P.; Mao, C.; Nuovo, G.J.; *et al*. MicroRNAs bind to Toll-like receptors to induce prometastatic inflammatory response. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E2110–E2116.
72. Ota, H.; Sakurai, M.; Gupta, R.; Valente, L.; Wulff, B.E.; Ariyoshi, K.; Iizasa, H.; Davuluri, R.V.; Nishikura, K. ADAR1 forms a complex with Dicer to promote microRNA processing and RNA-induced gene silencing. *Cell* **2013**, *153*, 575–589.
73. Lee, Y.; Ahn, C.; Han, J.; Choi, H.; Kim, J.; Yim, J.; Lee, J.; Provost, P.; Radmark, O.; Kim, S.; Kim, V.N. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **2003**, *425*, 415–419.
74. Han, J.; Lee, Y.; Yeom, K.H.; Nam, J.W.; Heo, I.; Rhee, J.K.; Sohn, S.Y.; Cho, Y.; Zhang, B.T.; Kim, V.N. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **2006**, *125*, 887–901.
75. Winter, J.; Jung, S.; Keller, S.; Gregory, R.I.; Diederichs, S. Many roads to maturity: MicroRNA biogenesis pathways and their regulation. *Nat. Cell Biol.* **2009**, *11*, 228–234.
76. Tang, R.; Li, L.; Zhu, D.; Hou, D.; Cao, T.; Gu, H.; Zhang, J.; Chen, J.; Zhang, C.Y.; Zen, K. Mouse miRNA-709 directly regulates miRNA-15a/16-1 biogenesis at the posttranscriptional level in the nucleus: Evidence for a microRNA hierarchy system. *Cell Res.* **2012**, *22*, 504–515.
77. Yi, R.; Qin, Y.; Macara, I.G.; Cullen, B.R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **2003**, *17*, 3011–3016.
78. Lund, E.; Guttinger, S.; Calado, A.; Dahlberg, J.E.; Kutay, U. Nuclear export of microRNA precursors. *Science* **2004**, *303*, 95–98.
79. Weinmann, L.; Hock, J.; Ivacevic, T.; Ohrt, T.; Mutze, J.; Schwille, P.; Kremmer, E.; Benes, V.; Urlaub, H.; Meister, G. Importin 8 is a gene silencing factor that targets argonaute proteins to distinct mRNAs. *Cell* **2009**, *136*, 496–507.
80. Gregory, R.I.; Chendrimada, T.P.; Cooch, N.; Shiekhattar, R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **2005**, *123*, 631–640.
81. Chendrimada, T.P.; Gregory, R.I.; Kumaraswamy, E.; Norman, J.; Cooch, N.; Nishikura, K.; Shiekhattar, R. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **2005**, *436*, 740–744.
82. Lee, Y.; Hur, I.; Park, S.Y.; Kim, Y.K.; Suh, M.R.; Kim, V.N. The role of PACT in the RNA silencing pathway. *EMBO J.* **2006**, *25*, 522–532.
83. MacRae, I.J.; Ma, E.; Zhou, M.; Robinson, C.V.; Doudna, J.A. *In vitro* reconstitution of the human RISC-loading complex. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 512–517.
84. Diederichs, S.; Haber, D.A. Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* **2007**, *131*, 1097–1108.
85. Nilsen, T.W. Mechanisms of microRNA-mediated gene regulation in animal cells. *Trends Genet.* **2007**, *23*, 243–249.
86. Luciano, D.J.; Mirsky, H.; Vendetti, N.J.; Maas, S. RNA editing of a miRNA precursor. *RNA* **2004**, *10*, 1174–1177.

87. Takeda, A.; Iwasaki, S.; Watanabe, T.; Utsumi, M.; Watanabe, Y. The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins. *Plant Cell Physiol.* **2008**, *49*, 493–500.
88. Valencia-Sanchez, M.A.; Liu, J.; Hannon, G.J.; Parker, R. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev.* **2006**, *20*, 515–524.
89. Carthew, R.W.; Sontheimer, E.J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **2009**, *136*, 642–655.
90. Blow, M.J.; Grocock, R.J.; van Dongen, S.; Enright, A.J.; Dicks, E.; Futreal, P.A.; Wooster, R.; Stratton, M.R. RNA editing of human microRNAs. *Genome Biol.* **2006**, *7*, R27.
91. Kawahara, Y.; Zinshteyn, B.; Sethupathy, P.; Iizasa, H.; Hatzigeorgiou, A.G.; Nishikura, K. Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* **2007**, *315*, 1137–1140.
92. Kawahara, Y.; Megraw, M.; Kreider, E.; Iizasa, H.; Valente, L.; Hatzigeorgiou, A.G.; Nishikura, K. Frequency and fate of microRNA editing in human brain. *Nucleic Acids Res.* **2008**, *36*, 5270–5280.
93. Vesely, C.; Tauber, S.; Sedlazeck, F.J.; von Haeseler, A.; Jantsch, M.F. Adenosine deaminases that act on RNA induce reproducible changes in abundance and sequence of embryonic miRNAs. *Genome Res.* **2012**, *22*, 1468–1476.
94. Alon, S.; Mor, E.; Vigneault, F.; Church, G.M.; Locatelli, F.; Galeano, F.; Gallo, A.; Shomron, N.; Eisenberg, E. Systematic identification of edited microRNAs in the human brain. *Genome Res.* **2012**, *22*, 1533–1540.
95. Warf, M.B.; Shepherd, B.A.; Johnson, W.E.; Bass, B.L. Effects of ADARs on small RNA processing pathways in *C. elegans*. *Genome Res.* **2012**, *22*, 1488–1498.
96. Ekdahl, Y.; Farahani, H.S.; Behm, M.; Lagergren, J.; Ohman, M. A-to-I editing of microRNAs in the mammalian brain increases during development. *Genome Res.* **2012**, *22*, 1477–1487.
97. Heale, B.S.; Keegan, L.P.; McGurk, L.; Michlewski, G.; Brindle, J.; Stanton, C.M.; Caceres, J.F.; O'Connell, M.A. Editing independent effects of ADARs on the miRNA/siRNA pathways. *EMBO J.* **2009**, *28*, 3145–3156.
98. Yang, W.; Chendrimada, T.P.; Wang, Q.; Higuchi, M.; Seeburg, P.H.; Shiekhattar, R.; Nishikura, K. Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat. Struct. Mol. Biol.* **2006**, *13*, 13–21.
99. Caudy, A.A.; Ketting, R.F.; Hammond, S.M.; Denli, A.M.; Bathoorn, A.M.; Tops, B.B.; Silva, J.M.; Myers, M.M.; Hannon, G.J.; Plasterk, R.H. A micrococcal nuclease homologue in RNAi effector complexes. *Nature* **2003**, *425*, 411–414.
100. Weissbach, R.; Scadden, A.D. Tudor-SN and ADAR1 are components of cytoplasmic stress granules. *RNA* **2012**, *18*, 462–471.
101. Kawahara, Y.; Zinshteyn, B.; Chendrimada, T.P.; Shiekhattar, R.; Nishikura, K. RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. *EMBO Rep.* **2007**, *8*, 763–769.
102. Cullen, B.R. Viral and cellular messenger RNA targets of viral microRNAs. *Nature* **2009**, *457*, 421–425.
103. Skalsky, R.L.; Cullen, B.R. Viruses, microRNAs, and host interactions. *Annu. Rev. Microbiol.* **2010**, *64*, 123–141.

104. Iizasa, H.; Wulff, B.E.; Alla, N.R.; Maragkakis, M.; Megraw, M.; Hatzigeorgiou, A.; Iwakiri, D.; Takada, K.; Wiedmer, A.; Showe, L.; Lieberman, P.; Nishikura, K. Editing of Epstein-Barr virus-encoded BART6 microRNAs controls their dicer targeting and consequently affects viral latency. *J. Biol. Chem.* **2010**, *285*, 33358–33370.
105. Lei, T.; Yuen, K.S.; Tsao, S.W.; Chen, H.; Kok, K.H.; Jin, D.Y. Perturbation of biogenesis and targeting of Epstein-Barr virus-encoded miR-BART3 microRNA by A-to-I editing. *J. Gen. Virol.* **2013**, doi:10.1099/vir.0.056226-0.
106. Gandy, S.Z.; Linnstaedt, S.D.; Muralidhar, S.; Cashman, K.A.; Rosenthal, L.J.; Casey, J.L. RNA editing of the human herpesvirus 8 kaposin transcript eliminates its transforming activity and is induced during lytic replication. *J. Virol.* **2007**, *81*, 13544–13551.
107. Pfeffer, S.; Sewer, A.; Lagos-Quintana, M.; Sheridan, R.; Sander, C.; Grasser, F.A.; van Dyk, L.F.; Ho, C.K.; Shuman, S.; Chien, M.; *et al*. Identification of microRNAs of the herpesvirus family. *Nat. Methods* **2005**, *2*, 269–276.
108. Choudhury, Y.; Tay, F.C.; Lam, D.H.; Sandanaraj, E.; Tang, C.; Ang, B.T.; Wang, S. Attenuated adenosine-to-inosine editing of microRNA-376a* promotes invasiveness of glioblastoma cells. *J. Clin. Investig.* **2012**, *122*, 4059–4076.
109. Maas, S.; Patt, S.; Schrey, M.; Rich, A. Underediting of glutamate receptor GluR-B mRNA in malignant gliomas. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 14687–14692.
110. Galeano, F.; Rossetti, C.; Tomaselli, S.; Cifaldi, L.; Lezzerini, M.; Pezzullo, M.; Boldrini, R.; Massimi, L.; Di Rocco, C.M.; Locatelli, F.; Gallo, A. ADAR2-editing activity inhibits glioblastoma growth through the modulation of the CDC14B/Skp2/p21/p27 axis. *Oncogene* **2012**, *32*, 998–1009.
111. Liang, H.; Landweber, L.F. Hypothesis: RNA editing of microRNA target sites in humans? *RNA* **2007**, *13*, 463–467.
112. Peng, Z.; Cheng, Y.; Tan, B.C.; Kang, L.; Tian, Z.; Zhu, Y.; Zhang, W.; Liang, Y.; Hu, X.; Tan, X.; *et al*. Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. *Nat. Biotechnol.* **2012**, *30*, 253–260.
113. Gu, T.; Buaas, F.W.; Simons, A.K.; Ackert-Bicknell, C.L.; Braun, R.E.; Hibbs, M.A. Canonical A-to-I and C-to-U RNA editing is enriched at 3'UTRs and microRNA target sites in multiple mouse tissues. *PLoS One* **2012**, *7*, e33720.
114. Wang, Q.; Hui, H.; Guo, Z.; Zhang, W.; Hu, Y.; He, T.; Tai, Y.; Peng, P.; Wang, L. ADAR1 regulates ARHGAP26 gene expression through RNA editing by disrupting miR-30b-3p and miR-573 binding. *RNA* **2013**, *19*, 1525–1536.
115. De Hoon, M.J.; Taft, R.J.; Hashimoto, T.; Kanamori-Katayama, M.; Kawaji, H.; Kawano, M.; Kishima, M.; Lassmann, T.; Faulkner, G.J.; Mattick, J.S.; *et al*. Cross-mapping and the identification of editing sites in mature microRNAs in high-throughput sequencing libraries. *Genome Res.* **2010**, *20*, 257–264.
116. Chiang, H.R.; Schoenfeld, L.W.; Ruby, J.G.; Auyeung, V.C.; Spies, N.; Baek, D.; Johnston, W.K.; Russ, C.; Luo, S.; Babiarz, J.E.; *et al*. Mammalian microRNAs: Experimental evaluation of novel and previously annotated genes. *Genes Dev.* **2010**, *24*, 992–1009.
117. Galeano, F.; Tomaselli, S.; Locatelli, F.; Gallo, A. A-to-I RNA editing: The "ADAR" side of human cancer. *Semin. Cell Dev. Biol.* **2012**, *23*, 244–250.

Reprinted from *IJMS*. Cite as: Li, J.; Xuan, Z.; Liu, C. Long Non-Coding RNAs and Complex Human Diseases. *Int. J. Mol. Sci.* **2013**, *14*, 18790-18808.

Review

Long Non-Coding RNAs and Complex Human Diseases

Jing Li [1,†], Zhenyu Xuan [2,*] and Changning Liu [1,*]

[1] Bioinformatics Research Group, Key Laboratory of Intelligent Information Processing, Advanced Computer Research Center, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China; E-Mail: lij.thu@gmail.com

[2] Department of Molecular and Cell Biology, Center for Systems Biology, University of Texas at Dallas, 800 W Campbell Road, Richardson, TX 75080, USA

[†] Present address: School of Medicine, Tsinghua University, Beijing 100084, China.

[*] Authors to whom correspondence should be addressed: E-Mails: zhenyu.xuan@utdallas.edu (Z.X.); lcn@ict.ac.cn (C.L.); Tel.: +1-972-883-2518 (Z.X.); +86-10-6260-1010 (C.L.); Fax: +1-972-883-5710 (Z.X.); +86-10-6260-1356 (C.L.).

Received: 20 June 2013; in revised form: 28 August 2013 / Accepted: 3 September 2013 / Published: 12 September 2013

Abstract: Long non-coding RNAs (lncRNAs) are a heterogeneous class of RNAs that are generally defined as non-protein-coding transcripts longer than 200 nucleotides. Recently, an increasing number of studies have shown that lncRNAs can be involved in various critical biological processes, such as chromatin remodeling, gene transcription, and protein transport and trafficking. Moreover, lncRNAs are dysregulated in a number of complex human diseases, including coronary artery diseases, autoimmune diseases, neurological disorders, and various cancers, which indicates their important roles in these diseases. Here, we reviewed the current understanding of lncRNAs, including their definition and subclassification, regulatory functions, and potential roles in different types of complex human diseases.

Keywords: non-coding RNA; long non-coding RNA; complex human disease

---

1. Introduction

While about 20,000 protein-coding genes, representing less than 2% of the human genome, have been reported [1,2], a large part of the genome can be transcribed into non-coding RNAs (ncRNAs),

which have little or no protein-coding capability [3,4]. Besides many widely studied classes of short ncRNAs, such as microRNAs (miRNA) and Piwi-interacting RNAs (piRNA) [5,6], one class of heterogeneous ncRNAs with lengths longer than 200 nucleotides, recently designated as long non-coding RNAs (lncRNAs), is increasingly attracting the attention of ncRNA researchers [7,8].

Over the past decade, furthered by the rapid progress in high-throughput sequencing technology, thousands of lncRNAs have been identified in mammalian transcriptomes [9,10]. A number of studies have revealed that lncRNAs can participate in various critical biological processes, such as chromatin remodeling, gene transcription, and protein transport and trafficking [11,12], implicating their impact on a wide range of complex human diseases [13,14]. However, despite some well-characterized lncRNAs, such as Xist and HOTAIR [15,16], little is known about the general features of most lncRNAs, such as gene structure, transcriptional regulation, and functional domains, and even less is known about their possible molecular mechanisms in different human diseases. Understanding the function of lncRNAs remains a significant challenge. Here, we review the current literature reporting on lncRNAs, including their definition, subclassification, regulatory functions, and roles in different types of complex human diseases.

2. Long Non-Coding RNAs: A Heterogeneous Class of RNAs

Long non-coding RNAs are most commonly defined as an RNA transcript more than 200 nucleotides (nt) long that cannot be translated into a protein [17]. However, this length threshold is not strict; it may vary 100 to 200 nt, or even longer. In different biochemical fractionation protocols, this threshold is primarily used to exclude most of the categories of small RNAs, such as small nucleolar RNA (snoRNA), microRNA (miRNA), Piwi-interacting RNA (piRNA), transfer RNA (tRNA), and small nuclear RNA (snRNA). Therefore, according to these simple criteria of transcript size and protein-coding capability, the designated lncRNAs contain a group of structurally and functionally heterogeneous RNAs, having a length that varies from approximately 200 nt to over 100 kb. They can undergo either splicing or not, with cellular locations in either nucleus or cytoplasm. They can be transcribed by RNA polymerase II or III, and they can play different functional and structural roles in different biological processes [18]. Based on these features, lncRNAs can be further categorized into different subgroups, as listed in Table 1. In past years, along with in-depth studies of lncRNAs, several lncRNA databases have been constructed (Table 2). These databases can facilitate further functional research on lncRNAs.

Table 1. Types of long non-coding RNAs.

Long non-coding RNA	Symbol	References
Long intergenic non-coding RNA	LincRNA	[19,20]
Long intronic non-coding RNA		[14,21]
Natural antisense transcript	NAT	[22–24]
Promoter-associated long RNA	PALR	[25]
Promoter upstream transcript	PROMPT	[26]
Repetitive element-associated non-coding RNA		[27–29]
Transcribed pseudogene		[30,31]
Transcribed ultraconserved region	T-UCR	[32]
Enhancer-like non-coding RNA	eRNA	[33]

Table 2. Public lncRNA databases.

Name	Website	Reference
ChIPBase	http://deepbase.sysu.edu.cn/chipbase/	[34]
fRNAdb	http://www.ncrna.org/frnadb/	[35]
LNCipedia	http://www.lncipedia.org/	[36]
lncRNAdb	http://www.lncrnadb.org/	[37]
NONCODE	http://www.noncode.org/	[9]
NRED	http://nred.matticklab.com/cgi-bin/ncrnadb.pl	[38]

Genome-wide transcriptional maps have shown that lncRNAs are pervasively transcribed throughout mammalian genomes [39,40]. Clusters of overlapping sense and antisense transcripts can be found inside of known genes, as well as in intergenic regions. Natural antisense transcripts (NATs), which have been largely discovered in human, mouse, and many other species, are endogenous RNA molecules that exhibit partial or complete complementarities to other transcripts [22–24]. NATs may regulate the expression level of their sense counterparts. Several plausible regulation models, like blocking translation by sense-antisense pairing or antisense RNA-directed chromatin remodeling, have been proposed [13,41]. However, little mechanistic information has supported these suppositions, and more intensive studies are needed. Unlike NATs, long intergenic non-coding RNAs (lincRNAs) are large multiexonic RNAs, which are transcribed from intergenic regions, and may act in trans within large ribonucleoprotein complexes [10,19,42]. For example, Huarte *et al*. have reported that lincRNA-p21, which can physically interact with hnRNP-K, serves as a repressor in p53-dependent transcriptional responses by modulating hnRNP-K localization to chromatin [43]. A recent comprehensive screen has identified dozens of lincRNAs, which bind to multiple chromatin regulatory proteins, to affect related gene expression procedures and function critically in the pathway controlling pluripotent embryonic stem cell (ES) state [20]. Long non-coding RNAs have also been transcribed from widespread repetitive elements. For instance, human Alu and mouse B2 RNAs are originally derived from short interspersed repeat elements (SINEs). They are transcribed by RNA polymerase III in response to environmental stresses, such as heat shock, and act as transcriptional repressors by directly targeting RNA polymerase II [27–29]. Another interesting example is lncRNA PTENP1, a biologically active pseudogene of the tumor suppressor gene Phosphatase and Tensin Homolog (*PTEN*). PTENP1 performs a tumor suppressive function by acting as an "endogenous miRNA sponge", which can positively regulate PTEN protein level via competing for PTEN-targeting miRNA binding [30,31].

Unlike protein-coding genes, which are usually conserved across species, most lncRNAs are poorly conserved and have been taken for transcriptional noise [44]. However, lack of conservation does not mean lack of function [45,46]. For example, two lncRNAs, Antisense Igf2r RNA (Air) and X Specific Transcript (Xist), are poorly conserved, but are well functional [47,48]. Their subjection to a series of recent and rapid adaptive selections may provide one explanation for the poor conservation of lncRNAs. For example, Highly Accelerated Region 1F (HAR1F), an lncRNA which is exclusively expressed in Cajal-Retzius neurons in the human neocortex, has undergone rapid evolutionary change in the human lineage since our last common ancestor of chimpanzee [49]. Moreover, for lncRNAs, which exert functions by secondary structures or short sequence motifs, we may only find small

conserved regions interspersed in long poorly conserved transcripts [50]. Besides those poorly conserved lncRNAs, non-coding transcripts can also be transcribed from ultraconserved genomic regions (UCRs). UCRs were first discovered in the sequence comparison of mouse, rat, and human genomes [51]. They are genomic elements longer than 200 bp with 100% identity between orthologous regions in these three genomes [52,53]. Genomic variations in UCRs, such as single nucleotide polymorphisms (SNPs), have been reported to be associated with increased cancer risk [54]. Genome-wide expression profiling has revealed that a large fraction of UCRs are transcribed (transcribed-UCRs, T-UCRs) with significant alteration at both DNA and RNA levels in adult chronic lymphocytic leukemias, as well as colorectal and hepatocellular carcinomas [32,55]. Recent study in neuroblastoma has discovered the relevance between expression levels of specific T-UCRs and important clinical-genetic parameters, suggesting that T-UCRs may be used as signatures associated with cancer diagnosis, prognosis, and treatment [55].

As a heterogeneous class of RNAs, lncRNAs have been implicated in the regulation of a number of complicated biological processes. A study in human cell lines suggests that about 30% of lncRNAs are specifically expressed in the nucleus [25]. Many of them are involved in chromatin remodeling complexes and mediate genomic silencing [10]. One of the most well-known examples is the participation of lncRNAs in X-chromosome inactivation (XCI), a process by which one of the two copies of the X chromosome present in female mammals is inactivated [15]. During XCI, Inactive X Specific Transcript (Xist), a 17-kb lncRNA transcribed from the XCI center, will accumulate on the inactive X chromosome and recruit Polycomb complexes for subsequent epigenetic modifications. Its antisense counterpart, Tsix, which is another lncRNA specifically expressed from the other X chromosome, can also interact with Polycomb complexes and maintain the activity of X chromosome [48]. Apart from in *trans* regulation, some lncRNAs can directly regulate gene expression in *cis*. By using knockdown approaches and reporter assays, Orom *et al*. have discovered an enhancer-like effect for a set of lncRNAs in human cell lines [56]. Depletion of these lncRNAs leads to decreased expression of their neighboring protein-coding genes. More interestingly, a recent genome-wide study of transcriptional enhancers in mouse has shown that some lncRNAs, termed as enhancer-RNAs (eRNAs), are transcribed from functional enhancers [33]. Although the function of eRNAs remains largely unclear, their close correlation with active enhancers suggests an important role of eRNAs in transcriptional activation. Long non-coding RNAs are also actively involved in diverse cytoplasmic processes. One of the well-studied examples is noncoding repressor of *NFAT* (NRON), an lncRNA repressor of nuclear factor of activated T cells (*NFAT*). Using an arrayed library of short hairpin RNAs and cell-based assays, Willingham *et al*. have identified that NRON interacts with multiple proteins, including members of the importin-beta superfamily, and possibly functions as a specific regulator of *NFAT* nuclear trafficking [57]. Moreover, lncRNAs are probably involved in stress-related signaling pathways. Utilizing whole-genome tiling arrays, Silva *et al*. identified a new class of long stress responsive non-coding transcripts (LSINCTs), which have increased expression in response to DNA damage induced by the tobacco carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) [58]. Interestingly, LSINCTs also have increased expression in a number of cancer-derived cell lines, indicating its stress response under a carcinogenic environment.

3. Long Non-Coding RNAs and Complex Human Diseases

Complex diseases are multifactorial or polygenic disorders of the body. They are likely caused by multiple genetic variants with low penetrance in combination with various environmental and lifestyle factors, and they do not simply obey the standard Mendelian patterns of inheritance [59,60]. Coronary artery diseases, autoimmune diseases, neurological disorders, various cancers, and many other diseases all belong within this classification. The recent discovery that lncRNAs can participate in a wide range of biological processes has attracted substantial scientific interest in their potential impact on these complex diseases. It has been reported that lncRNAs are dysregulated in a variety of complex human diseases and are closely associated with disease development and progression (Table 3). Here, we describe some of the well-characterized lncRNAs associated with different types of complex human diseases.

Table 3. Examples of lncRNAs dysregulated in complex human diseases.

LncRNA	Disease	References
aHIF	Multiple cancers	[61,62]
AK023948	Papillary thyroid carcinoma	[63]
ANRIL	Coronary artery disease; Multiple cancers	[64–68]
ASFMR1	Fragile X syndrome; Fragile X tremor ataxia syndrome	[69]
ATXN8OS	Spinocerebellar ataxia type 8	[70]
BACE1-AS	Alzheimer's disease	[71]
BC200	Alzheimer's disease; Multiple cancers	[72–74]
BIC	B-cell lymphoma	[75]
CUDR	Squamous carcinoma	[76]
DD3	Prostate cancer	[77,78]
FMR4	Fragile X syndrome; Fragile X tremor ataxia syndrome	[79]
GAS5	Breast cancer	[80]
H19	Multiple cancers	[81–85]
HOTAIR	Multiple cancers	[16,86]
HULC	Multiple cancers	[87,88]
Kcnq1ot1	Colon cancer	[89]
Kras1p	Prostate cancer	[30]
Linc-p21	Lung cancer	[43]
LOC285194	Osteosarcoma	[90]
MALAT-1	Multiple cancers	[91–93]
MEG3	Multiple cancers	[94–97]
MIAT	Myocardial infarction	[98]
ncRAN	Neuroblastoma	[99,100]
NDM29	Neuroblastoma	[101]
PCGEM1	Prostate cancer	[102]
PRINS	Psoriasis	[103]
PRNCR1	Prostate cancer	[104]
PTENP1	Prostate cancer	[30]
RMRP	Leukemia and lymphoma	[105]
SAS-ZFAT	Autoimmune thyroid disease	[106]

Table 3. *Cont.*

LncRNA	Disease	References
SPRY4-IT1	Melanoma	[107]
SRA-1	Breast cancer	[108,109]
TERC	Multiple cancers	[110]
Ube3a-as	Angelman syndrome	[111,112]
uc.73A	Colon cancer	[32]
UCA1	Bladder cancer	[113–115]
Zfas1	Breast cancer	[116]

3.1. LncRNAs in Coronary Artery Diseases

Facilitated by single nucleotide polymorphism (SNP) array, genome-wide association study (GWAS) is becoming one of the most powerful approaches to identify genetic variants susceptible to common diseases [117]. From these studies, a large amount of disease-associated SNPs are found to be mapped to non-coding genomic regions. While some of these SNPs could be associated with enhancers, it would not be surprising that many others are associated with lncRNAs. By using 52,608 haplotype-based SNP markers, the lncRNA called Myocardial Infarction Associated Transcript (MIAT) was first identified in a large-scale case-control association study of the samples from 3435 MI patients and 3774 controls [98]. MIAT dwells at a susceptible locus for myocardial infarction (MI) on chromosome 22q12.1. This study discovered six SNPs showing significant association with MI in this locus. The MIAT transcript is approximately 10 kb in length and has five exons. No translational product is encoded in MIAT based on *in vitro* translation assay, which indicates that it is an lncRNA. From *in vitro* functional analyses, MIAT transcription is increased by the minor variant of one SNP in exon 5. In contrast to the non-risk allele, the risk allele has more intense binding of nuclear protein(s). The study concluded that MIAT may play some roles in the pathogenesis of MI with altered expression by SNP [98].

Recent GWASs have identified a region on chromosome 9p that is associated with coronary artery disease (CAD) [118,119]. A long non-coding antisense RNA gene, named as Antisense non-coding RNA in the *INK4* locus (ANRIL), is a prime candidate for the chromosome 9p CAD locus [64]. Several recent GWASs showed that ANRIL has increased susceptibility to intracranial aneurysm, breast cancer, glioma, and basal cell carcinomas [65–68]. ANRIL is located in the *INK4b*/*ARF*/*INK4a* locus, and it is coregulated with *INK4a*, *INK4b* and *ARF*. Expression studies have confirmed that ANRIL is expressed in multiple atherosclerosis-related cell lines, including vascular endothelial cell, monocyte-derived macrophages and coronary smooth muscle cells [64,120]. Moreover, a mouse model study has demonstrated the pivotal role of ANRIL in the regulation of *INK4a/b* expression through a *cis*-acting mechanism and its implication in proliferation and senescence [121,122]. Interestingly, several studies discovered that Polycomb complexes are able to bind the *INK4*/*ARF* locus and alter expression of *INK4a* and *INK4b* [123,124]. Similar to lncRNA XIST, the ANRIL gene presents unusual masses of repetitive elements, as well as many binding sites for repressive transcription factors [120]. All these observations suggest that ANRIL, much like XIST, may regulate the expression of the INK4a/b transcript by recruiting Polycomb complex to the *INK4*/*ARF* locus and imposing a repressive chromatin state.

3.2. LncRNAs in Autoimmune Diseases

Long non-coding RNAs may also function in the regulation of downstream protein-coding genes, thus forming a complicated mutual regulation network with both coding and non-coding genes [125,126]. Recent studies have shown that autoimmune diseases, which result from an inappropriate immune response of the body against substances and tissues normally present in the body, have a complex genetic context that involves multiple protein-coding and non-coding genes. For example, in association study of 515 affected individuals and 526 controls, Shirasawa *et al*. discovered that the T allele of SNP Ex9b-SNP10 is correlated to increased risk for autoimmune thyroid disease (AITD) [106]. The Ex9b-SNP10 resides in intron 9 of the protein-coding gene *ZFAT* and the promoter region of an lncRNA, SAS-ZFAT, which is an antisense transcript of the ZFAT gene. With the existence of SNP Ex9b-SNP10, SAS-ZFAT expression is evidently upregulated, which, in turn, downregulates the expression level of its sense counterpart—truncated *ZFAT*. Since SAS-ZFAT is exclusively expressed in CD19+ B cells in peripheral blood lymphocytes, these results implicated that SAS-ZFAT might play a critical role in B cell function and determine susceptibility to AITD.

Another example is lncRNA PRINS, a Psoriasis Susceptibility-related non-coding RNA gene that harbors two Alu elements [103]. PRINS is transcribed by RNA polymerase II and is expressed at different levels in various human tissues. Real-time RT-PCR analysis showed that PRINS has higher expression in the uninvolved epidermis of psoriatic patients than in both psoriatic lesional and healthy epidermis, suggesting that PRINS plays a role in psoriasis susceptibility. *In silico* structural and homology studies have suggested that PRINS acts as a non-coding RNA. The RNA expression level of PRINS is decreased in the uninvolved psoriatic, but not healthy, epidermis with treatment of T-lymphokines that are known to precipitate psoriatic symptoms. Moreover, downregulating the RNA level of PRINS by RNAi can impair cell viability after serum starvation, but not under normal serum conditions. It was also discovered that PRINS could function as a "riboregulator" to regulate the expression of other genes involved in the proliferation and survival of cells.

3.3. LncRNAs in Neurological Disorders

Neurological disorders are diseases of the body's nervous systems, which include the central nervous system, the peripheral nervous system, and the autonomic nervous system. Previous transcriptome studies have shown a number of lncRNAs in the mammalian brain, and most of them exhibit particular expression profiles within specific neuroanatomical regions, cell types, or subcellular compartments, implicating that lncRNAs probably have a significant impact on neurological regulation [127].

Long non-coding RNAs may participate in the pathogenesis of fragile X syndrome (FXS) and fragile X tremor ataxia syndrome (FXTAS), both of which are caused by the aberrant expansion of CGG trinucleotide, repeat in the 5' UTR of protein-coding fragile-X mental retardation 1 gene (*FMR1*) [128,129]. Studies have shown that two lncRNAs, FMR4 and ASFMR1, are expressed from the *FMR1* locus. FMR4 is a primate-specific lncRNA that likely shares a bidirectional promoter with *FMR1* [79], while ASFMR1 is a spliced and poly-adenylated antisense transcript that overlaps the 5' UTR CGG repeat region of *FMR1* [69]. *In vitro* studies of FMR4 have shown that it may function to prevent neurons or their progenitors from apoptosis during the progress of development in human. FMR4 and

ASFMR1, as well as *FMR1*, may participate synergistically in neurological regulation in a RNA-protein interacting manner, since they are all silenced in FXS or upregulated in FXTAS patients. Dysregulation of these delicate interactions may result in various neurological disorders [69,79].

Long non-coding RNAs have also been reported to be dysregulated in different types of neurodegenerative diseases, such as spinocerebellar ataxia type 8 (SCA8) and Alzheimer's disease (AD). SCA8 is an autosomal dominant disorder caused by repeat expansion [130]. Two bidirectionally transcribed genes are located within the SCA8 expansion region: the protein-coding gene ataxin 8 (*ATXN8*), with a CAG expansion that encodes a polyglutamine expansion tract protein, and ataxin 8 opposite strand (ATXN8OS) which is an lncRNA with a CUG repeat. In studies of transgenic mice expressing SCA8 expansion, it was found that the ATXN8OS mutant is overexpressed and co-localized with Muscleblind-like splicing regulator 1 (*MBNL1*) in neurons, which can lead to dysregulation of MBNL1-mediated alternative splicing, loss of GABAergic inhibition within the granular cell layer, and set the stage for the occurrence of disease [70]. AD is a form of dementia, which is believed to be caused by the formation of amyloid plaques in neurons [131]. Studies have shown that BACE1-AS, an antisense lncRNA counterpart of protein-coding gene *BACE1*, is highly expressed in tissues from AD patients [71]. *BACE1* is an enzyme that is responsible for amyloid precursor protein (APP) cleavage into amyloid β peptides, which form amyloid plaques in the neurons of AD patients [132]. Upregulation of BACE1-AS promotes the stabilization of *BACE1* mRNA and boosts the expression of *BACE1* protein, which leads to the production of pathogenic amyloid β peptides and thus may speed up the pathogenesis of AD [71]. Another lncRNA involved in AD is BC200, which is expressed almost exclusively in neuronal cells. In AD patients, the level of BC200 becomes upregulated. The increased expression of BC200 was found to be correlated with the severity of AD [74]. In addition, studies of BC1, the mouse functional homolog of BC200, have shown that BC1 knockout mice exhibit behavioral changes, thus demonstrating an important role for BC1 in brain function [133]. All these results suggest that dysregulation of BC200 may contribute to AD susceptibility.

3.4. LncRNAs in Cancers

Cancer is a broad group of various diseases in which abnormal cells divide uncontrollably and tend to invade other tissues. Up to now, although hundreds of oncogenes and tumor suppressor genes have been identified, the exact cause of most cancers remains unknown or poorly understood. In recent years, researchers have increasingly come to recognize lncRNAs as major mediators in cancer pathogenesis [134]. Thus far, no concrete evidence has surfaced to indicate any lncRNAs as causal factors in cancer. However, many lncRNAs have been found to be differentially expressed in a variety of cancers and may act as either oncogenes, such as MALAT-1, HOTAIR, and ANRIL, or tumor suppressor genes, such as MEG3, lincRNA-p21, and PTENP1, in cancer development. Here, we discuss some examples of such lncRNAs.

Like protein-coding oncogenes, some lncRNAs can promote cell proliferation and induce tumorigenesis. Metastasis-associated lung adenocarcinoma transcript 1 (MALAT-1), which correlates with high metastasis and poor prognosis in non-small-cell lung cancer, is an abundant 8.7-kb lncRNA encoded in the human chromosome 11q13 [93]. MALAT-1 is broadly expressed in normal human tissues and is found to be upregulated in many solid tumors, such as lung, breast, prostate, liver, and

colon tumors [91,92,135]. MALAT-1 is believed to play a vital role in cell proliferation, migration, and invasion. By interacting with serine-arginine-rich splicing factor (*SR*), which is responsible for alternative splicing (AS) in a concentration- and phosphorylation-dependent manner, studies have shown that MALAT-1 can modulate the phosphorylation of *SR* proteins and thus regulate AS of selective pre-mRNAs [136]. MALAT1 is also involved in the regulation of cell mobility. RNAi-mediated silencing of MALAT1 impaired the *in vitro* migration of lung adenocarcinoma cells and reduced cell proliferation and invasive potential in a cervical cancer cell line [92]. Another onco-lncRNA example is *HOX* Antisense Intergenic RNA (HOTAIR). HOTAIR, a 2.2-kb spliced and poly-adenylated lncRNA, is transcribed from the antisense strand of the Homeobox C (*HOXC*) gene cluster on chromosome 12 [16]. Studies have shown that *HOTAIR* is overexpressed in breast tumors, hepatocellular carcinoma, and colorectal cancer [137–139]. A high level of HOTAIR expression is directly correlated with poor patient prognosis and metastasis. Recent studies revealed that HOTAIR is likely to work as a molecular scaffold to bind two distinct histone modification complexes, the Polycomb repressor complex 2 (PRC2) and the histone demethylase LSD1, facilitating their genome-wide retargeting to specific regions for coupled histone H3K27 methylation and H3K4 demethylation [140]. *In vitro* studies have shown that overexpression of HOTAIR in cell lines leads to the recruitment of PRC2 and LSD1 to over 800 additional loci, including those of tumor suppressor genes [16]. These observations indicate that dysregulation of HOTAIR may reprogram the epigenetic information to promote tumor cell invasion and subsequent metastasis.

Long non-coding RNAs can also act as tumor suppressor genes. One example is maternally expressed gene 3 (MEG3), a maternally imprinted RNA gene of approximately 1700 nucleotides [94]. Studies have revealed that MEG3 is expressed in many normal tissues, but not in the majority of human meningiomas or human meningioma cell lines [141]. Moreover, ectopic expression of MEG3 was found to suppress the growth of several human cancer cell lines, further supporting the effect of MEG3 on tumor suppression [95]. MEG3 was found to be a positive regulator of *p53*, a tumor suppressor protein [142]. In cells that are transfected with MEG3, *p53* protein level increases significantly, which results in dramatically stimulating the transcription of p53-dependent genes from a p53-responsive promoter. Studies have shown that MEG3 is also capable of inhibiting cell proliferation in the absence of *p53* [143]. These data suggest that MEG3 can function as a tumor suppressor through both p53-dependent and p53-independent pathways. MEG3 has a total of twelve isoforms from alternative splicing, all of which contain three distinct secondary folding motifs (M1, M2, and M3). Deletion analysis indicates that motifs M2 and M3 are important for *p53* activation. Furthermore, a hybrid MEG3 RNA, which contains a piece of unrelated sequence, but preserves the original secondary structure, retained the functions of both *p53* activation and growth suppression [144]. As a regulatory lncRNA, all of these experiments demonstrated that the proper conformation of MEG3 is critical to its biological functions.

4. Conclusions

Long non-coding RNAs are rapidly becoming a focal point for intensified research in the biological and medical sciences. Increasing evidence has indicated that lncRNAs play important roles in various critical biological processes and that they add a new layer of complexity to already complex human diseases. We believe that the further functional and mechanistic studies of these versatile macromolecules will expand our understanding of general principles in biological systems and provide new approaches to the diagnosis and treatment of complex human diseases.

Acknowledgments

This work was supported by National Natural Science Foundation of China (Grant No. 31000586) and The University of Texas at Dallas Startup Grant.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Consortium, E.P.; Dunham, I.; Kundaje, A.; Aldred, S.F.; Collins, P.J.; Davis, C.A.; Doyle, F.; Epstein, C.B.; Frietze, S.; Harrow, J.; *et al*. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74.
2. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; *et al*. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921.
3. Liu, C.; Bai, B.; Skogerbo, G.; Cai, L.; Deng, W.; Zhang, Y.; Bu, D.; Zhao, Y.; Chen, R. Noncode: An integrated knowledge database of non-coding RNAs. *Nucleic Acids Res.* **2005**, *33*, D112–D115.
4. Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M.C.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C.; *et al*. The transcriptional landscape of the mammalian genome. *Science* **2005**, *309*, 1559–1563.
5. Lau, N.C.; Seto, A.G.; Kim, J.; Kuramochi-Miyagawa, S.; Nakano, T.; Bartel, D.P.; Kingston, R.E. Characterization of the pirna complex from rat testes. *Science* **2006**, *313*, 363–367.
6. David, R. Small RNAs: MiRNA machinery disposal. *Nat. Rev.* **2013**, *14*, 4–5.
7. Wilusz, J.E.; Sunwoo, H.; Spector, D.L. Long noncoding RNAs: Functional surprises from the RNA world. *Genes Dev.* **2009**, *23*, 1494–1504.
8. Liao, Q.; Liu, C.; Yuan, X.; Kang, S.; Miao, R.; Xiao, H.; Zhao, G.; Luo, H.; Bu, D.; Zhao, H.; *et al*. Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network. *Nucleic Acids Res.* **2011**, *39*, 3864–3878.
9. Bu, D.; Yu, K.; Sun, S.; Xie, C.; Skogerbo, G.; Miao, R.; Xiao, H.; Liao, Q.; Luo, H.; Zhao, G.; *et al*. Noncode v3.0: Integrative annotation of long noncoding RNAs. *Nucleic Acids Res.* **2012**, *40*, D210–D215.

10. Khalil, A.M.; Guttman, M.; Huarte, M.; Garber, M.; Raj, A.; Rivea Morales, D.; Thomas, K.; Presser, A.; Bernstein, B.E.; van Oudenaarden, A.; *et al*. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11667–11672.
11. Mercer, T.R.; Mattick, J.S. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat. Struct. Mol. Biol.* **2013**, *20*, 300–307.
12. Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914.
13. Tufarelli, C.; Stanley, J.A.; Garrick, D.; Sharpe, J.A.; Ayyub, H.; Wood, W.G.; Higgs, D.R. Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat. Genet.* **2003**, *34*, 157–165.
14. Tahira, A.C.; Kubrusly, M.S.; Faria, M.F.; Dazzani, B.; Fonseca, R.S.; Maracaja-Coutinho, V.; Verjovski-Almeida, S.; Machado, M.C.; Reis, E.M. Long noncoding intronic RNAs are differentially expressed in primary and metastatic pancreatic cancer. *Mol. Cancer* **2011**, *10*, 141.
15. Chow, J.; Heard, E. X inactivation and the complexities of silencing a sex chromosome. *Curr. Opin. Cell Biol.* **2009**, *21*, 359–366.
16. Gupta, R.A.; Shah, N.; Wang, K.C.; Kim, J.; Horlings, H.M.; Wong, D.J.; Tsai, M.C.; Hung, T.; Argani, P.; Rinn, J.L.; *et al*. Long non-coding RNA hotair reprograms chromatin state to promote cancer metastasis. *Nature* **2010**, *464*, 1071–1076.
17. Kung, J.T.; Colognori, D.; Lee, J.T. Long noncoding RNAs: Past, present, and future. *Genetics* **2013**, *193*, 651–669.
18. Ponting, C.P.; Oliver, P.L.; Reik, W. Evolution and functions of long noncoding RNAs. *Cell* **2009**, *136*, 629–641.
19. Guttman, M.; Amit, I.; Garber, M.; French, C.; Lin, M.F.; Feldser, D.; Huarte, M.; Zuk, O.; Carey, B.W.; Cassady, J.P.; *et al*. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **2009**, *458*, 223–227.
20. Guttman, M.; Donaghey, J.; Carey, B.W.; Garber, M.; Grenier, J.K.; Munson, G.; Young, G.; Lucas, A.B.; Ach, R.; Bruhn, L.; *et al*. LincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **2011**, *477*, 295–300.
21. Louro, R.; Smirnova, A.S.; Verjovski-Almeida, S. Long intronic noncoding RNA transcription: Expression noise or expression choice? *Genomics* **2009**, *93*, 291–298.
22. Yin, Y.; Zhao, Y.; Wang, J.; Liu, C.; Chen, S.; Chen, R.; Zhao, H. Anticode: A natural sense-antisense transcripts database. *BMC Bioinf.* **2007**, *8*, 319.
23. Katayama, S.; Tomaru, Y.; Kasukawa, T.; Waki, K.; Nakanishi, M.; Nakamura, M.; Nishida, H.; Yap, C.C.; Suzuki, M.; Kawai, J.; *et al*. Antisense transcription in the mammalian transcriptome. *Science* **2005**, *309*, 1564–1566.
24. Lehner, B.; Williams, G.; Campbell, R.D.; Sanderson, C.M. Antisense transcripts in the human genome. *Trends Genet.* **2002**, *18*, 63–65.
25. Kapranov, P.; Cheng, J.; Dike, S.; Nix, D.A.; Duttagupta, R.; Willingham, A.T.; Stadler, P.F.; Hertel, J.; Hackermuller, J.; Hofacker, I.L.; *et al*. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **2007**, *316*, 1484–1488.

26. Preker, P.; Nielsen, J.; Kammler, S.; Lykke-Andersen, S.; Christensen, M.S.; Mapendano, C.K.; Schierup, M.H.; Jensen, T.H. RNA exosome depletion reveals transcription upstream of active human promoters. *Science* **2008**, *322*, 1851–1854.
27. Liu, W.M.; Chu, W.M.; Choudary, P.V.; Schmid, C.W. Cell stress and translational inhibitors transiently increase the abundance of mammalian sine transcripts. *Nucleic Acids Res.* **1995**, *23*, 1758–1765.
28. Espinoza, C.A.; Goodrich, J.A.; Kugel, J.F. Characterization of the structure, function, and mechanism of b2 RNA, an ncRNA repressor of RNA polymerase II transcription. *RNA* **2007**, *13*, 583–596.
29. Mariner, P.D.; Walters, R.D.; Espinoza, C.A.; Drullinger, L.F.; Wagner, S.D.; Kugel, J.F.; Goodrich, J.A. Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. *Mol. Cell* **2008**, *29*, 499–509.
30. Poliseno, L.; Salmena, L.; Zhang, J.; Carver, B.; Haveman, W.J.; Pandolfi, P.P. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **2010**, *465*, 1033–1038.
31. Salmena, L.; Poliseno, L.; Tay, Y.; Kats, L.; Pandolfi, P.P. A ceRNA hypothesis: The rosetta stone of a hidden RNA language? *Cell* **2011**, *146*, 353–358.
32. Calin, G.A.; Liu, C.G.; Ferracin, M.; Hyslop, T.; Spizzo, R.; Sevignani, C.; Fabbri, M.; Cimmino, A.; Lee, E.J.; Wojcik, S.E.; *et al*. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. *Cancer Cell* **2007**, *12*, 215–229.
33. Kim, T.K.; Hemberg, M.; Gray, J.M.; Costa, A.M.; Bear, D.M.; Wu, J.; Harmin, D.A.; Laptewicz, M.; Barbara-Haley, K.; Kuersten, S.; *et al*. Widespread transcription at neuronal activity-regulated enhancers. *Nature* **2010**, *465*, 182–187.
34. Yang, J.H.; Li, J.H.; Jiang, S.; Zhou, H.; Qu, L.H. Chipbase: A database for decoding the transcriptional regulation of long non-coding RNA and microRNA genes from chip-seq data. *Nucleic Acids Res.* **2013**, *41*, D177–D187.
35. Mituyama, T.; Yamada, K.; Hattori, E.; Okida, H.; Ono, Y.; Terai, G.; Yoshizawa, A.; Komori, T.; Asai, K. The functional RNA database 3.0: Databases to support mining and annotation of functional RNAs. *Nucleic Acids Res.* **2009**, *37*, D89–D92.
36. Volders, P.J.; Helsens, K.; Wang, X.; Menten, B.; Martens, L.; Gevaert, K.; Vandesompele, J.; Mestdagh, P. Lncipedia: A database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Res.* **2013**, *41*, D246–D251.
37. Amaral, P.P.; Clark, M.B.; Gascoigne, D.K.; Dinger, M.E.; Mattick, J.S. LncRNAdb: A reference database for long noncoding RNAs. *Nucleic Acids Res.* **2011**, *39*, D146–D151.
38. Dinger, M.E.; Pang, K.C.; Mercer, T.R.; Crowe, M.L.; Grimmond, S.M.; Mattick, J.S. Nred: A database of long noncoding RNA expression. *Nucleic Acids Res.* **2009**, *37*, D122–D126.
39. Okazaki, Y.; Furuno, M.; Kasukawa, T.; Adachi, J.; Bono, H.; Kondo, S.; Nikaido, I.; Osato, N.; Saito, R.; Suzuki, H.; *et al*. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **2002**, *420*, 563–573.
40. Bertone, P.; Stolc, V.; Royce, T.E.; Rozowsky, J.S.; Urban, A.E.; Zhu, X.; Rinn, J.L.; Tongprasit, W.; Samanta, M.; Weissman, S.; *et al*. Global identification of human transcribed sequences with genome tiling arrays. *Science* **2004**, *306*, 2242–2246.

41. Knee, R.; Murphy, P.R. Regulation of gene expression by natural antisense RNA transcripts. *Neurochem. Inter.* **1997**, *31*, 379–392.
42. Cabili, M.N.; Trapnell, C.; Goff, L.; Koziol, M.; Tazon-Vega, B.; Regev, A.; Rinn, J.L. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **2011**, *25*, 1915–1927.
43. Huarte, M.; Guttman, M.; Feldser, D.; Garber, M.; Koziol, M.J.; Kenzelmann-Broz, D.; Khalil, A.M.; Zuk, O.; Amit, I.; Rabani, M., *et al*. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **2010**, *142*, 409–419.
44. Wang, J.; Zhang, J.; Zheng, H.; Li, J.; Liu, D.; Li, H.; Samudrala, R.; Yu, J.; Wong, G.K. Mouse transcriptome: Neutral evolution of "non-coding" complementary DNAs. *Nature* **2004**, *431*, doi:10.1038/nature03016.
45. Kutter, C.; Watt, S.; Stefflova, K.; Wilson, M.D.; Goncalves, A.; Ponting, C.P.; Odom, D.T.; Marques, A.C. Rapid turnover of long noncoding RNAs and the evolution of gene expression. *PLoS Genet.* **2012**, *8*, e1002841.
46. Pang, K.C.; Frith, M.C.; Mattick, J.S. Rapid evolution of noncoding RNAs: Lack of conservation does not mean lack of function. *Trends Genet.* **2006**, *22*, 1–5.
47. Nagano, T.; Mitchell, J.A.; Sanz, L.A.; Pauler, F.M.; Ferguson-Smith, A.C.; Feil, R.; Fraser, P. The air noncoding RNA epigenetically silences transcription by targeting g9a to chromatin. *Science* **2008**, *322*, 1717–1720.
48. Leeb, M.; Steffen, P.A.; Wutz, A. X chromosome inactivation sparked by non-coding RNAs. *RNA Biol.* **2009**, *6*, 94–99.
49. Pollard, K.S.; Salama, S.R.; Lambert, N.; Lambot, M.A.; Coppens, S.; Pedersen, J.S.; Katzman, S.; King, B.; Onodera, C.; Siepel, A.; *et al*. An RNA gene expressed during cortical development evolved rapidly in humans. *Nature* **2006**, *443*, 167–172.
50. Leontis, N.B.; Westhof, E. Analysis of RNA motifs. *Curr. Opin. Struct. Biol.* **2003**, *13*, 300–308.
51. Bejerano, G.; Pheasant, M.; Makunin, I.; Stephen, S.; Kent, W.J.; Mattick, J.S.; Haussler, D. Ultraconserved elements in the human genome. *Science* **2004**, *304*, 1321–1325.
52. Chen, C.T.; Wang, J.C.; Cohen, B.A. The strength of selection on ultraconserved elements in the human genome. *Am. J. Hum. Genet.* **2007**, *80*, 692–704.
53. Katzman, S.; Kern, A.D.; Bejerano, G.; Fewell, G.; Fulton, L.; Wilson, R.K.; Salama, S.R.; Haussler, D. Human genome ultraconserved elements are ultraselected. *Science* **2007**, *317*, 915.
54. Catucci, I.; Verderio, P.; Pizzamiglio, S.; Manoukian, S.; Peissel, B.; Barile, M.; Tizzoni, L.; Bernard, L.; Ravagnani, F.; Galastri, L.; *et al*. Snps in ultraconserved elements and familial breast cancer risk. *Carcinogenesis* **2009**, *30*, 544–545.
55. Scaruffi, P.; Stigliani, S.; Moretti, S.; Coco, S.; de Vecchi, C.; Valdora, F.; Garaventa, A.; Bonassi, S.; Tonini, G.P. Transcribed-ultra conserved region expression is associated with outcome in high-risk neuroblastoma. *BMC Cancer* **2009**, *9*, 441.
56. Orom, U.A.; Derrien, T.; Beringer, M.; Gumireddy, K.; Gardini, A.; Bussotti, G.; Lai, F.; Zytnicki, M.; Notredame, C.; Huang, Q.; *et al*. Long noncoding RNAs with enhancer-like function in human cells. *Cell* **2010**, *143*, 46–58.

57. Willingham, A.T.; Orth, A.P.; Batalov, S.; Peters, E.C.; Wen, B.G.; Aza-Blanc, P.; Hogenesch, J.B.; Schultz, P.G. A strategy for probing the function of noncoding RNAs finds a repressor of nfat. *Science* **2005**, *309*, 1570–1573.
58. Silva, J.M.; Perez, D.S.; Pritchett, J.R.; Halling, M.L.; Tang, H.; Smith, D.I. Identification of long stress-induced non-coding transcripts that have altered expression in cancer. *Genomics* **2010**, *95*, 355–362.
59. Hunter, D.J. Gene-environment interactions in human diseases. *Nat. Rev.* **2005**, *6*, 287–298.
60. Liu, L.; Li, Y.; Tollefsbol, T.O. Gene-environment interactions and epigenetic basis of human diseases. *Curr. Issues Mol. Biol.* **2008**, *10*, 25–36.
61. Thrash-Bingham, C.A.; Tartof, K.D. Ahif: A natural antisense transcript overexpressed in human renal cancer and during hypoxia. *J. Natl. Cancer Inst.* **1999**, *91*, 143–151.
62. Cayre, A.; Rossignol, F.; Clottes, E.; Penault-Llorca, F. Ahif but not hif-1alpha transcript is a poor prognostic marker in human breast cancer. *Breast Cancer Res.* **2003**, *5*, R223–R230.
63. He, H.; Nagy, R.; Liyanarachchi, S.; Jiao, H.; Li, W.; Suster, S.; Kere, J.; de la Chapelle, A. A susceptibility locus for papillary thyroid carcinoma on chromosome 8q24. *Cancer Res.* **2009**, *69*, 625–631.
64. Broadbent, H.M.; Peden, J.F.; Lorkowski, S.; Goel, A.; Ongen, H.; Green, F.; Clarke, R.; Collins, R.; Franzosi, M.G.; Tognoni, G.; *et al*. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked snps in the anril locus on chromosome 9p. *Hum. Mol. Genet.* **2008**, *17*, 806–814.
65. Yasuno, K.; Bilguvar, K.; Bijlenga, P.; Low, S.K.; Krischek, B.; Auburger, G.; Simon, M.; Krex, D.; Arlier, Z.; Nayak, N.; *et al*. Genome-wide association study of intracranial aneurysm identifies three new risk loci. *Nat. Genet.* **2010**, *42*, 420–425.
66. Shete, S.; Hosking, F.J.; Robertson, L.B.; Dobbins, S.E.; Sanson, M.; Malmer, B.; Simon, M.; Marie, Y.; Boisselier, B.; Delattre, J.Y.; *et al*. Genome-wide association study identifies five susceptibility loci for glioma. *Nat. Genet.* **2009**, *41*, 899–904.
67. Stacey, S.N.; Sulem, P.; Masson, G.; Gudjonsson, S.A.; Thorleifsson, G.; Jakobsdottir, M.; Sigurdsson, A.; Gudbjartsson, D.F.; Sigurgeirsson, B.; Benediktsdottir, K.R.; *et al*. New common variants affecting susceptibility to basal cell carcinoma. *Nat. Genet.* **2009**, *41*, 909–914.
68. Turnbull, C.; Ahmed, S.; Morrison, J.; Pernet, D.; Renwick, A.; Maranian, M.; Seal, S.; Ghoussaini, M.; Hines, S.; Healey, C.S.; *et al*. Genome-wide association study identifies five new breast cancer susceptibility loci. *Nat. Genet.* **2010**, *42*, 504–507.
69. Ladd, P.D.; Smith, L.E.; Rabaia, N.A.; Moore, J.M.; Georges, S.A.; Hansen, R.S.; Hagerman, R.J.; Tassone, F.; Tapscott, S.J.; Filippova, G.N. An antisense transcript spanning the cgg repeat region of fmr1 is upregulated in premutation carriers but silenced in full mutation individuals. *Hum. Mol. Genet.* **2007**, *16*, 3174–3187.
70. Moseley, M.L.; Zu, T.; Ikeda, Y.; Gao, W.; Mosemiller, A.K.; Daughters, R.S.; Chen, G.; Weatherspoon, M.R.; Clark, H.B.; Ebner, T.J.; *et al*. Bidirectional expression of cug and cag expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nat. Genet.* **2006**, *38*, 758–769.

71. Faghihi, M.A.; Modarresi, F.; Khalil, A.M.; Wood, D.E.; Sahagan, B.G.; Morgan, T.E.; Finch, C.E.; St Laurent, G., III; Kenny, P.J.; Wahlestedt, C. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat. Med.* **2008**, *14*, 723–730.
72. Chen, W.; Bocker, W.; Brosius, J.; Tiedge, H. Expression of neural bc200 RNA in human tumours. *J. Pathol.* **1997**, *183*, 345–351.
73. Iacoangeli, A.; Lin, Y.; Morley, E.J.; Muslimov, I.A.; Bianchi, R.; Reilly, J.; Weedon, J.; Diallo, R.; Bocker, W.; Tiedge, H. Bc200 RNA in invasive and preinvasive breast cancer. *Carcinogenesis* **2004**, *25*, 2125–2133.
74. Mus, E.; Hof, P.R.; Tiedge, H. Dendritic bc200 RNA in aging and in Alzheimer's disease. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10679–10684.
75. Eis, P.S.; Tam, W.; Sun, L.; Chadburn, A.; Li, Z.; Gomez, M.F.; Lund, E.; Dahlberg, J.E. Accumulation of mir-155 and bic RNA in human b cell lymphomas. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 3627–3632.
76. Tsang, W.P.; Wong, T.W.; Cheung, A.H.; Co, C.N.; Kwok, T.T. Induction of drug resistance and transformation in human cancer cells by the noncoding RNA cudr. *RNA* **2007**, *13*, 890–898.
77. Bussemakers, M.J.; van Bokhoven, A.; Verhaegh, G.W.; Smit, F.P.; Karthaus, H.F.; Schalken, J.A.; Debruyne, F.M.; Ru, N.; Isaacs, W.B. Dd3: A new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer Res.* **1999**, *59*, 5975–5979.
78. De Kok, J.B.; Verhaegh, G.W.; Roelofs, R.W.; Hessels, D.; Kiemeney, L.A.; Aalders, T.W.; Swinkels, D.W.; Schalken, J.A. Dd3(pca3), a very sensitive and specific marker to detect prostate tumors. *Cancer Res.* **2002**, *62*, 2695–2698.
79. Khalil, A.M.; Faghihi, M.A.; Modarresi, F.; Brothers, S.P.; Wahlestedt, C. A novel RNA transcript with antiapoptotic function is silenced in fragile x syndrome. *PLoS One* **2008**, *3*, e1486.
80. Mourtada-Maarabouni, M.; Pickard, M.R.; Hedge, V.L.; Farzaneh, F.; Williams, G.T. Gas5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene* **2009**, *28*, 195–208.
81. Brannan, C.I.; Dees, E.C.; Ingram, R.S.; Tilghman, S.M. The product of the h19 gene may function as an RNA. *Mol. Cell. Biol.* **1990**, *10*, 28–36.
82. Gabory, A.; Jammes, H.; Dandolo, L. The h19 locus: Role of an imprinted non-coding RNA in growth and development. *BioEssays* **2010**, *32*, 473–480.
83. Matouk, I.J.; DeGroot, N.; Mezan, S.; Ayesh, S.; Abu-lail, R.; Hochberg, A.; Galun, E. The h19 non-coding RNA is essential for human tumor growth. *PLoS One* **2007**, *2*, e845.
84. Yang, F.; Bi, J.; Xue, X.; Zheng, L.; Zhi, K.; Hua, J.; Fang, G. Up-regulated long non-coding RNA h19 contributes to proliferation of gastric cancer cells. *FEBS J.* **2012**, *279*, 3159–3165.
85. Luo, M.; Li, Z.; Wang, W.; Zeng, Y.; Liu, Z.; Qiu, J. Long non-coding RNA h19 increases bladder cancer metastasis by associating with ezh2 and inhibiting e-cadherin expression. *Cancer Lett.* **2013**, *333*, 213–221.
86. Rinn, J.L.; Kertesz, M.; Wang, J.K.; Squazzo, S.L.; Xu, X.; Brugmann, S.A.; Goodnough, L.H.; Helms, J.A.; Farnham, P.J.; Segal, E.; *et al*. Functional demarcation of active and silent chromatin domains in human hox loci by noncoding RNAs. *Cell* **2007**, *129*, 1311–1323.

87. Panzitt, K.; Tschernatsch, M.M.; Guelly, C.; Moustafa, T.; Stradner, M.; Strohmaier, H.M.; Buck, C.R.; Denk, H.; Schroeder, R.; Trauner, M.; *et al*. Characterization of hulc, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. *Gastroenterology* **2007**, *132*, 330–342.
88. Matouk, I.J.; Abbasi, I.; Hochberg, A.; Galun, E.; Dweik, H.; Akkawi, M. Highly upregulated in liver cancer noncoding RNA is overexpressed in hepatic colorectal metastasis. *Eur. J. Gastroenterol. Hepatol.* **2009**, *21*, 688–692.
89. Tanaka, K.; Shiota, G.; Meguro, M.; Mitsuya, K.; Oshimura, M.; Kawasaki, H. Loss of imprinting of long qt intronic transcript 1 in colorectal cancer. *Oncology* **2001**, *60*, 268–273.
90. Pasic, I.; Shlien, A.; Durbin, A.D.; Stavropoulos, D.J.; Baskin, B.; Ray, P.N.; Novokmet, A.; Malkin, D. Recurrent focal copy-number changes and loss of heterozygosity implicate two noncoding RNAs and one tumor suppressor gene at chromosome 3q13.31 in osteosarcoma. *Cancer Res.* **2010**, *70*, 160–171.
91. Ying, L.; Chen, Q.; Wang, Y.; Zhou, Z.; Huang, Y.; Qiu, F. Upregulated malat-1 contributes to bladder cancer cell migration by inducing epithelial-to-mesenchymal transition. *Mol. Biosyst.* **2012**, *8*, 2289–2294.
92. Tano, K.; Mizuno, R.; Okada, T.; Rakwal, R.; Shibato, J.; Masuo, Y.; Ijiri, K.; Akimitsu, N. Malat-1 enhances cell motility of lung adenocarcinoma cells by influencing the expression of motility-related genes. *FEBS Lett.* **2010**, *584*, 4575–4580.
93. Ji, P.; Diederichs, S.; Wang, W.; Boing, S.; Metzger, R.; Schneider, P.M.; Tidow, N.; Brandt, B.; Buerger, H.; Bulk, E.; *et al*. Malat-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* **2003**, *22*, 8031–8041.
94. Miyoshi, N.; Wagatsuma, H.; Wakana, S.; Shiroishi, T.; Nomura, M.; Aisaka, K.; Kohda, T.; Surani, M.A.; Kaneko-Ishino, T.; Ishino, F. Identification of an imprinted gene, meg3/gtl2 and its human homologue meg3, first mapped on mouse distal chromosome 12 and human chromosome 14q. *Genes Cells* **2000**, *5*, 211–220.
95. Zhang, X.; Zhou, Y.; Mehta, K.R.; Danila, D.C.; Scolavino, S.; Johnson, S.R.; Klibanski, A. A pituitary-derived meg3 isoform functions as a growth suppressor in tumor cells. *J. Clin. Endocrinol. Metab.* **2003**, *88*, 5119–5126.
96. Braconi, C.; Kogure, T.; Valeri, N.; Huang, N.; Nuovo, G.; Costinean, S.; Negrini, M.; Miotto, E.; Croce, C.M.; Patel, T. MicroRNA-29 can regulate expression of the long non-coding RNA gene meg3 in hepatocellular cancer. *Oncogene* **2011**, *30*, 4750–4756.
97. Ying, L.; Huang, Y.; Chen, H.; Wang, Y.; Xia, L.; Chen, Y.; Liu, Y.; Qiu, F. Downregulated meg3 activates autophagy and increases cell proliferation in bladder cancer. *Mol. Biosyst.* **2013**, *9*, 407–411.
98. Ishii, N.; Ozaki, K.; Sato, H.; Mizuno, H.; Saito, S.; Takahashi, A.; Miyamoto, Y.; Ikegawa, S.; Kamatani, N.; Hori, M.; *et al*. Identification of a novel non-coding RNA, miat, that confers risk of myocardial infarction. *J. Hum. Genet.* **2006**, *51*, 1087–1099.
99. Yu, M.; Ohira, M.; Li, Y.; Niizuma, H.; Oo, M.L.; Zhu, Y.; Ozaki, T.; Isogai, E.; Nakamura, Y.; Koda, T.; *et al*. High expression of ncran, a novel non-coding RNA mapped to chromosome 17q25.1, is associated with poor prognosis in neuroblastoma. *Int. J. Oncol.* **2009**, *34*, 931–938.

100. Zhu, Y.; Yu, M.; Li, Z.; Kong, C.; Bi, J.; Li, J.; Gao, Z.; Li, Z. Ncran, a newly identified long noncoding RNA, enhances human bladder tumor growth, invasion, and survival. *Urology* **2011**, *77*, e511–e515.
101. Castelnuovo, M.; Massone, S.; Tasso, R.; Fiorino, G.; Gatti, M.; Robello, M.; Gatta, E.; Berger, A.; Strub, K.; Florio, T.; *et al*. An alu-like RNA promotes cell differentiation and reduces malignancy of human neuroblastoma cells. *FASEB J.* **2010**, *24*, 4033–4046.
102. Petrovics, G.; Zhang, W.; Makarem, M.; Street, J.P.; Connelly, R.; Sun, L.; Sesterhenn, I.A.; Srikantan, V.; Moul, J.W.; Srivastava, S., Elevated expression of pcgem1, a prostate-specific gene with cell growth-promoting function, is associated with high-risk prostate cancer patients. Oncogene 2004, 23, 605-611.
103. Sonkoly, E.; Bata-Csorgo, Z.; Pivarcsi, A.; Polyanka, H.; Kenderessy-Szabo, A.; Molnar, G.; Szentpali, K.; Bari, L.; Megyeri, K.; Mandi, Y.; *et al*. Identification and characterization of a novel, psoriasis susceptibility-related noncoding RNA gene, prins. *J. Biol. Chem.* **2005**, *280*, 24159–24167.
104. Chung, S.; Nakagawa, H.; Uemura, M.; Piao, L.; Ashikawa, K.; Hosono, N.; Takata, R.; Akamatsu, S.; Kawaguchi, T.; Morizono, T.; *et al*. Association of a novel long non-coding RNA in 8q24 with prostate cancer susceptibility. *Cancer Sci.* **2011**, *102*, 245–252.
105. Maida, Y.; Yasukawa, M.; Furuuchi, M.; Lassmann, T.; Possemato, R.; Okamoto, N.; Kasim, V.; Hayashizaki, Y.; Hahn, W.C.; Masutomi, K. An RNA-dependent RNA polymerase formed by tert and the rmrp RNA. *Nature* **2009**, *461*, 230–235.
106. Shirasawa, S.; Harada, H.; Furugaki, K.; Akamizu, T.; Ishikawa, N.; Ito, K.; Ito, K.; Tamai, H.; Kuma, K.; Kubota, S.; *et al*. Snps in the promoter of a b cell-specific antisense transcript, sas-zfat, determine susceptibility to autoimmune thyroid disease. *Hum. Mol. Genet.* **2004**, *13*, 2221–2231.
107. Khaitan, D.; Dinger, M.E.; Mazar, J.; Crawford, J.; Smith, M.A.; Mattick, J.S.; Perera, R.J. The melanoma-upregulated long noncoding RNA spry4-it1 modulates apoptosis and invasion. *Cancer Res.* **2011**, *71*, 3852–3862.
108. Leygue, E. Steroid receptor RNA activator (sra1): Unusual bifaceted gene products with suspected relevance to breast cancer. *Nucl. Recept. Signaling* **2007**, *5*, e006.
109. Chooniedass-Kothari, S.; Hamedani, M.K.; Troup, S.; Hube, F.; Leygue, E. The steroid receptor RNA activator protein is expressed in breast tumor tissues. *Int. J. Cancer* **2006**, *118*, 1054–1059.
110. Cao, Y.; Bryan, T.M.; Reddel, R.R. Increased copy number of the tert and terc telomerase subunit genes in cancer cells. *Cancer Sci.* **2008**, *99*, 1092–1099.
111. Mishra, A.; Godavarthi, S.K.; Jana, N.R. Ube3a/e6-ap regulates cell proliferation by promoting proteasomal degradation of p27. *Neurobiol. Dis.* **2009**, *36*, 26–34.
112. Numata, K.; Kohama, C.; Abe, K.; Kiyosawa, H. Highly parallel snp genotyping reveals high-resolution landscape of mono-allelic ube3a expression associated with locus-wide antisense transcription. *Nucleic Acids Res.* **2011**, *39*, 2649–2657.
113. Wang, F.; Li, X.; Xie, X.; Zhao, L.; Chen, W. Uca1, a non-protein-coding RNA up-regulated in bladder carcinoma and embryo, influencing cell growth and promoting invasion. *FEBS Lett.* **2008**, *582*, 1919–1927.

114. Wang, X.S.; Zhang, Z.; Wang, H.C.; Cai, J.L.; Xu, Q.W.; Li, M.Q.; Chen, Y.C.; Qian, X.P.; Lu, T.J.; Yu, L.Z.; *et al*. Rapid identification of uca1 as a very sensitive and specific unique marker for human bladder carcinoma. *Clin. Cancer Res.* **2006**, *12*, 4851–4858.
115. Yang, C.; Li, X.; Wang, Y.; Zhao, L.; Chen, W. Long non-coding RNA uca1 regulated cell cycle distribution via creb through pi3-k dependent pathway in bladder carcinoma cells. *Gene* **2012**, *496*, 8–16.
116. Askarian-Amiri, M.E.; Crawford, J.; French, J.D.; Smart, C.E.; Smith, M.A.; Clark, M.B.; Ru, K.; Mercer, T.R.; Thompson, E.R.; Lakhani, S.R.; *et al*. Snord-host RNA zfas1 is a regulator of mammary development and a potential marker for breast cancer. *RNA* **2011**, *17*, 878–891.
117. Visscher, P.M.; Brown, M.A.; McCarthy, M.I.; Yang, J. Five years of gwas discovery. *Am. J. Hum. Genet.* **2012**, *90*, 7–24.
118. McPherson, R.; Pertsemlidis, A.; Kavaslar, N.; Stewart, A.; Roberts, R.; Cox, D.R.; Hinds, D.A.; Pennacchio, L.A.; Tybjaerg-Hansen, A.; Folsom, A.R.; *et al*. A common allele on chromosome 9 associated with coronary heart disease. *Science* **2007**, *316*, 1488–1491.
119. Samani, N.J.; Erdmann, J.; Hall, A.S.; Hengstenberg, C.; Mangino, M.; Mayer, B.; Dixon, R.J.; Meitinger, T.; Braund, P.; Wichmann, H.E.; *et al*. Genomewide association analysis of coronary artery disease. *N. Engl. J. Med.* **2007**, *357*, 443–453.
120. Jarinova, O.; Stewart, A.F.; Roberts, R.; Wells, G.; Lau, P.; Naing, T.; Buerki, C.; McLean, B.W.; Cook, R.C.; Parker, J.S.; *et al*. Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. *Arterioscler. Thromb. Vasc. Biol.* **2009**, *29*, 1671–1677.
121. Pasmant, E.; Sabbagh, A.; Vidaud, M.; Bieche, I. Anril, a long, noncoding RNA, is an unexpected major hotspot in gwas. *FASEB J.* **2011**, *25*, 444–448.
122. Visel, A.; Zhu, Y.; May, D.; Afzal, V.; Gong, E.; Attanasio, C.; Blow, M.J.; Cohen, J.C.; Rubin, E.M.; Pennacchio, L.A. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature* **2010**, *464*, 409–412.
123. Yap, K.L.; Li, S.; Munoz-Cabello, A.M.; Raguz, S.; Zeng, L.; Mujtaba, S.; Gil, J.; Walsh, M.J.; Zhou, M.M. Molecular interplay of the noncoding RNA anril and methylated histone h3 lysine 27 by polycomb cbx7 in transcriptional silencing of ink4a. *Mol. Cell* **2010**, *38*, 662–674.
124. El Messaoudi-Aubert, S.; Nicholls, J.; Maertens, G.N.; Brookes, S.; Bernstein, E.; Peters, G. Role for the mov10 RNA helicase in polycomb-mediated repression of the ink4a tumor suppressor. *Nat. Struct. Mol. Biol.* **2010**, *17*, 862–868.
125. He, S.; Su, H.; Liu, C.; Skogerbo, G.; He, H.; He, D.; Zhu, X.; Liu, T.; Zhao, Y.; Chen, R. MicroRNA-encoding long non-coding RNAs. *BMC Genom.* **2008**, *9*, 236.
126. Zhao, Y.; He, S.; Liu, C.; Ru, S.; Zhao, H.; Yang, Z.; Yang, P.; Yuan, X.; Sun, S.; Bu, D.; *et al*. MicroRNA regulation of messenger-like noncoding RNAs: A network of mutual microRNA control. *Trends Genet.* **2008**, *24*, 323–327.
127. Mercer, T.R.; Dinger, M.E.; Sunkin, S.M.; Mehler, M.F.; Mattick, J.S. Specific expression of long noncoding RNAs in the mouse brain. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 716–721.
128. Hagerman, P.J.; Hagerman, R.J. Fragile x-associated tremor/ataxia syndrome—An older face of the fragile x gene. *Nat. Clin. Pract. Neurol.* **2007**, *3*, 107–112.
129. Jacquemont, S.; Hagerman, R.J.; Hagerman, P.J.; Leehey, M.A. Fragile-x syndrome and fragile x-associated tremor/ataxia syndrome: Two faces of fmr1. *Lancet Neurol.* **2007**, *6*, 45–55.

130. Nemes, J.P.; Benzow, K.A.; Moseley, M.L.; Ranum, L.P.; Koob, M.D. The sca8 transcript is an antisense RNA to a brain-specific transcript encoding a novel actin-binding protein (klhl1). *Hum. Mol. Genet.* **2000**, *9*, 1543–1551.
131. Berchtold, N.C.; Cotman, C.W. Evolution in the conceptualization of dementia and Alzheimer's disease: Greco-roman period to the 1960s. *Neurobiol. Aging* **1998**, *19*, 173–189.
132. Fleck, D.; Garratt, A.N.; Haass, C.; Willem, M. Bace1 dependent neuregulin processing: Review. *Curr. Alzheimer Res.* **2012**, *9*, 178–183.
133. Lewejohann, L.; Skryabin, B.V.; Sachser, N.; Prehn, C.; Heiduschka, P.; Thanos, S.; Jordan, U.; Dell'Omo, G.; Vyssotski, A.L.; Pleskacheva, M.G.; *et al*. Role of a neuronal small non-messenger RNA: Behavioural alterations in bc1 RNA-deleted mice. *Behav. Brain Res.* **2004**, *154*, 273–289.
134. Qiu, M.T.; Hu, J.W.; Yin, R.; Xu, L. Long noncoding RNA: An emerging paradigm of cancer research. *Tumour Biol.* **2013**, *34*, 613–620.
135. Xu, C.; Yang, M.; Tian, J.; Wang, X.; Li, Z. Malat-1: A long non-coding RNA and its important 3' end functional motif in colorectal cancer metastasis. *Int. J. Oncol.* **2011**, *39*, 169–175.
136. Tripathi, V.; Ellis, J.D.; Shen, Z.; Song, D.Y.; Pan, Q.; Watt, A.T.; Freier, S.M.; Bennett, C.F.; Sharma, A.; Bubulya, P.A.; *et al*. The nuclear-retained noncoding RNA malat1 regulates alternative splicing by modulating sr splicing factor phosphorylation. *Mol. Cell* **2010**, *39*, 925–938.
137. Lu, L.; Zhu, G.; Zhang, C.; Deng, Q.; Katsaros, D.; Mayne, S.T.; Risch, H.A.; Mu, L.; Canuto, E.M.; Gregori, G.; *et al*. Association of large noncoding RNA hotair expression and its downstream intergenic cpg island methylation with survival in breast cancer. *Breast Cancer Res. Treat.* **2012**, *136*, 875–883.
138. Ishibashi, M.; Kogo, R.; Shibata, K.; Sawada, G.; Takahashi, Y.; Kurashige, J.; Akiyoshi, S.; Sasaki, S.; Iwaya, T.; Sudo, T.; *et al*. Clinical significance of the expression of long non-coding RNA hotair in primary hepatocellular carcinoma. *Oncol. Rep.* **2013**, *29*, 946–950.
139. Kogo, R.; Shimamura, T.; Mimori, K.; Kawahara, K.; Imoto, S.; Sudo, T.; Tanaka, F.; Shibata, K.; Suzuki, A.; Komune, S.; *et al*. Long noncoding RNA hotair regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res.* **2011**, *71*, 6320–6326.
140. Tsai, M.C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J.K.; Lan, F.; Shi, Y.; Segal, E.; Chang, H.Y. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **2010**, *329*, 689–693.
141. Zhang, X.; Gejman, R.; Mahta, A.; Zhong, Y.; Rice, K.A.; Zhou, Y.; Cheunsuchon, P.; Louis, D.N.; Klibanski, A. Maternally expressed gene 3, an imprinted noncoding RNA gene, is associated with meningioma pathogenesis and progression. *Cancer Res.* **2010**, *70*, 2350–2358.
142. Cho, Y.; Gorina, S.; Jeffrey, P.D.; Pavletich, N.P. Crystal structure of a p53 tumor suppressor-DNA complex: Understanding tumorigenic mutations. *Science* **1994**, *265*, 346–355.
143. Zhou, Y.; Zhong, Y.; Wang, Y.; Zhang, X.; Batista, D.L.; Gejman, R.; Ansell, P.J.; Zhao, J.; Weng, C.; Klibanski, A. Activation of p53 by meg3 non-coding RNA. *J. Biol. Chem.* **2007**, *282*, 24731–24742.
144. Zhang, X.; Rice, K.; Wang, Y.; Chen, W.; Zhong, Y.; Nakayama, Y.; Zhou, Y.; Klibanski, A. Maternally expressed gene 3 (meg3) noncoding ribonucleic acid: Isoform structure, expression, and functions. *Endocrinology* **2010**, *151*, 939–947.

Reprinted from *IJMS*. Cite as: Hadjiargyrou, M.; Delihas, N. The Intertwining of Transposable Elements and Non-Coding RNAs. *Int. J. Mol. Sci.* **2013**, *14*, 13307-13328.

Review

The Intertwining of Transposable Elements and Non-Coding RNAs

Michael Hadjiargyrou [1] and Nicholas Delihas [2,*]

[1] Department of Life Sciences, Theobald Science Center, Room 420, New York Institute of Technology, Old Westbury, NY 11568, USA; E-Mail: mhadji@nyit.edu

[2] Department of Molecular Genetics and Microbiology, School of Medicine, Stony Brook University, Stony Brook, NY 11794, USA

***** Author to whom correspondence should be addressed; E-Mail: Nicholas.delihas@stonybrook.edu; Tel.: +1-631-286-9427; Fax: +1-631-632-9797.

Received: 20 May 2013; in revised form: 5 June 2013 / Accepted: 5 June 2013 / Published: 26 June 2013

Abstract: Growing evidence shows a close association of transposable elements (TE) with non-coding RNAs (ncRNA), and a significant number of small ncRNAs originate from TEs. Further, ncRNAs linked with TE sequences participate in a wide-range of regulatory functions. Alu elements in particular are critical players in gene regulation and molecular pathways. Alu sequences embedded in both long non-coding RNAs (lncRNA) and mRNAs form the basis of targeted mRNA decay via short imperfect base-pairing. Imperfect pairing is prominent in most ncRNA/target RNA interactions and found throughout all biological kingdoms. The piRNA-Piwi complex is multifunctional, but plays a major role in protection against invasion by transposons. This is an RNA-based genetic immune system similar to the one found in prokaryotes, the CRISPR system. Thousands of long intergenic non-coding RNAs (lincRNAs) are associated with endogenous retrovirus LTR transposable elements in human cells. These TEs can provide regulatory signals for lincRNA genes. A surprisingly large number of long circular ncRNAs have been discovered in human fibroblasts. These serve as "sponges" for miRNAs. Alu sequences, encoded in introns that flank exons are proposed to participate in RNA circularization via Alu/Alu base-pairing. Diseases are increasingly found to have a TE/ncRNA etiology. A single point mutation in a SINE/Alu sequence in a human long non-coding RNA leads to brainstem atrophy and death. On the other hand, genomic clusters of repeat sequences as well as lncRNAs function in epigenetic regulation. Some clusters are unstable, which can lead to formation

of diseases such as facioscapulohumeral muscular dystrophy. The future may hold more surprises regarding diseases associated with ncRNAs andTEs.

Keywords: non-coding RNAs; transposable elements; microRNAs; Alu sequences; endogenous retrovirus LTR; epigenetics; disease formation

1. Introduction

The genome is a dynamic entity, ever-changing as a result of endogenous DNA movement or the acquisition of exogenous DNA leading to genomic rearrangements. Such events have contributed to the plasticity and evolution of the genome and all of its complexity, much of which has slowly come to light over the past decades but whose pace has certainly accelerated in the last few years as a result of breakthroughs in genomic technologies, development of newer sequencing techniques, and availability of data in public databases by scientists all over the world. This new and vast genomic knowledge has led to a revolutionary and unparalleled in depth examination of the genome, transcriptome, proteome, interactome, *etc.* Indeed, the concept and definition of a gene may have to be altered [1,2].

However, one key question before us is what new functional loci and regulatory mechanisms have been formed during genomic evolution, especially as they pertain to the genesis of non-coding RNAs (ncRNAs), ncRNA regulatory roles and their association and interaction with transposable elements (TEs). In this review, we outline recent advances in origins of microRNAs (miRNA) and functional properties of ncRNAs as they pertain to their interaction with TEs, especially in humans. What emerges is a fascinating new picture of interconnected molecular interactions and regulatory pathways.

1.1. Non-coding RNAs

The development and use of new sequencing techniques, such as RNA-Seq has greatly increased our discovery of new RNAs [3,4]. The Encyclopedia of DNA Elements (ENCODE), an international project with the intent goal of determining functional elements of the entire human genome, has employed these techniques and found thousands of new RNA transcripts [2,5]. Surprisingly, about 75%–85% of the human genome is transcribed into primary and processed transcripts [2]; yet only 1.2% of the human genome encodes proteins. This suggests that most of the human genome space is devoted to RNA synthesis that is not devoted to protein-coding. Functions for most non-coding RNA (ncRNA) transcripts are unknown, but the future may hold fascinating prospects of finding new roles and molecular pathways. For example, thousands of circular ncRNAs (cirRNA) have recently been identified; they represent scrambled coding sequences that originate from exons (nonrandom products of RNA splicing) and are involved in small ncRNA regulation [6,7]. These cirRNAs are transcripts that do not encode proteins but have a regulatory role in the cell and thus are regulatory ncRNAs. A significant number of ncRNAs stem from non-protein-coding regions of the genome (intergenic regions), but many also originate from protein-coding regions as antisense transcripts, or from intron regions, and as just mentioned, from scrambled coding sequences. Many ncRNAs target mRNAs and induce their degradation. On the other hand, others are associated with regulation of transcription. Indeed the biological significance of regulation by RNA was grossly underestimated in the past.

Given the barrage of published studies on newly discovered ncRNAs, especially with eukaryotes, their classification and subclassification is indeed very challenging. However, Di Leva and Garofalo [8] used a simple classification system and presented three basic categories: (1) Housekeeping RNAs (rRNAs, tRNAs, snRNAs and snoRNAs); (2) short non-coding RNAs that are less than 200 nucleotides that include but are not limited to microRNAs (miRNAs), Piwi-interacting RNAs (piRNAs) and retrotransposon-derived ncRNAs, and (3) long non-coding RNAs (lncRNAs) that are greater than 200 nucleotides. lncRNAs are currently divided into long intergenic ncRNAs (lincRNAs) as they are encoded in intergenic regions, transcripts from introns, long ncRNA that are antisense transcripts in coding regions but do not encode proteins, and circular RNA transcripts from coding regions that have scrambled exon sequences and also do not encode proteins. Recent identification and classification of long ncRNAs lists additional categories [5].

A common theme that prevails in target RNA regulation by ncRNAs is the formation of short intermolecular RNA/RNA base-paired stems that contain Watson–Crick pairs, imperfect pairing (bulged and looped out positions), and non-canonical pairs. Imperfect ncRNA/target RNA pairing was determined experimentally with the first discovered and functionally characterized non-coding RNA in prokaryotes [9–11]. This is a type of ncRNA/target RNA binding that prevails throughout all biological kingdoms, although RNA binding proteins are also key factors in stable binding. Most of the ncRNAs discussed in this paper interact with their target RNAs via imperfect pairing.

1.2. Transposable Elements

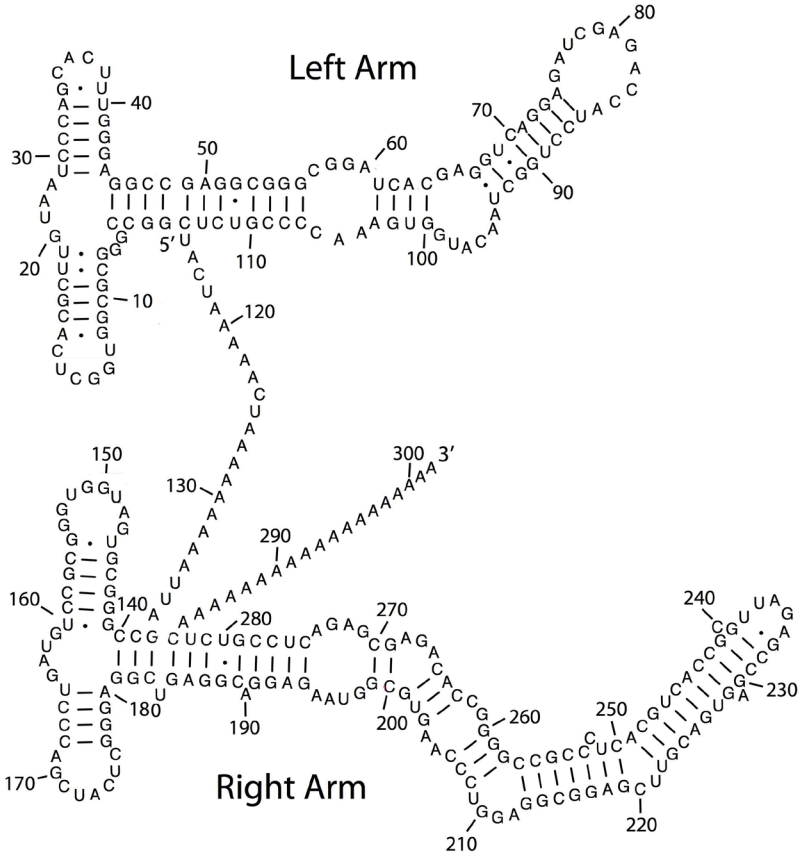
TEs are defined as mobile genetic elements (pieces of DNA capable of moving to new locations); they also constitute the "mobilome" in that they can impact cell transcription [12]. TEs are characterized as either Class I retrotransposons or Class II DNA transposons. Retrotransposons are further subdivided based on the presence of their long terminal repeat (LTR) that contains the element's functions for mobility and regulatory sequences. LTRs flank endogenous retroviruses (ERV) and are capable of transposition. ERVs are mostly inactive viruses due to accumulation of mutations, but LTRs are active transposons, encode for all the essential factors for mobility and can multiply within a cell independent of the ERV. They carry promoter and enhancer sequences enabling host genes to be transcribed, as well as lncRNA genomic sequences. There are about 500,000 copies of LTR sequences in Homo sapiens, which make up about 8 percent of the human genome.

The other type of Class I retrotransposons are composed of the Long Interspersed Nuclear Elements (LINEs) as well as the Short Interspersed Nuclear Elements (SINEs). LINEs are mobile, whereas SINEs are non-autonomous DNA transposable elements and require LINEs for their mobility and propagation [13]. Duplicate copies are generated during mobility, with sequences identical to the original element inserted in a new location on the genome. In time, such copies may accumulate mutations independently and therefore will differ in sequence from their original sequence leading to increased divergence.

SINEs do not encode proteins. Alu sequences are classified as SINE elements and are about 300 nucleotides. They originated from the 7SL RNA transcript via retrotranscription into DNA. 7SL RNA forms part of the signal recognition particle. A hallmark of Alu sequences is that Alu RNAs fold into specific stable stem-loop structures, albeit with extensive imperfect base-pairing (Figure 1). Alus

are highly abundant in mammalian cells, e.g., there ~10^6 copies of Alu in the human genome that make up ~11 percent of the genome [14], but most of these cannot be mobilized due to accumulation of mutations. Alu sequences are also found embedded in lncRNAs, where they are found to directly participate in base-pairing to target mRNAs (see Section 3.1). Additionally they are also found in piRNA genetic clusters (see Section 3.9).

Figure 1. Secondary structural model of Alu RNA. Modified from [15] with permission from Dr. Jennifer Doudna.



All-in-all, repeat sequences comprise 50%–75% of the human genome [14,16]. Repeats are broadly classified either as TE repeats or tandem repeats [17], but the major fraction of repeats represent transposable elements, either active or inactive. Tandem repeats represent a rather heterogeneous group, but some may have originated from TEs [18]. Repeats are regions where there is high recombination and this may sometimes result in genetic abnormalities.

In this review we focus on the origins of ncRNAs, the microRNAs from TEs and the interaction of ncRNA with TEs, primarily as found in mammalian tissues. Evidence is rapidly accumulating to show that this intimate association plays a central role in molecular and genetic mechanisms, such as RNA-based immunity. Furthermore, Cowley and Oakley have already described some of the impact of TEs in the promotion of human transcript diversity [12].

2. TE Origins of miRNAs

microRNAs (miRNA) are small non-coding regulatory RNAs molecules that function post-transcriptionally by binding to the 3'UTR of target mRNAs and ultimately inducing inhibition of

target mRNA function. While the majority of miRNAs originate from intergenic genomic sequences, some arise from genes and TEs. The molecular origins of many miRNAs support the hypothesis that miRNA hairpin generation is based on the insertion of two related TEs flanking a single genomic locus (see below). As such, transcription that occurs across this locus leads to the biogenesis of functional miRNAs. One of the earliest studies to indicate that a number of mammalian miRNAs are derived from TEs utilized a bioinformatics approach, where the authors analyzed the Sanger miRNAi database using a software program that specifically detects well characterized repeats [19]. Specifically, 11 different miRNA precursors contained repeat sequences (4 derived from LINE-2 repeats and others with SINEs, LTRs and simple repeats). The majority of these miRNAs are highly conserved across human, mouse and rat, but some are confined to only one or two species.

In a subsequent and more in depth study, Piriyapongsa and Jordan investigated the relationship between human miRNAs and TEs by comparing the genomic locations of experimentally characterized human miRNA genes with the locations of annotated genomic TE sequences [20]. A correlation was observed, and nt sequence comparisons showed a high identity between seven members of the family of miRNAs hsa-mir-548 and the miniature inverted repeat transposable element (MITE), Made1. By use of human genome tilling arrays that visualize genomic expression, one Made1 element was found to be inserted into a transcriptionally active intergenic site. Made1 and other MITEs have palindromic sequences, and when transcribed, show a segment that has an imperfect stem-loop RNA structure. As RNAi-related enzymes can recognize this type of imperfect stem-loop and process it into the 22 bp mature miRNA sequences, the authors proposed that Made1 TE transcripts are processed into hsa-mir-548 miRNAs. The expression date and high sequence identities strongly support the proposed TE origin of several hsa-mir-548 family members.

In a related study, Piriyapongsa *et al*. used comparative genomic sequence data from the UCSC Genome Browser and evaluated the evolution of TE-derived human miRNAs [21]. They found 55 experimentally characterized human miRNA genes that were derived from TEs (LINE and SINE, LTRs and DNA transposons). Sequence comparisons showed that on average, TE-derived miRNAs are less conserved than non-TE-derived miRNAs. Further, a subset of these, are related to the ancient L2 and MIR families. Results also predicted an additional 85 novel TE-derived miRNA genes. Lastly, for some of the TE-derived miRNAs and their putative target genes, a comparison of expression patterns (miRNA *vs.* mRNA) was performed and revealed a number of them to have anti-correlated expression, consistent with regulation via mRNA degradation and thus supporting their regulatory function.

Examination of fourteen previously identified marsupial (Monodelphis domestica) specific miRNAs and their flanking sequences revealed that half of these miRNAs evolved from marsupial-specific TEs [22]. More specifically, six of these TE sequences were identified as LINEs and one as a Mariner DNA transposon. In a subsequent study, Yuan and colleagues also investigated another placental-specific miRNA gene family (miR-1302) that at the time of the analysis had 11 members that were distributed in the human genome (present in the miRBase) [23]. They demonstrated that all members of this family were derived from a single transposon (MER53 element). MER53 is a TE with a 193-bp consensus sequence and is characterized by the presence of terminal inverted repeats and TA target site duplications that can form palindromic structures [24]. Further analysis of the phylogenetic distribution and evolution dynamics of the miR-1302 family identified 36 potential paralogs of

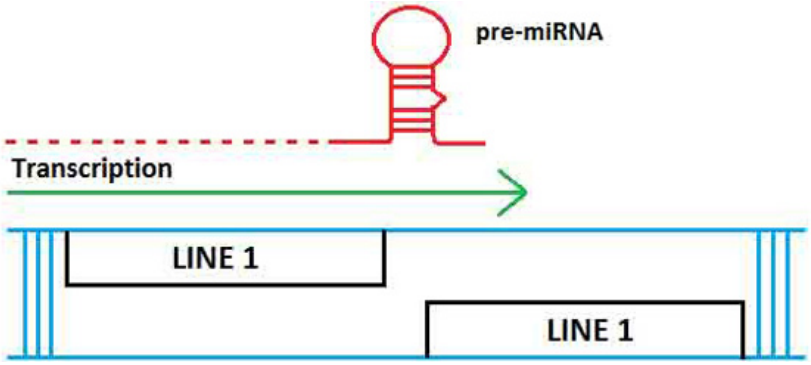
MER53-derived miR-1302 genes in the human genome and another 58 potential orthologs in placental mammals and showed that these members of the hsa-mir-1302 family emerged within the last 180 million years since placental mammals diverged from marsupials. Lastly, the authors also explored the targets of the mature human miR-1302 and found 1835 genes with predicted function in transportation, localization, system development processes and their regulation, as well as in binding and in transcription regulation [23].

Genome-wide studies were performed using a comparative genomics approach in order to identify human miRNA paralogs (in mouse and rhesus) in segmental duplication pair data [25]. Of ~1000 miRNA genes and ~1000 mature sequences from human, ~700 miRNA genes and ~1000 mature sequences from mouse, and ~500 miRNA genes and ~500 mature sequences from rhesus, they identified 228 novel miRNA homologs in the rhesus genome and 22 novel miRNA homologs in the mouse genome (by using miRBase 16). Further, they also found 12 and 2 novel miRNA paralogs in the human and mouse genome, respectively, but none were found in the rhesus genome. In a separate analysis, the authors also examined the coverage density of repetitive elements, and if it was at least 50% in a miRNA gene or 100% in one of the associated mature miRNA sequences, then the miRNA gene was considered to be a RdmiR. Using this rule, the study identified a large number of miRNAs genes that overlap with repeats (TEs: LINEs, SINEs and LTRs) and other types of repetitive elements (DNA transposons, specifically MADE1 elements) within the three genomes; 226 (human), 115 (rhesus) and 141 (mouse). The study also identified a smaller number of possible repeat derived miRNAs, which they termed RrmiRs. Lastly, a computational analysis was conducted to investigate the functions of 19 of the conserved RrmiR families (between the three genomes), by identifying their target genes and it was found that the most significant targets are involved in transcriptional regulation, central nervous system development, and negative regulation of biological process. Collectively, the results of this study suggest that repetitive elements contribute to the de novo origin of miRNAs, and that large segmentation duplication events most likely accelerate the expansion of miRNA families (including RdmiRs).

A more recent study involved a comprehensive analysis of the genomic events responsible for the formation of ~15,000 annotated miRNAs against the principle datasets for TEs and ncRNAs and found 2392 (~15%) TE-based miRNAs [26]. The majority of these TE-based miRNAs may have originated via the proposed mechanism depicted in Figure 2.

The authors further investigated the exact TE origins of these 2392 miRNAs and showed that DNA transposons comprise the TE most frequently responsible for miRNA generation (891); others were: LTR Retrotransposon (414), Non-LTR Retrotransposon (814), LINE (312), SINE (353), Satellite (137) and others (136). This last category ("Other") had significant sequence identity to known noncoding RNA sequences (e.g., snoRNAs, scaRNAs, tRNAs). Lastly, a hypothetical scheme proposes that the regulatory miRNAs may have arisen via selective subfunctionalization created by the associated benefit of regulating host genes containing portions of TEs.

Figure 2. Schematic of proposed origin of TE-based miRNAs. When two related but not identical LINE1elements insert themselves near each other, but on opposite strands of the DNA (in blue), they can create a precursor miRNA containing an imperfect stem-loop upon transcription (red line). The pre-miRNA is shown above the arrow and transcription is indicated from the positive strand LINE1. The stem is potentially recognized and processed by the endogenous RNAi machinery. The pre-miRNA stem-loop depicted is representational. Modified from [26].



Using the miRBASE database, a more recent study sought to map all miRNA precursors to several genomes and to determine the repetition and dispersion of the corresponding loci, as well as the repetitive elements overlapping these loci. To facilitate this analysis an automatic method called ncRNAclassifier was used in order to classify the relationship of TEs with pre-ncRNAs [27]. By applying this method, a correlation between the number of pre-ncRNA candidates and the presence of TEs was determined using six genomes (frog, human, mouse, nematode, rat and sea squirt). The results indicate that 235 and 68 mis-annotated pre-miRNAs correspond completely to TEs out of 1426 human and 721 mouse pre-miRNAs of miRBase (10.0 release), respectively. Further, the various types of TEs involved were also identified and include (MADE1 and other MITEs, DNA transposons, LTR/ERV, CR1/RTE, L1, SINE, other non-LTR). Lastly, the authors suggest that the ncRNAclassifier can be openly used to determine if a given ncRNA hairpin sequence corresponded to a TE sequence.

An investigation of the TE origins of miRNAs focused on the MER (MEdium Reiteration frequency), interspersed repeats in the genomes of primates, rodentia, and lagomorpha) transposon-derived miRNAs in human genome. Once again, a bioinformatics approach was undertaken to identify the specific miRNAs that are derived from palindromic MERs, by analyzing MER paralogs in human genome. Results from this study identified three miRNAs derived from MER96 located on chromosome 3, and MER91C paralogs located on chromosome 8 and chromosome 17 [28]. More importantly, this study also experimentally validated the interactions between these MER-derived miRNAs with AGO1, AGO2, and AGO3 proteins (involved in gene silencing and act as the catalytic component of the RNA induced silencing complex [RISC]).

Lastly, there are additional classes of small ncRNAs that originated from TEs and/or consist of TE sequences, e.g., certain piRNAs and the specialized SINE and Alu transcripts that function as small ncRNAs; these are discussed below with respect to their functional roles.

In summary, a sizable proportion of miRNAs appear to be derived from TEs. It is highly probably that future bioinformatic analyses will increase the number of miRNA-transposable element

relationship as not all miRNAs have been discovered and most likely, all consensus repetitive elements have not yet been described. As ncRNAs serve such a critical regulatory role, TE colonization of the genome has given rise to a number of regulatory processes, several of which we discuss here.

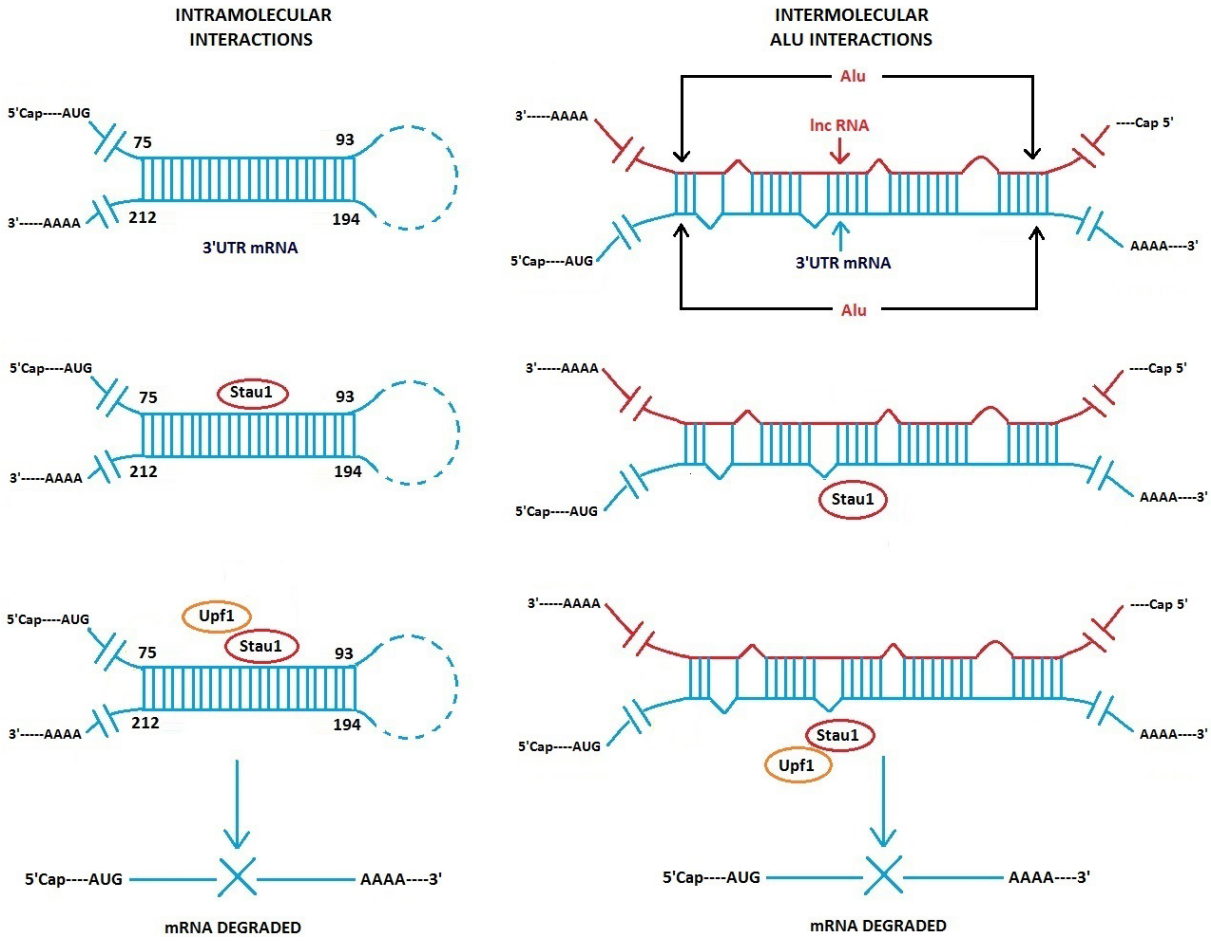
3. Interaction of TEs with ncRNAs—Functional and Disease-Related Significance

3.1. Alu Element Embedded in Long ncRNAs and mRNAs—Crucial Role in Target mRNA Decay

Gong and Maguat revealed the importance of Alu intermolecular base-pairing to lncRNA-induced degradation of mRNA [29,30]. By computational analysis, Alu sequences are found present in ~380 lncRNAs in HeLa cells [30]. In addition, mRNAs were identified that contain an Alu sequence in their 3'UTR regions. Certain mRNAs are targets of the double-stranded RNA binding protein Staufen1 (Stau1), which can induce degradation of the mRNAs. Co-immunoprecipitation experiments show an Alu-containing lncRNA, which originates from chromosome 11, binds to and decreases the abundance of target messages, *i.e.*, plasminogen activator inhibitor type 1 (SERPINE 1) mRNA and an mRNA that encodes an unknown protein termed FLJ21870. Both these mRNAs have an Alu sequence in their 3'UTRs. By secondary structure modeling, it was shown that Alu sequences in lncRNA can base-pair to Alu sequences in the 3'UTR of target mRNAs by intermolecular base-pairing with a stable $-\Delta G$ (Figure 3). This pairing is imperfect and contains bulged positions. The intermolecular stem structure, formed by the interaction between the Alu sequences present in the lncRNA and in the 3' UTR of the mRNA serves as the binding site for Stau1, which subsequently recruits UPF1, a protein required to initiate mRNA decay. Several hundred other lncRNAs contain Alu sequences and have the potential to base-pair with Alu-containing mRNAs, but possible functions of the majority of the several hundred Alu-containing lncRNAs are unknown.

Alu/Alu sequence pairing is not the only interaction seen with Stau1-binding mRNAs, but it represents an important “variation on a theme”. In previous experiments with another Stau1-binding mRNA that contains no Alu sequences but a perfect 19 bp stem in the 3'UTR, it was shown that Stau1 binds the perfect 19 bp stem that is formed intramolecularly between distal sequences in the 3'UTR of the mRNA [31] (Figure 3, left). ARF1 is an ADP-ribosylation factor 1 protein, and the *arf1* mRNA transcript has a Stau1-binding site. The 19 bp stem is phylogenetically conserved in different mammalian species. The question remains, how many Stau1-binding mRNAs are targets for decay via intramolecular bp within a message and how many by intermolecular Alu/Alu pairing. As mentioned, there are several hundred lncRNAs that have Alu sequences. In addition, in a related topic, it should also be pointed out that inverted Alu elements are found in many human mRNA 3'UTR sequences. These can form double-stranded intramolecular stems, and they appear to affect mRNA translation efficiency [32].

Figure 3. Schematic of binding of Stau1 to 3' UTR mRNA intramolecular stem (*arf1* mRNA) (**left**) and to intermolecular stem formed by Alu sequence in lncRNA with Alu sequence within 3'UTR mRNA. Upf1is an RNA helicase. The two RNA duplex stems shown are not drawn to scale. Modified from Gong and Maquat [30].



Thus, mRNAs that have a Stau1 binding site, whether it is a formed intermolecularly via Alu/Alu lncRNA/mRNA imperfect base-pairing or by perfect Watson–Crick intramolecular pairing within the mRNA can be targeted for degradation. This raises an interesting question concerning the specificity of Stau1 and its recognition sites on duplex RNA stems and imperfect *vs.* perfect double stranded stems. The probability of mistakes in recognition must be very low, yet two or more types of RNA tertiary structures are recognized with great accuracy. Crystal structures of perfect and imperfect double-stranded RNA/Stau1-protein complexes would be of major interest. Gleghorn *et al.* have already determined crystal structures of Stau1 and showed that dimerization of Stau1 occurs by a degenerate dsRNA-binding domain on Stau1 [33]. And in another recent study, it was revealed that dimerization can also involve Stau2 [34].

In another study from this laboratory, rodents appear to use the same mechanism of mRNA regulation involving intermolecular imperfect base-pairing between lncRNAs and mRNAs, only this occurs via SINE elements B1, B2 or B4 found at 3'UTRs and in lncRNAs [35].

3.2. Long Non-Coding Antisense RNA Controls mRNA Translation—Importance of Embedded SINE/Alu Repeats

Carrieri *et al.* determined that a long ncRNA transcript, which is partially antisense to ubiquitin carboxy-terminal hydrolase L1 (Uchl1), is essential for increased translation of uch1 mRNA [36]. Uch1 is a neuron associated protein in mammals and may be involved in neurological disease formation [37,38]. Two segments of the antisense lncRNA are crucial for function: the 5' end of the lncRNA transcript that overlaps the uch1 sense transcript, and SINEB2 and Alu repeat segments located downstream on the antisense lncRNA. By using a bioinformatics approach, 31 antisense lncRNAs have been pinpointed that contain SINE/Alu sequences in their 3' end half regions. These RNAs can potentially base-pair to sense transcripts via their 5' ends to the 3'UTR of mRNAs, in a similar manner as uch1 sense RNA/antisense RNA transcript pairing occurs. However the mechanism by which the SINE sequences act to control translation of uch1 mRNA has not been determined, but a hint comes from data showing the orientation of the SINE in the antisense lncRNA is important in rescue experiments [36]. This implies a possible RNA/RNA base-pairing mechanism. This work may define a separate class of regulatory lncRNAs in mammals that appears to differ from the Alu-containing lncRNAs described by Gong and Maquat [29].

3.3. Point Mutation in LINE-1/Alu Element Embedded in a lncRNA Results in Lethal Brain Disease

A primate conserved LINE-1 sequence is found embedded in a lncRNA that maps to human chromosome 8p22 [39]. This LINE sequence also overlaps with an Alu sequence. The lncRNA constitutes a unique transcript originating from an intron. This RNA most likely has regulatory functions. A rare single point mutation, A to G in the LINE-1/Alu sequence is associated with brainstem cell atrophy, a genetic abnormality that results in lethal infantile encephalopathy in humans [39]. The LINE-1 is a degenerate retrotransposon and assumed not to be mobile. It was experimentally determined that in patient brain tissues, the expression of the mutant lncRNA was reduced nearly 10-fold relative to unmutated RNA control levels. mRNAs of two genes that map in the same locus as the lnRNA were found to be unchanged. In addition, knockdown experiments against wild-type lncRNA using siRNA showed a significant increase in apoptotic cells.

Several hypotheses have been presented to explain the drastic phenotypic effects of the single base-pair change in the LINE-1/Alu sequence [39]. One is that piRNAs accidentally target the lncRNA transcript via base-pairing with the mutated LINE-1/Alu sequence and induce silencing of the lncRNA. Another involves inadvertent SRP protein recognition of the mutated sequence, which resides in a conserved internal loop of the Alu secondary structure embedded in the lncRNA. This loop is also present in the 7SL RNA secondary structure.

Thus, this is an example of a point mutation in an embedded TE in a lncRNA sequence that produces human disease. What is not known is what exact role the unique lncRNA plays in normal cell functions and the normal function of the embedded LINE-1/Alu sequence, although there may be involvement in regulatory networks during brain development [39].

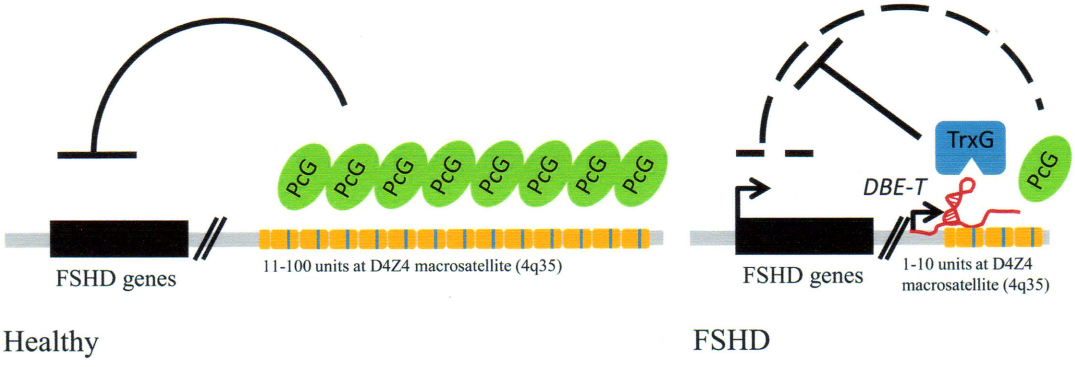
3.4. Facioscapulohumeral Muscular Dystrophy (FSHD)-Involvement of Tandem Repeats and Long ncRNA

Repeat elements in primate genomes are primarily of two types: repeats of transposable element integration into the genome and tandem repeats (microsatellites, minisatellites and macrosatellites). There is evidence that ~25 percent of minisatellite tandem repeats are derived from TEs [18]. The evolutionary origin of most macrosatellites is uncertain, however the macrosatellite tandem repeat at human chromosome region 4q35 discussed below has an interesting origin and possibly related to a retrotransposition. It represents a remnant of an unprocessed mRNA from a primate ancestral retrogene. The retrogene was lost but the retrotranscribed unprocessed mRNA (which included introns) was retained. It proceeded through an expansion in the last 25 million years and was greatly multi-copied in Homo sapiens [40–43].

The tandem repeat macrosatellite sequences at chromosome region 4q35 and a chromatin associated long ncRNA are intimately involved in a human genetic disorder termed facioscapulohumeral muscular dystrophy (FSHD) [17,44]. This is a process perhaps best described as a "cause and effect" mechanism involving loss of repeat sequences, the activation of a lncRNA and resultant epigenetic changes in FAHD patients.

Normal individuals carry a certain number of repeats (11–100) in the facioscpulohumeral (FSHD) locus. FSHD patients have a decreased number of repeats [17]. The FSHD locus maps to the chromosomal region 4q35. The region containing the repeat sequences is termed D4Z4 (Figure 4). A lncRNA termed DBE-T is partly encoded by the D4Z4 repeat locus at its 3' end. The genomic repeat sequences serve as binding sites for Polycomb proteins (PcG). PcG proteins have multiple functions but are epigenetic suppressors and are needed to suppress the *FSHD* genes that are silent in normal individuals. However, when there is a loss of genomic repeats resulting in less than 11 repeats, as is found in FSHD patients, DBE-T lncRNA transcription is activated. The RNA is normally silent and is only detected in FSHD patients. DBE-T lncRNA recruits Ash1L, a histone methytransferase to the FHDS locus, which subsequently results in transcription of locus genes due to epigenetic changes in chromatin. Thus these are changes involving the loss of repeat sequences in the D4Z4 region, transcription of a lncRNA with resultant chromatin remodeling and transcription of normally silent genes from the FSHD locus (Figure 4).

Figure 4. Involvement of repeat units (in orange) and lncRNA (in red) in FSHD muscular dystrophy. DBE-T, lncRNA; TRG, Trithorax proteins; other terms are defined in text**.** From Casa and Gabellini [17]. Reprinted with permission of publisher.



This fascinating study raises questions regarding the cell's reliance of repeat sequences for crucial epigenetic regulation. The D4Z4 region is highly variable in normal individuals, again, 11–100 copies of repeats. Genetic rearrangements appear to have take place in the 4q35 region in FSHD patients [45]. It seems that the cell's reliance on variable number of repeat sequences, as in the D4Z4 region for regulation appears to be a flawed mechanism used for epigenetic silencing. There are also other genetic abnormalities involving deletions in unstable chromosomal regions that have repeat sequences, as in chromosome locus 22q11.2 in DiGeorge Syndrome [46]. Although these repeat regions widely differ in structure and properties, e.g., DiGeorge syndrome can involve rearrangements between palindromic AT-rich repeats, it appears that genomic repeat sequences may constitute weak points in terms of maintenance of genomic fidelity, even though they have their important functions in the cell.

3.5. Embedded Alu Sequences Can Take Part in Alternative Splicing and A to I Editing in Human mRNAs

As opposed to other embedded Alu sequences described in this manuscript, the association of Alus with alternate splicing and RNA editing represent nuclear processes that Alu sequences participate in. Alu sequences are found in exons in about 5% of alternatively spliced mRNAs [47]. The presence of an Alu in exons of pre-messenger RNA transcripts can provide alternative splice sites and parts of the embedded Alu sequences can be incorporated into the processed mRNA [48–51].

Approximately 45 percent of Alu sequences are found in introns in both 5' to 3' and reverse orientations and are present in multiple copies [47,52]. These Alu sequences can potentially form double-stranded stems within a transcript when two Alu RNAs are in antiparallel orientation [53]. This enables RNA editing to take place [54] and can lead to premature stop codons or changes in codon reading [47]. Alus are prominent targets for RNA editing [54]. This is an example of how Alu RNA secondary structure can participate in altering molecular processes.

3.6. Human Endogenous Retrovirus (HERV) LTR Transcripts

About 8% of the human genome consists of human endogenous human retroviruses (HERV) [55]. HERVs cannot produce a viable virus due to mutations, but its associated LTR transposons serve a vital role in cell transcription. In addition, the human genome contains several thousand copies of single long terminal repeats (sLTRs), which originally stem from HERV [55]. These sLTRs carry no viral genes but can function as promoters and enhancers when found upstream of genes. However some sLTRs are situated in introns and are transcribed into RNA. Xu and co-workers, while studying the expression of HERV-9 U3 sLTR show that sLTR RNA transcripts are both sense and antisense RNAs, but the U3 sLTR antisense transcript can bind key transcription factors involved in cell proliferation. The sense sLTR RNA does not bind transcription factors. Importantly, malignant cells express lower levels of antisense sLTR RNA relative to sense transcripts than normal cells. The antisense sLTR RNA, which is ~550 nt appears to be a novel sLTR RNA species. The authors propose that the antisense sLTR lncRNA serves as a trap for some cell proliferation transcription factors. This may have significance in terms of a possible lack of inhibition of growth in cancer cells. Thus this is an example of a regulatory lncRNA that is encoded by a transposon (sLTR), but binds to and

inactivates proteins and not other RNAs. Importantly, this ncRNA may play a crucial role in cell proliferation [55].

3.7. Interrelatedness between HERV LTRs and Intergenic Long Non-Coding RNAs

In a different study concerning HERV LTRs, other TEs and lncRNAs, Kelly and Rinn provide a comprehensive analysis of human TE sequences in long intergenic non-coding RNAs (lincRNAs) and conversely, the presence of lincRNAs sequences in transposons [56]. About 7700 lincRNAs overlap TEs and about 1530 lincRNAs are devoid of TEs; thus about 80 percent of human lincRNAs are associated with TEs. lincRNAs display a strikingly non-random association with transposable elements; the majority overlap human endogenous retrovirus (HERV) LTRs and a small minority are associated with LINE or SINE elements.

Interesting observations were made on the orientation of HERV transposons relative to lincRNAs and expression in specific cells. A large number of HERV LTRs are situated at the transcriptional start sites (TSS) of lincRNAs and in the sense orientation. This suggests that HERV LTRs provide regulatory signals for lincRNAs. lincRNAs display a marked stem cell specificity in expression, but lincRNAs that have no LTR associations are expressed highest in testes. On the other hand, lincRNAs that contain Alu sequences are expressed in all cell lines but testes. lincRNAs most likely function in specific tissues, but Alu-containing lincRNAs may be deleterious in testes, as they are not expressed in these tissues. Thus, there is a tissue-specificity in expression [56]. TEs may work "hand-in-hand" with lincRNAs as functional units in particular cells.

3.8. Regulatory Non-Coding Circular RNAs

Circular RNAs were first characterized in human and other mammalian cells about 20 years ago [57,58], however they were initially detected in electron micrographs over 30 years ago [59]. These RNAs consist of scrambled protein coding exons, *i.e.*, the order of exons is not the same as in the genomic sequence of protein coding regions. Scrambled exon sequences were discovered in RNA transcripts in rodents and humans [60]. Subsequently, additional cirRNAs were found [61–63]. Recently, by using deep sequencing of RNA techniques and a bioinformatics approach, Saltzman *et al.* [64] discovered several hundred circular RNAs in human cells and surprisingly, Jeck *et al.* [65] using circular enrichment techniques as well as bioinformatics determined that greater than 14% of human fibroblast gene transcripts are cirRNAs (over 25,000 circular transcripts).

Although circular RNAs arise from protein-coding regions, they do not encode proteins. They are thus a separate class of long non-coding RNAs. Functions of cirRNAs were not elucidated for over 20 years since their discovery. However, the field has now moved dramatically, with two laboratories determining that some circular RNAs serve as "sponges" that can bind approximately 70 microRNAs and thus inactivate the microRNAs [6,7]. This shows that circular RNAs have regulatory functions, *i.e.*, they "regulate the regulator", the microRNAs.

Via bioinformatics analyses, it was shown that Alu elements are found in upstream and downstream introns that straddle the exons that are circularized, and that Alu sequences tended to be inverted and thus complementary [65]. Intron pairing may contribute and be essentially to circularization of exons by complementary base-pairing between Alu elements in the upstream and downstream introns. If this

is so, then Alu elements play a major role in formation of circularized RNAs. Related to this, there is precedent for Alu pairing in intons during alternative splicing [53].

3.9. piRNAs—Known Regulators of TEs

piRNAs are a class of small non-coding RNAs that are 26–31 nt. They interact with Piwi proteins, hence their name. The Piwi family is regulatory proteins that were originally defined in Drosophila as P-element induced wimpy testis [66]. piRNAs are abundantly found in germ line cells, especially in mammals, e.g., several million piRNAs are found in mammalian testes. Genetic regions that encode piRNAs consist of clusters. These clusters have repeats of piRNA sequences and there can be as many as 1000 copies of piRNAs in a cluster. piRNAs are processed from long precursors transcripts but little is known of the biogenesis of piRNAs and the number and functions of the associated proteins.

Some piRNA clusters consist of transposon or remnants of transposon sequences. Thus piRNAs can have sequences complementary to transposon sequences and can recognize their targets by base-pairing, either by perfect or imperfect base-pairing. A major role of the piRNA/Piwi protein complex in germ line cells is to protect cells from invading transposons. This is a type of "genetic immune system" that is found in both eukaryotes and prokaryotes. For example, the CRISPR complex in bacteria and archae functions by a comparable mechanism (albeit with significant variations on a theme) to protect cells from invasion by plasmids and viruses [67–69]. Both the piRNA/Piwi and CRISPR immune complexes function by an RNA-based mechanism.

piRNA functions have been studied in detail in *Drosophila*, *C. elegans* and mammalian cells. Functions are complex and may differ in different species, but a large fraction of piRNAs represent antisense transcripts to transposon transcripts [70–72]. A basic mechanism of action of piRNAs, first deduced in *Drosophila*, has the following scenario: when the cell is previously exposed to a TE but now experiences an overload of this transposon, piRNAs containing complementary sequences to the TE will base pair with and induce degradation of the TE RNA via the Piwi proteins. When the cell encounters a transposon that it has not been exposed to before, the TE by chance, may incorporate into the DNA in a piRNA-encoding cluster and thus its sequence can become part of the piRNA cluster (however, we are not aware that the probability of incorporation has been experimentally determined). Via the same mechanism mentioned above, piRNA transcripts that are antisense to the new transposon RNA induce degradation of the TE RNA via the piRNA–Piwi complex [71]. Thus, this is an immune system that helps keep transposons in check [70]. It is of interest that cell survival depends in part, on the probability of incorporation of the TE into a piRNA cluster, *vs.* the probability of insertion into and inactivation of an essential gene. However, other protective mechanisms also operate to limit TE activity.

Additional processes of piRNAs have been determined. In nematodes, piRNAs detect a TE sequence via imperfect base-pairing and then induce another small RNA class, termed 22G-RNAs to silence a transposon [73]. Some processes involve epigenetic mechanisms. For example, in Drosophila, nuclear piRNAs can target a transposon and thus direct Piwi proteins to repression chromatin and thus transcription of the TE [74]. Additionally, piRNAs may also induce the methylation of TE LINE-1 DNA in humans. This can prevent transcription of the transposon and thus assure that the TE DNA will remain dormant and not be expressed [75].

The piRNA/Piwi complex is also essential in genetic imprinting in the case involving DNA methylation of the imprinted locus Ras protein-specific guanine nucleotide-releasing factor 1 (Rasgfr1)

locus in mouse germ line cells [76]. Mutants affecting piRNA expression correlate with defects in DNA methylation of Rasgrf1. The differentially methylated region (DMR) associated with Rasgfr1 contains a LINE1 retrotransposon and sequences consisting of 23–31 nt small RNAs; these correspond to piRNAs. Yet, a different locus on chromosome 7 also has a region that produces piRNAs that have a good match to the Rasgfr1 DMR piRNA sequences. The authors propose that piRNAs generated from chromosome 7 target the retrotransposon in the DMR of the imprinted Rasgrf1 locus and that chromosome 7 piRNAs may direct methylation of the Rasgrf1 locus [76]. Thus piRNA, in addition to being a post-transcriptional regulator may also be involved in epigenetic regulation of chromosomal genes and genetic imprinting.

3.10. SINE/Alu Transcripts Function as ncRNAs in Gene Regulation at the Transcription Level

The non-autonomous retrotransposon SINE sequences are transcribed into small ncRNAs, but like some piRNAs, they can function at the level of transcription. A SINE transcript termed B2, which is found in the nucleus of mice was shown to be 177 nt. This B2 RNA transcript is conserved in rodents. This RNA binds polymerase II during the heat shock response, disrupts the polymerase/promotor interaction and represses transcription from protein gene promoters [77–79]. It is unclear whether the ncRNAs also interact with the promoter sequence. The human counter part, an Alu RNA, functions in a similar manner during heat shock, even though it's nucleotide sequence and secondary structure differs from B2 RNA [47,77]. Expression of both these small ncRNAs is increased during the heat shock response.

In other studies, a processed human Alu RNA has a sequence that is identical to that of a piRNA present in mammalian testes [80]. The processed transcript, termed piAluRNA is found in the nucleus of human adult stem cells, appears to interact with several nuclear proteins and may be involved in several processes. These include transcription, chromatin organization, organelle organization, DNA repair and cell cycle control [80]. An RNA affinity assay with synthetic oligonucleotides representing a segment of the piAluRNA and high-resolution mass spectrometry-LC-MS were used to identify interacting proteins. Functional studies are needed, but the current binding data strongly suggest the involvement of piAluRNA in several of these nuclear functions. These studies may greatly extend the roles of small ncRNAs in cells.

It is important to point out that binding and repression of proteins by ncRNAs also occurs in prokaryotes. For example, the bacterial ncRNA 6S RNA regulates RNA polymerase by binding sigma factor 70 factor and subsequently repressing RNA polymerase activity from sigma 70 promoters [81–83]. Thus, this is another example of the basic principles of molecular regulation that encompass all biological kingdoms.

4. Conclusions

We presented examples of ncRNAs originating from TEs, such as miRNAs derived from MADE1 TE. In addition, there are piRNAs that consist of TE sequences and processed SINE and/or Alu transcripts that function as small ncRNAs. Some findings show TE-derived miRNAs to be less conserved than non-TE-derived miRNAs [21], which may imply a species-specific function of TE-derived ncRNAs. There is growing evidence for the meshing of TE sequences with ncRNAs

involving both structure and function, and this association has resulted in formation of new regulatory pathways. It is obvious that TE transcripts are of an enormous asset to organisms, either as embedded sequences in ncRNAs or as individual RNAs. The interaction between TEs and ncRNAs could be looked at as a "symbiotic partnership" between the cell and transposable elements involving structure and function.

Cells offer TEs the means to multiple and maintain stability. Nevertheless, when active, TEs become overabundant and can become a threat to survival. The cell then elicits mechanisms to limit their replication. In this process, cells often use the TE sequences themselves to limit proliferation, as in the case with piRNAs involved in the genetic immune system [71].

Of special significance is that ~80% of long intergenic non-coding RNAs are associated with TEs, and in a nonrandom fashion [56]. They most likely serve functional roles, e.g., retrotransposon LTRs may provide regulatory signals for associated lincRNAs. This may be the tip of the iceberg in terms of TE/lincRNA functions as there are thousands of transcripts found in humans.

Most known ncRNA/target RNA interactions consist of short imperfect base-paired stems, and many miRNAs can bind to and regulate multiple target sequences. But this raises the question of the probability of making mistakes and targeting the wrong RNA. Other factors such as RNA-binding proteins also contribute to RNA recognition and stable formation, but the probability of mistakes must be very low, as imperfect RNA/RNA interactions appear to be highly specific. For example, the Alu-ncRNA/Alu-target RNA recognition must be as stable and specific as that of the intramolecular Watson–Crick base-paired stem recognized by Stau-1 in mRNA-induced degradation in human cells [29,30]. These stems are an interesting example of divergent RNA/RNA structures that bind the same RNA-binding protein. Three-dimensional RNP structures are needed to understand duplex conformations and protein binding sites on the two types of RNA stems.

Multiple ncRNA/target RNA pairings by the same ncRNA are also found in prokaryotic interactions with binding of ncRNAs to different target sequences and with different predicted base-pairings, e.g., see [84,85]. Thus stable short ncRNA/target RNA sequence pairings and multiple targeting are found in prokaryotes and eukaryotes. There is a beauty in the specificity and stability of small imperfect RNA/RNA pairings, and with employment of transposable elements in this binding process, and at least in eukaryotes, the cell appears to also use TE sequences to a significant extent in this form of binding.

It is of interest that some TE ncRNAs have been shown to bind proteins. For example, the lncRNA from an LTR retrotransposon situated in an intron binds transcription factors involved in cell proliferation [55]. The B2 ncRNA from an Alu transcript binds polymerase II [47]. The piAlu RNA binds nuclear proteins [80]. Thus transposon RNA transcripts show versatility in function in that these can also repress nuclear protein functions. The 6S RNA in bacteria, although not a TE transcript, is an example of a prokaryotic ncRNA that binds and inhibits the bacterial polymerase enzyme. This adds to the universality of ncRNA-related regulatory mechanisms in biological species. The 6S RNA was the first ncRNA to be sequenced [86], albeit its function was not determined until ~30 years later [81]!

The LINE-1/Alu element in a human lncRNA plays a pivotal role in formation of disease. A single mutation in the embedded TE causes human brainstem atrophy [39]. Whether the embedded LINE-1/Alu element is more prone to mutation than the rest of the lncRNA is not known, but this shows that a point mutation in a TE is the cause of atrophy and death and not a mutation in a protein

gene. This adds to the spectrum of mutations that cause human disease, *i.e.*, non-protein-coding genomic sequences can be important factors in human disease. Related to this, the deregulation of lncRNA transcription in diseases such as cancer has recently been highlighted [87,88]. As there are thousands of ncRNAs associated with TEs whose functions have not been determined, the future may possibly hold some interesting surprises with respect to diseases that may have an aberrant ncRNA etiology.

Acknowledgements

We thank Kiran Kumar Govindaiah for aid with the figure drawings, and Drs. Stefanie Mortimer and Jennifer Doudna for providing the Alu secondary structure drawing. We also are grateful to Dr. Davide Gabellini for clarification of the evolution of and molecular mechanisms pertaining to the FSDH repeat region. ND acknowledges support from the Department of Molecular Genetics and Microbiology, Stony Brook University.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Mattick, J.S. Rocking the foundations of molecular genetics. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 16400–16401.
2. Djebali, S.; Davis, C.A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F.; *et al.* Landscape of transcription in human cells. *Nature* **2012**, *489*, 101–108.
3. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **2009**, *10*, 57–63.
4. Habegger, L.; Sboner, A.; Gianoulis, T.A.; Rozowsky, J.; Agarwal, A.; Snyder, M.; Gerstein, M. RSEQtools: A modular framework to analyze RNA-Seq data using compact, anonymized data summaries. *Bioinformatics* **2011**, *27*, 281–283.
5. Derrien, T.; Johnson, R.; Bussotti, G.; Tanzer, A.; Djebali, S.; Tilgner, H.; Guernec, G.; Martin, D.; Merkel, A.; Knowles, D.G. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.* **2012**, *22*, 1775–1789.
6. Memczak, S.; Jens, M.; Elefsinioti, A.; Torti, F.; Krueger, J.; Rybak, A.; Maier, L.; Mackowiak, S.D.; Gregersen, L.H.; Munschauer, M.N.; *et al.* Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **2013**, *495*, 333–338.
7. Hansen, T.B.; Jensen, T.I.; Clausen, B.H.; Bramsen, J.B.; Finsen, B.; Damgaard, C.K.; Kjems, J. Natural RNA circles function as efficient microRNA sponges. *Nature* **2013**, *495*, 384–388.
8. Di Leva, G.; Garofalo, M. Non-Coding RNAs and Cancer. In *Oncogene and Cancer—From Bench to Clinic*; Siregar, Y., Ed.; InTech: New York, NY, USA, 2013; Chapter 14, pp. 317–358.

9. Mizuno, T.; Chou, M.Y.; Inouye, M. A unique mechanism regulating gene expression: Translational inhibition by a complementary RNA transcript (micRNA). *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 1966–1970.
10. Andersen, J.; Forst, S.A.; Zhao, K.; Inouye, M.; Delihas, N. The function of micF RNA: micF RNA is a major factor in the thermal regulation of OmpF protein in *Escherichia coli*. *J. Biol. Chem.* **1989**, *264*, 17961–17970.
11. Schmidt, M. Zheng, P.; Delihas, N. Secondary structures of *Escherichia coli* antisense *micF* RNA, the 5'-end of the target ompF mRNA, and the RNA/RNA Duplex. *Biochemistry* **1995**, *34*, 3621–3631.
12. Cowley, M.; Oakey, R.J. Transposable elements re-wire and fine-tune the transcriptome. *PLoS Genet.* **2013**, *9*, e1003234.
13. Goodier, J.L.; Kazazian, H.H., Jr. Retrotransposons revisited: The restraint and rehabilitation of parasites. *Cell* **2008**, *135*, 23–35.
14. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; *et al.* Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921.
15. Doudna, J.A. The Doudna Lab. Translational control by mRNA secondary structure—Alu-element regulated miRNA interactions. Available online: http://doudna.berkeley.edu/ (accessed on 24 June 2013).
16. De Koning, A.P.; Gu, W.; Castoe, T.A.; Batzer, M.A.; Pollock, D.D. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* **2011**, *7*, e1002384.
17. Casa, V.; Gabellini, D. A repetitive elements perspective in Polycomb epigenetics. *Front. Genet.* **2012**, *3*, 199.
18. Ahmed, M.; Liang, P. Transposable elements are a significant contributor to tandem repeats in the human genome. *Comp. Funct. Genomics* **2012**, *12*, 947089.
19. Smalheiser, N.R.; Torvik, V.I. Mammalian microRNAs derived from genomic repeats. *Trends Genet.* **2005**, *21*, 322–326.
20. Piriyapongsa, J.; Jordan, I.K. A family of human microRNA genes from miniature inverted-repeat transposable elements. *PLoS One* **2007**, *2*, e203.
21. Piriyapongsa, J.; Mariño-Ramírez, L.; Jordan, I.K. Origin and evolution of human microRNAs from transposable elements. *Genetics* **2007**, *176*, 1323–1337.
22. Devor, E.J.; Peek, A.S.; Lanier, W.; Samollow, P.B. Marsupial-specific microRNAs evolved from marsupial-specific TEs. *Gene* **2009**, *448*, 187–191.
23. Yuan, Z.; Sun, X.; Jiang, D.; Ding, Y.; Lu, Z.; Gong, L.; Liu, H.; Xie, J. Origin and evolution of a placental-specific microRNA family in the human genome. *BMC Evol. Biol.* **2010**, *10*, 346.
24. Kapitonov, V.V.; Jurka, J. MER53, a non-autonomous DNA transposon associated with a variety of functionally related defense genes in the human genome. *DNA Seq.* **1998**, *8*, 277–288.
25. Yuan, Z.; Sun, X.; Liu, H.; Xie, J. MicroRNA genes derived from repetitive elements and expanded by segmental duplication events in mammalian genomes. *PLoS One* **2011**, *6*, e17666.
26. Borchert, G.M.; Holton, N.W.; Williams, J.D.; Hernan, W.L.; Bishop, I.P.; Dembosky, J.A.; Elste, J.E.; Gregoire, N.S.; Kim, J.A.; Koehler, W.W.; *et al.* Comprehensive analysis of microRNA genomic loci identifies pervasive repetitive-element origins. *Mob. Genet. Element* **2011**, *1*, 8–17.

27. Tempel, S.; Pollet, N.; Tahi, F. ncRNAclassifier: A tool for detection and classification of transposable element sequences in RNA hairpins. *BMC Bioinforma.* **2012**, *13*, 246.
28. Ahn, K.; Gim, J.A.; Ha, H.S.; Han, K.; Kim, H.S. The novel MER transposon-derived miRNAs in human genome. *Gene* **2013**, *512*, 422–428.
29. Gong, C.; Maquat, L.E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **2011**, *470*, 284–288.
30. Gong, C.; Maquat, L.E. "Alu" strious long ncRNAs and their role in shortening mRNA half-lives. *Cell Cycle* **2011**, *10*, 1882–1883.
31. Kim, Y.K.; Furic, L.; Parisien, M.; Major, F.; DesGroseiller, L.; Maquat, L.E. Staufen1 regulates diverse classes of mammalian transcripts. *EMBO J.* **2007**, *26*, 2670–2681.
32. Capshew, C.R.; Dusenbury, K.L.; Hundley, H.A. Inverted Alu dsRNA structures do not affect localization but can alter translation efficiency of human mRNAs independent of RNA editing. *Nucleic Acids Res.* **2012**, *40*, 8637–8645.
33. Gleghorn, M.L.; Gong, C.; Kielkopf, C.L.; Maquat, L.E. Staufen1 dimerizes through a conserved motif and a degenerate dsRNA-binding domain to promote mRNA decay. *Nat. Struct. Mol. Biol.* **2013**, *20*, 515–524.
34. Park, E.; Gleghorn, M.L.; Maquat, L.E. Staufen2 functions in Staufen1-mediated mRNA decay by binding to itself and its paralog and promoting UPF1 helicase but not ATPase activity. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 405–412.
35. Wang, J.; Gong, C.; Maquat, L.E. Control of myogenesis by rodent SINE-containing lncRNAs. *Genes Dev.* **2013**, *27*, 793–804.
36. Carrieri, C.; Cimatti, L.; Biagioli, M.; Beugnet, A.; Zucchelli, S.; Fedele, S.; Pesce, E.; Ferrer, I.; Collavin, L.; Santoro, C. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* **2012**, *491*, 454–457.
37. Choi, J.; Levey, A.I.; Weintraub, S.T.; Rees, H.D.; Gearing, M.; Chin, L.S.; Li, L. Oxidative modifications and down-regulation of ubiquitin carboxyl-terminal hydrolase L1 associated with idiopathic Parkinson's and Alzheimer's diseases. *J. Biol. Chem*. **2004**, *279*, 13256–13264.
38. Barrachina, M.; Castaño, E.; Dalfó, E.; Maes, T.; Buesa, C.; Ferrer, I. Reduced ubiquitin C-terminal hydrolase-1 expression levels in dementia with Lewy bodies. *Neurobiol. Dis.* **2006**, *22*, 265–273.
39. Cartault, F.; Munier, P.; Benko, E.; Desguerre, I.; Hanein, S.; Boddaert, N.; Bandiera, S.; Vellayoudom, J.; Krejbich-Trotot, P.; Bintner, M. Mutation in a primate-conserved retrotransposon reveals a noncoding RNA as a mediator of infantile encephalopathy. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 4980–4985.
40. Hewitt, J.E.; Lyle, R.; Clark, L.N.; Valleley, E.M.; Wright, T.J.; Wijmenga, C.; van Deutekom, J.C.; Francis, F.; Sharpe, P.T.; Hofker, M.; *et al.* Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum. Mol. Genet.* **1994**, *8*, 1287–1295.
41. Clapp, J.; Mitchell, L.M.; Bolland, D.J.; Fantes, J.; Corcoran, A.E.; Scotting, P.J.; Armour, J.A.; Hewitt, J.E. Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet*. **2007**, *81*, 264–279.

42. Snider, L.; Asawachaicharn, A.; Tyler, A.E.; Geng, L.N.; Petek, L.M.; Maves, L.; Miller, D.G.; Lemmers, R.J.; Winokur, S.T.; Tawil, R. RNAtranscripts, miRNA-sizedfragments and proteinsproduced from D4Z4units: New candidates for the pathophysiology of facioscapulohumeral dystrophy. *Hum. Mol. Genet*. **2009**, *18*, 2414–2430.
43. Cabianca, D.S.; Gabellini, D. FSHD: Copy number variations on the theme of muscular dystrophy *J. Cell Biol.* **2010**, *191*, 1049–1060.
44. Cabianca, D.S.; Casa, V.; Bodega, B.; Xynos, A.; Ginelli, E.; Tanaka, Y.; Gabellini, D. A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* **2012**, 819–831.
45. Wijmenga, C.; Hewitt, J.E.; Sandkuijl, L.A.; Clark, L.N.; Wright, T.J.; Dauwerse, H.G.; Gruter, A.M.; Hofker, M.H.; Moerer, P.; Williamson, R.; *et al.* Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat. Genet.* **1992**, *2*, 26–30.
46. Emanuel, B.S. Molecular mechanisms and diagnosis of chromosome 22q11.2 rearrangements. *Dev. Disabil. Res. Rev*. **2008**, *14*, 11–18.
47. Ponicsan, S.L.; Kugel, J.F.; Goodrich, J.A. Genomic gems: SINE RNAs regulate mRNA production. *Curr. Opin. Genet. Dev*. **2010**, *20*, 149–155.
48. Sorek, R.; Ast, G.; Graur, D. Alu-containing exons are alternatively spliced. *Genome Res.* **2002**, *12*, 1060–1067.
49. Ram, O.; Schwartz, S.; Ast, G. Multifactorial interplay controls the splicing profile of Alu-derived exons. *Mol. Cell Biol*. **2008**, *28*, 3513–3525.
50. Gal-Mark, N.; Schwartz, S.; Ast, G. Alternative splicing of Alu exons—Two arms are better than one. *Nucleic Acids Res*. **2008**, *36*, 2012–2023.
51. Schwartz, S.; Gal-Mark, N.; Kfir, N.; Oren, R.; Kim, E.; Ast, G. Alu exonization events reveal features required for precise recognition of exons by the splicing machinery. *PLoS Comput. Biol.* **2009**, *5*, e1000300.
52. Dagan, T.; Sorek, R.; Sharon, E.; Ast, G.; Graur, D. AluGene: A database of Alu elements incorporated within protein-coding genes. *Nucleic Acids Res.* **2004**, *32*, D489–D492.
53. Lev-Maor, G.; Ram, O.; Kim, E.; Sela, N.; Goren, A.; Levanon, E.Y.; Ast, G. Intronic Alus influence alternative splicing. *PLoS Genet.* **2008**, *4*, e1000204.
54. Athanasiadis, A.; Rich, A.; Maas, S. Widespread A-to-I RNA editing of alu-containing mRNAs in the human transcriptome. *PLoS Biol.* **2004**, *2*, e391.
55. Xu, L.; Elkahloun, A.G.; Candotti, F.; Grajkowski, A.; Beaucage, S.L.; Petricoin, E.F.; Calvert, V.; Juhl, H.; Mills, F.; Mason, K. A novel function of RNAs arising from the long terminal repeat of human endogenous retrovirus 9 in cell cycle arrest. *J. Virol.* **2013**, *87*, 25–36.
56. Kelley, D.; Rinn, J. Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol.* **2012**, *13*, R107.
57. Cocquerelle, C.; Mascrez, B.; Hetuin, D.; Bailleul, B. Mis-splicing yields circular RNA molecules. *FASEB J.* **1993**, *7*, 155–160.
58. Capel, B.; Swain, A.; Nicolis, S.; Hacker, A.; Walter, M.; Koopman, P.; Goodfellow, P.; Lovell-Badge, R. Circular transcripts of the testis-determining gene Sry in adult mouse testis. *Cell* **1993**, *73*, 1019–1030.

59. Hsu, M.-T.; Coca-Prados, M. Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature* **1979**, *280*, 339–340.
60. Nigro, J.M.; Cho, K.R.; Fearon, E.R.; Kern, S.E.; Ruppert, J.M.; Oliner, J.D.; Kinzler, K.W.; Vogelstein, B. Scrambled exons. *Cell* **1991**, *64*, 607–613.
61. Zaphiropoulos, P.G. Circular RNAs from transcripts of the rat cytochrome P450 *2C24* gene: Correlation with exon skipping. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6536–6541.
62. Li X.-F.; Lytton, J. A circularized sodium-calcium exchanger exon 2 transcript. *J. Biol. Chem*. **1999**, *274*, 8153–8160.
63. Gualandi, F.; Trabanelli, C.; Rimessi, P.; Calzolari, E.; Toffolatti, L.; Patarnello, T.; Kunz, G.; Muntoni, F.; Ferlini, A. Multiple exon skipping and RNA circularisation contribute to the severe phenotypic expression of exon 5 dystrophin deletion. *J. Med. Genet.* **2003**, *40*, e100.
64. Salzman, J.; Gawad, C.; Wang, P.L.; Lacayo, N.; Brown, P.O. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* **2012**, *7*, e30733.
65. Jeck, W.R.; Sorrentino, J.A.; Wang, K.; Slevin, M.K.; Burd, C.E.; Liu, J.; Marzluff, W.F.; Sharpless, N.E. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **2013**, *19*, 141–157.
66. Lin, H.; Spradling, A.C. A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. *Development* **1997**, *124*, 2463–2476.
67. Karginov, F.V.; Hannon, G.J. The CRISPR system: Small RNA-guided defense in bacteria and archaea. *Mol. Cell* **2010**, *37*, 7–19.
68. Wiedenheft, B.; van Duijn, E.; Bultema, J.B.; Waghmare, S.P.; Zhou, K.; Barendregt, A.; Westphal, W.; Heck, A.J.; Boekema, E.J.; Dickman, M.J.; *et al.* RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10092–10097.
69. Sashital, D.G.; Jinek, M.; Doudna, J.A. An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. *Nat. Struct. Mol. Biol.* **2011**, *18*, 680–687.
70. Brennecke, J.; Aravin, A.A.; Stark, A.; Dus, M.; Kellis, M.; Sachidanandam, R.; Hannon, G.J. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **2007**, *128*, 1089–1103.
71. Malone, C.D.; Hannon, G.J. Small RNAs as guardians of the genome. *Cell* **2009**, *136*, 656–668.
72. Simonelig, M. Developmental functions of piRNAs and transposable elements: A *Drosophila* point-of-view. *RNA Biol.* **2011**, *8*, 754–759.
73. Bagijn, M.P.; Goldstein, L.D.; Sapetschnig, A.; Weick, E.M.; Bouasker, S.; Lehrbach, N.J.; Simard, M.J.; Miska, E.A. Function, targets, and evolution of *Caenorhabditis elegans* piRNAs. *Science* **2012**, *337*, 574–578.
74. Le Thomas, A.; Rogers, A.K.; Webster, A.; Marinov, G.K.; Liao, S.E.; Perkins, E.M.; Hur, J.K.; Aravin, A.A.; Tóth, K.F. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev.* **2013**, *27*, 390–399.
75. Sigurdsson, M.I.; Smith, A.V.; Bjornsson, H.T.; Jonsson, J.J. The distribution of a germline methylation marker suggests a regional mechanism of LINE-1 silencing by the piRNA-PIWI system. *BMC Genet.* **2012**, *13*, 31.

76. Watanabe, T.; Tomizawa, S.; Mitsuya, K.; Totoki, Y.; Yamamoto, Y.; Kuramochi-Miyagawa, S.; Iida, N.; Hoki, Y.; Murphy, P.J.; Toyoda, A.; *et al*. Role for piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse *Rasgrf1* locus. *Science* **2011**, *332*, 848–852.
77. Espinoza, C.A.; Goodrich, J.A.; Kugel, J.F. Characterization of the structure, function, and mechanism of B2 RNA, an ncRNA repressor of RNA polymerase II transcription. *RNA* **2007**, *13*, 583–596.
78. Yakovchuk, P.; Goodrich, J.A.; Kugel, J.F. B2 RNA and Alu RNA repress transcription by disrupting contacts between RNA polymerase II and promoter DNA within assembled complexes. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5569–5574.
79. Yakovchuk, P.; Goodrich, J.A.; Kugel, J.F. B2 RNA represses TFIIH phosphorylation of RNA polymerase II. *Transcription* **2011**, *2*, 45–49.
80. Blackwell, B.J.; Lopez, M.F.; Wang, J.; Krastins, B.; Sarracino, D.; Tollervey, J.R.; Dobke, M.; Jordan, I.K.; Lunyak, V.V. Protein interactions with piALU RNA indicates putative participation of retroRNA in the cell cycle, DNA repair and chromatin assembly. *Mob. Genet. Elements* **2012**, *2*, 26–35.
81. Wassarman, K.M.; Storz, G. 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* **2000**, *101*, 613–623.
82. Wassarman, K.M. 6S RNA: A regulator of transcription. *Mol. Microbiol.* **2007**, *65*, 1425–1431.
83. Decker, K.B.; Hinton, D.M. The secret to 6S: Regulating RNA polymerase by ribo-sequestration. *Mol. Microbiol.* **2009**, *73*, 137–140.
84. Holmqvist, E.; Unoson, C.; Reimegård, J.; Wagner, E.G. A mixed double negative feedback loop between the sRNA MicF and the global regulator. *Lrp. Mol. Microbiol.* **2012**, *4*, 414–427.
85. Corcoran, C.P.; Podkaminski, D.; Papenfort, K.; Urban, J.H.; Hinton, J.C.; Vogel, J. Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol. Microbiol.* **2012**, *84*, 428–445.
86. Brownlee, G.G. Sequence of 6S RNA of *E. coli*. *Nat. New Biol.* **1971**, *229*, 147–149.
87. Hauptman, N.; Glavač, D. Long non-coding RNA in cancer. *Int. J. Mol. Sci.* **2013**, *14*, 4655–4669.
88. Gutschner, T.; Diederichs, S. The Hallmarks of Cancer: A long non-coding RNA point of view. *RNA Biol.* **2012**, *9*, 703–719.

Reprinted from *IJMS*. Cite as: Hou, J.; Zhao, J. MicroRNA Regulation in Renal Pathophysiology. *Int. J. Mol. Sci.* **2013**, *14*, 13078-13092.

Review

MicroRNA Regulation in Renal Pathophysiology

Jianghui Hou [1,*] and Dan Zhao [2]

[1] Renal Division and Center for Investigation of Membrane Excitability Diseases, Washington University in St. Louis, 660 South Euclid Avenue, St. Louis, MO 63110, USA

[2] Division of Pharmacology, PLA 85th Hospital, 1328 Hua Shan Road, Shanghai 20052, China; E-Mail: zdandzjw@sina.com

***** Author to whom correspondence should be addressed; E-Mail: jhou@wustl.edu; Tel.: +1-314-362-5685; Fax: +1-314-362-8237.

Received: 7 May 2013; in revised form: 5 June 2013 / Accepted: 6 June 2013 / Published: 25 June 2013

Abstract: MicroRNAs are small, noncoding RNA molecules that regulate a considerable amount of human genes on the post-transcriptional level, and participate in many key biological processes. MicroRNA deregulation has been found associated with major kidney diseases. Here, we summarize current knowledge on the role of microRNAs in renal glomerular and tubular pathologies, with emphasis on the mesangial cell and podocyte dysfunction in diabetic nephropathy, the proximal tubular cell survival in acute kidney injury, the transport function of the thick ascending limb in Ca^{++} imbalance diseases, and the regulation of salt, K^{+} and blood pressure in the distal tubules. Identification of microRNAs and their target genes provides novel therapeutic candidates for treating these diseases. Manipulation of microRNA function with its sense or antisense oligonucleotide enables coordinated regulation of the entire downstream gene network, which has effectively ameliorated several renal disease phenotypes. The therapeutic potentials of microRNA based treatments, though promising, are confounded by major safety issues related to its target specificity, which remain to be fully elucidated.

Keywords: microRNA; kidney; diabetic nephropathy; hypercalciuria; hypertension

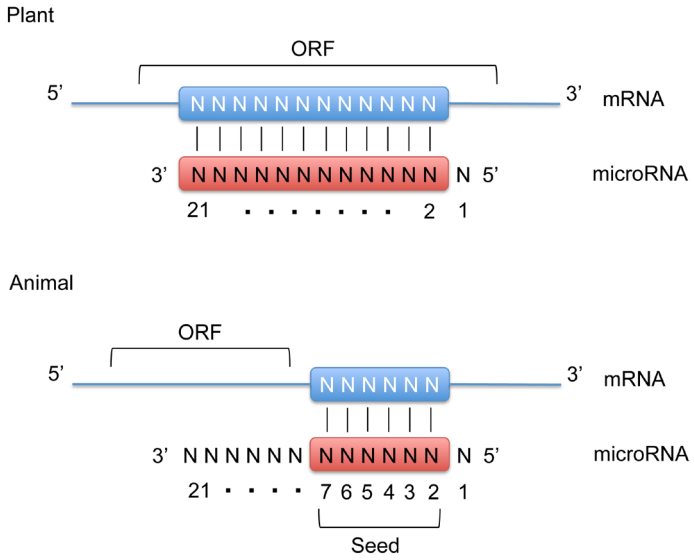
---

1. Introduction

MicroRNAs comprise a large family of 21–22-nucleotide-long RNAs that have emerged as key post-transcriptional regulators of gene expression in animals and plants [1,2]. In animals, microRNAs are predicted to control the activity of ~50% of all protein-coding genes [3]. Functional studies indicate that microRNAs participate in the regulation of almost every cellular process, and are intrinsically associated with many human pathologies.

In animals, microRNAs are processed from longer hairpin transcripts, known as pre-microRNA, by the RNase III-like enzymes Drosha and Dicer, whereas in plants only Dicer is responsible for microRNA processing [4,5]. One strand of the hairpin duplex is loaded into an Argonaute family protein (AGO) to form the core of microRNA-induced silencing complexes (miRISCs). miRISCs silence the expression of target genes through mRNA decay and translational repression. The target recognitions are achieved through base-pairing complementarity between the loaded microRNA and the target mRNA that contains a partially or fully complementary sequence [4,5]. Unlike plant microRNAs, that recognize fully complementary binding sites within the open reading frame (ORF), animal microRNAs recognize partially complementary binding sites generally located in the 3'-untranslated region (UTR) (Figure 1). For most microRNA binding sites, the complementarity is limited to the seed sequence found in the 5'-end of microRNA from nucleotide 2 to 7. The partial recognition between microRNA and its target is sufficient to trigger silencing.

Figure 1. MicroRNA target recognition mechanism. Plant microRNAs recognize fully complementary binding sites within the open reading frame (ORF) of the target mRNA. Animal microRNAs recognize partially complementary binding sites located in the 3'-untranslated region (3'-UTR). For most microRNA binding sites, the complementarity is limited to the seed sequence containing nucleotides #2–7 on the microRNA molecule. Note that in both plant and animal microRNAs the 5'-terminal nucleotide (#1) is not involved in target recognition.



The significance of microRNA in renal pathophysiology has been demonstrated in Dicer knockout animal models. During kidney development, the global knockout of Dicer in nephron progenitor cells

results in a marked decrease in nephron number [6]. Conditional removal of Dicer from the ureteric lineage results in cystic kidney disease [7]. Podocyte-specific loss of Dicer function causes proteinuria, foot process effacement, and glomerulosclerosis [8–10]. The inducible deletion of another microRNA processing enzyme, Drosha, in mature podocytes from two- to three-month-old mice, results in a similar phenotype, demonstrating a post-developmental need for microRNA activity in podocytes [11]. Deletion of Dicer in renin-secreting juxtaglomerular cells results in a selective loss of these juxtaglomerular cells, suggesting a role in cell fate determination [12]. In the proximal tubule, microRNA appear to promote cellular injury because a selective loss of Dicer in animals after three weeks of age confers resistance to ischemia-reperfusion injury [13].

2. MicroRNA in Glomerular Diseases

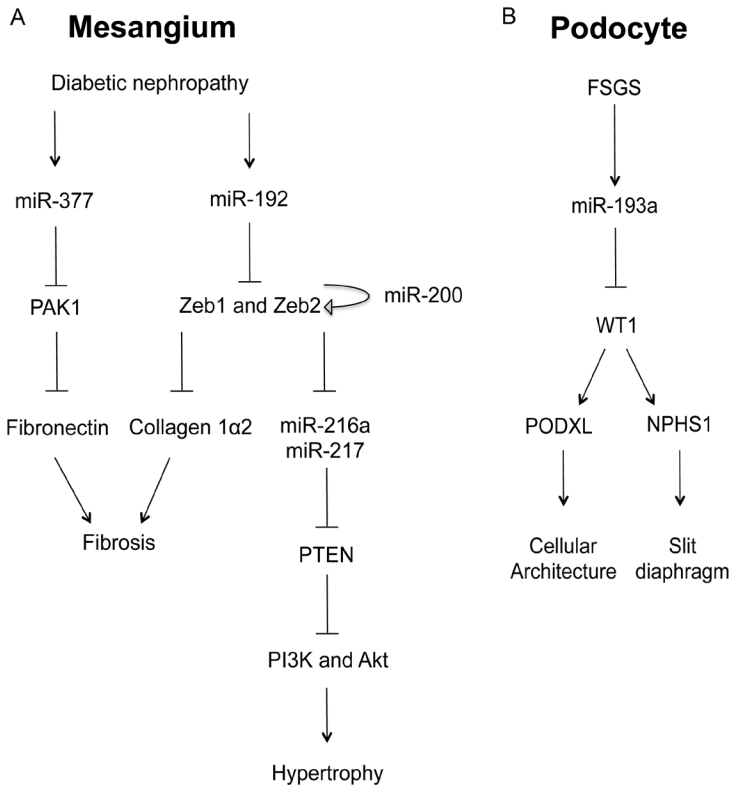
Hypertrophy and expansion in the glomerular mesangium, along with podocyte dysfunction and accumulation of extracellular matrix proteins, are major features of diabetic nephropathy, glomerulonephritis, glomerulosclerosis, and many other types of glomerular pathologies. Studies of microRNAs and their targets in glomeruli will provide critical insights of the pathogenesis of glomerular diseases and reveal new therapeutic targets for pharmacological intervention.

2.1. MicroRNA in Glomerular Mesangial Cell

Among the microRNAs highly expressed in the kidney [14,15], several key microRNAs (miR-192, miR-200b, miR-200c, miR-216a, and miR-217) were upregulated in glomerular mesangium of diabetic mouse models (type I (streptozotocin (STZ)-induced) and type2 (db/db)) (*vide infra*). *In vitro*, TGF-β-induced miR-192 was shown to increase the gene expression of collagen 1α2 by reducing the expression of two E-box repressors (Zeb1 and Zeb2) that control collagen 1α2 gene activation [16]. *In vivo*, the miR-192 and Collagen 1α2 levels were substantially increased in the mesangial cells of STZ-induced diabetic mice, as well as of db/db diabetic mice, suggesting a role in glomerular basement membrane thickening (Figure 2A) [16]. miR-216a and miR-217 are downstream targets of miR-192 through Zeb1/2 mediated mechanisms [17]. In diabetic mesangium, the cellular levels of both microRNAs were increased, resulting in the silencing of their target—PTEN (phosphatase and tensin homolog). PTEN is a phosphatidylinositol-3,4,5-triphosphate (PIP3)-phosphatase that inhibits the phosphoinositide-3-kinase pathway (PI3K) and thus prevents Akt activation [18]. Akt activation by miR-216a and miR-217 led to glomerular mesangial cell expansion and hypertrophy, another hallmark of diabetic nephropathy (Figure 2A). MiR-377 was also upregulated in cultured human and mouse mesangial cells by glucose or TGF-β treatment or in type I diabetes [19]. Overexpression of miR-377 in mesangial cells *in vitro* increased fibronectin protein production, another component of the glomerular extracellular matrix. Mechanistically, miR-377 silenced the expression of serine/threonine protein kinase PAK1, also known as p21 protein (Cdc42/Rac)-activated kinase 1, and superoxide dismutase, both of which negatively controlled fibronectin protein production (Figure 2A) [19]. The miR-200 family members are separated into two clusters located in different genomic loci. Among them, miR-200b and c are regulated by the miR-192 targets—Zeb1/2 through E-boxes in the promoters of their host genes [20,21]. miR-200b and c also target the transcripts of Zeb1/2 to auto-regulate their

own expression and amplify the signaling response leading to collagen expression and glomerular fibrosis (Figure 2A) [22].

Figure 2. The role of microRNA in glomerular diseases. (**A**) In the mesangial cell, diabetic conditions increases the levels of miR-377 and miR-192, both of which promote fibrosis and hypertrophy through the signaling cascades involving PAK1 and Zeb1/2; (**B**) In the podocyte, focal segmental glomerulosclerosis (FSGS) induces the expression of miR-193a, which in turn inhibits WT1, a master regulator of podocyte homeostasis through podocalyxin (PODXL) and nephrin (NPHS1).



IgA nephropathy is a leading cause of idiopathic glomerulonephritis and is characterized by mesangial deposition of IgA. The levels of glomerular microRNAs were deranged in patients with IgA nephropathy [23]. Among the analyzed microRNAs, the expression level of miR-200c was profoundly downregulated in these patients and negatively correlated with proteinuria, while the level of miR-192 was significantly upregulated and positively correlated with glomerulosclerosis [23].

2.2. MicroRNA in Glomerular Podocyte

Podocytes in the glomerular basement membrane are critical in the maintenance of structure and function of the glomerular filtration barrier. To study an overall role of miRNAs in podocyte biology, two independent lines of Dicer KO mice were generated for podocytes [9,10]. Mutant mice developed proteinuria by three weeks after birth and progressed rapidly to end-stage kidney disease. Multiple abnormalities were observed in glomeruli of mutant mice, including foot process effacement, irregular and split areas of the glomerular basement membrane, podocyte apoptosis and depletion, mesangial expansion, capillary dilation, and glomerulosclerosis [10]. Cytoskeletal dynamics was significantly

altered in mutant animals, including early loss of synaptopodin and downregulation of the ERM protein family (ezrin-radixin-moesin) at three weeks [9]. Gene profiling revealed upregulation of 190 genes in glomeruli isolated from mutant mice at the onset of proteinuria. Target sequences for 16 microRNA were significantly enriched in the 3'-untranslated regions of 190 upregulated genes [10]. Bioinformatic approaches were used to validate six of the eight top-candidate microRNAs, which were miR-28, miR-34a, and four members of the miR-30 family (miR-30c-1, miR-30b, miR-30d, and miR-30c-2) [10]. The miR-30 family was shown to target four genes known to be functional in podocytes, including genes that mediate podocyte apoptosis (receptor for advanced glycation end product and immediate early response 3 protein) and cytoskeletal disruption (vimentin and heat shock protein 20).

Focal segmental glomerulosclerosis (FSGS) is a devastating glomerular diseases caused by podocyte dysfunction. Deranged expression of several podocyte specific genes (WT1, NPHS1, ACTN4, and TRPC6), accompanied by collapse of normal podocyte shape and podocyte foot process effacement, presents as major pathogenic origins for FSGS. The importance of microRNAs in FSGS has been demonstrated by Gebeshuber and colleagues in a recent study [24]. Through transgenic screening in mice, Gebeshuber *et al*. have identified miR-193a as a powerful inducer of FSGS. Mechanistically, miR-193a silences the Wilms' tumor (WT1) gene, which encodes a transcriptional factor and acts as a master regulator for podocye homeostasis [24]. In normal podocytes, WT1 positively regulates the expression of several key genes crucial for podocyte architecture, e.g., podocalyxin (PODXL) and for slit diaphragm formation, e.g., nephrin (NPHS1). The level of miR-193a was consistently higher in isolated glomeruli from FSGS patients compared to normal kidneys, which provides an important mechanism for FSGS pathogenesis (Figure 2B) [24]. Deranged expression of selected microRNAs also causes podocyte abnormalities under diabetic conditions such as apoptosis and fibrosis [25,26]. In diabetic podocytes, the level of miR-195 is significantly elevated [25] while miR-29 is reduced [26]. MiR-195 targets the BCL2 gene and contributes to podocyte apoptosis via an increase in caspase-3 protein [25]. MiR-29, on the other hand, represses the expression of collagens I and IV, at both the mRNA and protein levels, by targeting the 3'-untranslated region of these genes [26].

3. MicroRNA in the Proximal Nephron

Renal ischemia-reperfusion injury (IRI) is a major cause of acute kidney injury (AKI), which is often associated with renal failure and high mortality rates [27]. The pathogenic role of microRNAs in AKI was interrogated by Wei *et al.* [13] using a proximal tubule specific Dicer KO model. Despite normal development, histology, and function of the kidney, these conditional KO mice are remarkably resistant to renal IRI, showing significantly better renal function, less tissue damage, lower tubular apoptosis rate, and higher survival rates. Microarray analyses in wildtype animals undergoing the same IRI procedure revealed changes in microRNA expression levels in the proximal tubule. Among the 173 microRNAs detected in renal cortex, miRNA-132, -362, -379, -668, and -687 showed continuous change during 12–48 h of reperfusion [13]. Another study by Godwin *et al.* demonstrated similar changes in microRNA expression during renal IRI in laboratory mice [28]. Consistent with the pathogenic hypothesis of microRNA in AKI, miR-34a was upregulated in mouse proximal tubular

cells within a few hours of cisplatin-induced nephrotoxicity [29]. Inhibition of p53 with pifithrin-α abrogated the induction of miR-34a during cisplatin treatment. Ablation of miR-34a with antisense oligonucleotides led to increased apoptosis and reduced cell survival [29]. MiR-34a is therefore a novel downstream target of p53 in response to cisplatin induced DNA damage.

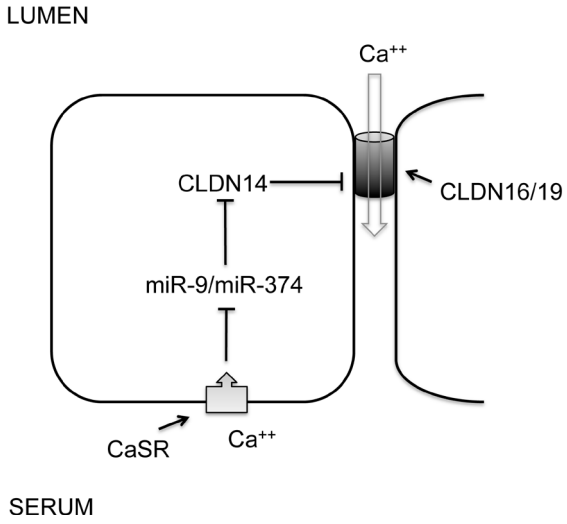
In contrast to its role in the glomerular mesangium, miR-192 appears to play a protective role against tubulointerstitial fibrosis in the proximal tubule of diabetic patients [30]. Krupa and colleagues performed a systematic study to profile microRNA expression levels in renal biopsies from patients with established diabetic nephropathy and identified 12 microRNAs showing significant differences from normal kidneys [30]. Among them, miR-192 showed the greatest change, the level of which was consistently lower in diabetic patients. In individual biopsies, miR-192 expression was inversely correlated with the progression of tubulointerstitial fibrosis and the loss of GFR. Mechanistically, overexpression of miR-192 in cultured proximal tubular cells suppressed the expression of Zeb1 and 2, opposing TGF-β-mediated epithelial-to-mesenchymal transition (EMT) [30].

4. MicroRNA in the Loop of Henle

The thick ascending limb of Henle's loop (TALH) is responsible for extracellular Ca^{++} and Mg^{++} homeostasis. A signaling cascade made of CaSR and claudins senses extracellular Ca^{++}/Mg^{++} differences and makes corresponding changes in renal excretion rates. CaSR—the Ca^{++} sensing receptor is a member of the G protein-coupled receptors (GPCRs). Mutations in the CaSR gene cause familial hypocalciuric hypercalcemia (FHH) and neonatal severe hyperparathyroidism (NSHPT), two inherited conditions characterized by altered calcium homeostasis [31]. Claudins are integral membrane proteins of the tight junction that is responsible for the paracellular transport of ions and solutes between apical and basolateral membranes. Mutations in the claudin genes, claudin-16 and claudin-19, cause familial hypomagnesemia with hypercalciuria and nephrocalcinosis (FHHNC) [32,33]. Synonymous sequence variants in the claudin-14 gene are associated with hypercalciuric kidney stones and reduced bone mineral density [34]. Several *in vitro* studies have revealed that that claudin-16 and -19 form a heteromeric cation channel in the TALH [35], whereas claudin-14 interacts with the claudin-16 and inhibits its cation permeability as a regulatory subunit [36]. Although the promoter activity and the mRNA level of claudin-14 are very high in the kidney, its protein level is surprisingly low. Gong *et al.* have identified two microRNA molecules, miR-9 and miR-374 from TALH cells, both of which recognize partially complementary binding sites located in 3'-UTRs of claudin-14 mRNA (Figure 3) [36]. Treatments with antisense oligonucleotide against miR-9 or miR-374 revealed that both microRNAs suppressed claudin-14 translation and induced its mRNA decay in a synergistic manner [36]. High Ca^{++} intake significantly downregulated the expression levels of miR-9 and miR-374 in TALH cells, which in turn causes a reciprocal increase in claudin-14 expression level. Deletion of CaSR from TALH cells abolished extracellular Ca^{++} induced changes in microRNA and claudin-14 [36]. The dietary regulation of microRNA suggests a physiological role for microRNA based signaling in the TALH of the kidney. The observed association between claudin-14 and hypercalciuric nephrolithiasis [34] can be explained by claudin-14 deregulation that escapes microRNA suppression, inhibits claudin-16/-19 channel permeabilities and phenocopies FHHNC to variable degrees. FHHNC patients [33,37] and animal models [38,39] with claudin-16 or claudin-19 mutations are known to have hypercalciuria, nephrocalcinosis, and nephrolithiasis. The regulation

of microRNA by CaSR may occur on several layers: microRNA transcription, processing, or degradation [4]. Transcriptional regulation is undoubtedly the most specific way for individual microRNA. The promoters of both miR-9 (miR-9-3 locus) and miR-374 genes contain a canonical myc-binding site (E-box: CACGTG). The transcription of miR-9-3 is upregulated by myc in human breast cancer cells [40]; miR-421/-374 cluster is upregulated by myc in HeLa cells [41]. Although these studies did not prove a role for myc in CaSR signaling, both CaSR and claudin have been implicated in tumorigenesis and metastasis [42–44].

Figure 3. The role of microRNA in renal CaSR signaling. A feedback loop of CaSR signaling in the thick ascending limb of Henle's loop (TALH) regulates urinary excretion of Ca^{++} through transcriptional regulation of miR-9 and miR-374. MicroRNAs in turn negatively regulate the expression level of claudin-14 through mRNA decay and translational repression. Claudin-14 directly binds to claudin-16 and inhibits its cation permeability to urinary Ca^{++}.



In addition to Ca^{++} metabolism, microRNAs also play a role in salt and fluid handling in the TALH. Mladinov *et al.* performed microarray assays in microdissected nephron segments and identified 12 microRNAs that were highly enriched in the TALH compared to the glomerulus and the proximal tubule [45]. Among them, miR-192 was found to target the Na^{+}/K^{+}-ATPase β1 subunit gene (Atp1b1). High salt diet increased the expression level of miR-192 in the TALH, which in turn suppressed Atp1b1 gene expression [45]. Knockdown of miR-192, *in vivo* with antisense oligonucleotide, upregulated the Atp1b1 protein level in the kidney, causing antidiuresis under high salt dietary condition. Contrasting with common microRNA target areas, miR-192 appeared to target Atp1b1 through the 5'- rather than 3'-untranslated region, although its binding sites were present within both regions [45].

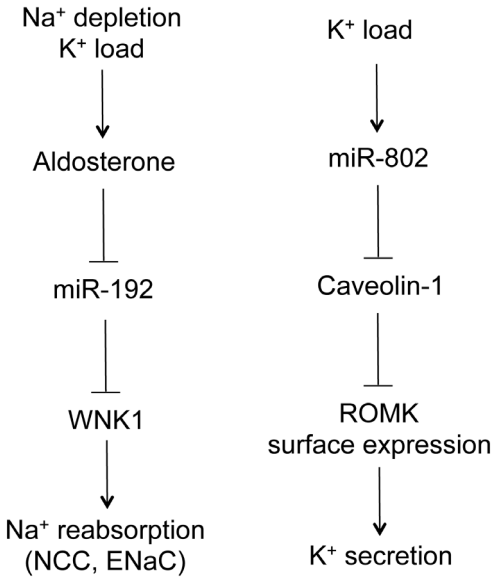
5. MicroRNA in the Distal Nephron

The aldosterone-sensitive distal nephron (ASDN) encompasses the distal convoluted tubule (DCT), the connecting tubule, and the collecting duct. It is collectively responsible for the reabsorption of approximately 5% of filtered NaCl, and plays a vital role in the regulation of extracellular fluid volume (ECFV) and blood pressure [46]. The involvement of microRNA in hypertensive kidney disease was

first studied by Sequeira-Lopez *et al.* in the juxtaglomerular cell [12]. The juxtaglomerular cell is found between the vascular pole of the renal corpuscle and the macula densa of the DCT. Its primary role in blood pressure control is the synthesis of renin. Renin is the key regulated step that initiates an enzymatic cascade that leads to angiotensin and aldosterone generation (collectively known as the rennin-angiotensin-aldosterone system). Conditional knockout of Dicer in juxtaglomerular cells resulted in a pronounced reduction in the number of juxtaglomerular cells accompanied by decreased expression of *Ren1* and *Ren2*, decreased plasma renin concentration, decreased blood pressure, and striking nephrovascular abnormalities, including striped corticomedullary fibrosis [12].

Two studies reported the role of microRNA in the regulation of ion transport by the ASDN. Elvira-Matelot *et al.* have shown that miR-192 expression is strongly reduced in the kidneys of mice treated with salt depletion, potassium load, or chronic aldosterone infusion, whereas its level is not modified by a high salt diet (Figure 4) [47]. The serine-threonine kinase WNK1 is the target of miR-192 when assayed *in vitro* and *ex vivo*. Its gene expression was reciprocally regulated by aldosterone, potassium and salt as to miR-192 in the kidney [47]. WNK1, belonging to the WNK (with no lysine-K) serine-threonine kinase subfamily, is essential for the coordinated regulation of Na^+ and K^+ transport in the kidney [48]. Mutations in WNK1 cause familial hyperkalemic hypertension (FHHt), a rare mendelian form of human hypertension [49].

Figure 4. The role of microRNA in Na^+ and K^+ handling by the distal nephron. Na^+ depletion or K^+ load increases serum aldosterone level, which decreases miR-192 levels in the aldosterone-sensitive distal nephron (ASDN). MiR-192 suppresses WNK1 gene expression, which in turn regulates Na^+ reabsorption through the NCC and ENaC channels. K^+ load also increases miR-802 levels in the ASDN. MiR-802 suppresses caveolin-1 gene expression, which regulates renal outer medullary potassium (ROMK) channel surface expression through endocytosis. ROMK mediates K^+ secretion to urine.



The renal outer medullary potassium channel (ROMK) is an ATP-dependent potassium channel encoded by the gene Kir1.1 and responsible for K^+ secretion along the ASDN. Its membrane localization is regulated by dietary K^+ intake through endocytosis-mediated mechanisms. Lin *et al.*

have revealed a physiological role for microRNA in ROMK endocytosis [50]. The ROMK channel is physically associated with caveolin-1, the principal protein component of caveolae. Expression of caveolin-1 varied inversely with the expression of ROMK in the plasma membrane, and caveolin-1 inhibited ROMK channel activity [50]. Caveolin-1 is the molecular target of miR-802 in the collecting duct. *In vitro*, expression of miR-802 suppressed the expression of caveolin-1; *in vivo*, the level of miR-802 was inversely correlated with that of caveolin-1 [50]. High K^+ intake stimulated the transcription of miR-802, which in turn decreased the expression of caveolin-1 and increased the membrane localization of ROMK, leading to higher K^+ excretion (Figure 4) [50].

MicroRNAs are also regulated by tonicity in the ASDN of the kidney. Exposure of the inner medullary collecting duct cells (mIMCD3) to a hypertonic solution induces a decrease in miR-200b and miR-717 expression level as early as two hours [51]. A common target for both microRNAs is the gene encoding the tonicity responsive element binding protein (TonEBP), also known as the osmotic response element binding protein (OREBP) [51]. TonEBP regulates many aspects of tonicity induced cellular responses, such as accumulation of inorganic osmolytes and expression of the HSP70 osmoprotective chaperone protein [52]. Depletion of microRNAs by knocking-down Dicer significantly increases TonEBP protein expression, while overexpression of miR-200b and miR-717 in mIMCD3 cells suppresses TonEBP expression on both the transcriptional and post-transcriptional levels [51]. *In vivo* in the mouse kidney, furosemide induced diuresis significantly upregulates miR-200b and miR-717, but downregulates TonEBP in the medulla [51]. The inverse correlation between the expression of microRNA and TonEBP suggests a signaling mechanism for renal cell adaptation to urinary osmolality.

6. MicroRNA as Therapeutic Candidates

Antisense oligonucleotides targeting specific microRNAs, termed “antagomirs”, were successfully used to silence endogenous microRNAs *in vivo* in experimental animals [53]. Antagomirs can be modified with methoxyethyl group (MOE) [53] or locked nucleic acid (LNA) [54] to enhance their *in vivo* stability. Once administered systemically, antagomirs are absorbed by major organs such as the liver, lungs, skin, spleen, lymph nodes, bone marrow, and kidney, excepting the brain [53–56]. The effect of a single intravenous dose of antagomir can last up to 21 d in liver and kidney [54,55]. MiR-21 regulates the ERK-MAPK signaling pathway to stimulate fibroblast survival and growth factor secretion [57]. Antagomirs against miR-21 was used to block fibrosis in cardiovascular diseases [57] and in pulmonary diseases [58]. The same strategy was also proven effective for treating renal fibrosis in chronic kidney diseases [57]. The liver-expressed miR-122 is essential for hepatitis C virus (HCV) RNA accumulation [59]. Antagomir treatments in nonhuman primates with chronic HCV infection generated long-lasting suppression of HCV viremia with no evidence of viral resistance or other adverse effects in the treated animals [60]. The most effective treatment of kidney diseases with antagomirs was demonstrated in the case of miR-192. Anti-miR-192 treatments ameliorated glomerular fibrosis in mouse models of diabetic nephropathy through a concomitant repression of collagen and fibronectin levels in the mesangial cells [61]. Another approach to interfere with the binding of a microRNA to its cognate target mRNA is the use of “microRNA sponge”. A “microRNA sponge” is a construct encoding an mRNA (e.g., the GFP mRNA) that contains a tandem repeat of microRNA

binding sites in the 3'-UTR [62]. The sponges can be introduced *in vivo* to target organs via lentivirus or retrovirus mediated transgenesis.

The major limitation to microRNA therapeutics is the off-target effect, which occurs on two distinct levels. Most current miRNA modulators display a high degree of target specificity, as point mutations made into an antagomir can completely abolish its regulatory effect both *in vitro* and *in vivo* [63]. Nevertheless, the antagomir may still target a class of related microRNAs that differ only at one or two nucleotide loci. Additional ambiguity derives from the relatively modest inhibitory effects of individual microRNAs on mRNA targets. For example, a complete depletion of a microRNA often results in maximal increases (<1.5-fold) in the expression levels of its mRNA targets, suggesting that it is the cumulative impact of small changes in the expression of myriad targets—rather than pronounced changes in a single target—that mediates the biological actions of microRNAs. A more perplexing matter is the undesired effect of systemically administrated antagomirs in off-target organs or off-target cells in the target organ. Local administration of antagomirs represents a potential solution. For example, intracerebral injection of anti-miR-16 has been demonstrated to reduce miR-16 levels in the brain, which was, however, immune to any systemic manipulation [63]. Coupling antagomirs to tissue specific antibodies provides another means to manipulating microRNA levels locally. Upstream signals intervening with microRNA expression or maturation may be unique to each target organ, thus providing new layers of specificity in microRNA-based therapeutics. Despite these potential challenges, microRNA therapeutics have recently entered Phase I and Phase II clinical trials of Santaris Pharma's antagomir anti-miR-122, *miravirsen*, for the treatment of HCV.

In some cases, a gain-of-microRNA function is beneficial to treating diseases. This can be achieved by use of a synthetic double-stranded precursor microRNA molecule, known as the "miRNA mimic". In acute myocardial infarction, for example, miR-29 is significantly downregulated in fibroblast cells of the heart, resulting in cardiac fibrosis. Introducing miR-29 mimics profoundly reduced collagen expression and ameliorated fibrotic phenotypes in the heart [64].

MicroRNAs have also been found in blood and urine samples and have emerged as potential biomarkers for various diseases including kidney diseases [65]. Nevertheless, no association study is currently available to link the expression levels of these markers with disease progression and their power for predicting disease susceptibility is still debated.

7. Conclusions

MicroRNA has emerged to be a critical regulator underlying a diverse range of renal pathophysiologies. Owing to its short seed sequence, a cognate microRNA regulates multiple gene targets, making it a powerful signaling molecule to coordinate various cellular functions. Manipulation of tissue microRNA level as a novel therapeutic approach has been proven effective in several renal disease models. Systematic identification of downstream microRNA target genes will improve our knowledge in renal pathology and supply additional candidates for therapeutic intervention. What remain largely unknown are the mechanisms that endogenously control microRNA metabolism and the clinical evidence that derangement of such mechanisms causes human diseases.

Acknowledgments

This work was supported by the National Institutes of Health Grants RO1DK084059 and P30 DK079333, and American Heart Association Grant 0930050N to JH.

Conflict of Interests

The authors declare no conflict of interests.

References

1. Bartel, D.P. MicroRNAs: Target recognition and regulatory functions. *Cell* **2009**, *136*, 215–233.
2. Carthew, R.W.; Sontheimer, E.J. Origins and mechanisms of miRNAs and siRNAs. *Cell* **2009**, *136*, 642–655.
3. Rajewsky, N. microRNA target predictions in animals. *Nat. Genet.* **2006**, *38*, S8–S13.
4. Krol, J.; Loedige, I.; Filipowicz, W. The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* **2010**, *11*, 597–610.
5. Voinnet, O. Origin, biogenesis, and activity of plant microRNA. *Cell* **2009**, *136*, 669–687.
6. Ho, J.; Pandey, P.; Schatton, T.; Sims-Lucas, S.; Khalid, M.; Frank, M.H.; Hartwig, S.; Kreidberg, J.A. The pro-apoptotic protein Bim is a microRNA target in kidney progenitors. *J. Am. Soc. Nephrol.* **2011**, *22*, 1053–1063.
7. Nagalakshmi, V.K.; Ren, Q.; Pugh, M.M.; Valerius, M.T.; McMahon, A.P.; Yu, J. Dicer regulates the development of nephrogenic and ureteric compartments in the mammalian kidney. *Kidney Int.* **2011**, *79*, 317–330.
8. Ho, J.; Ng, K.H.; Rosen, S.; Dostal, A.; Gregory, R.I.; Kreidberg, J.A. Podocyte-specific loss of functional microRNAs leads to rapid glomerular and tubular injury. *J. Am. Soc. Nephrol.* **2008**, *19*, 2069–2075.
9. Harvey, S.J.; Jarad, G.; Cunningham, J.; Goldberg, S.; Schermer, B.; Harfe, B.D.; McManus, M.T.; Benzing, T.; Miner, J.H. Podocyte-specificdeletionofdicer alters cytoskeletal dynamics and causes glomerular disease. *J. Am. Soc. Nephrol.* **2008**, *19*, 2150–2158.
10. Shi, S.; Yu, L.; Chiu, C.; Sun, Y.; Chen, J.; Khitrov, G.; Merkenschlager, M.; Holzman, L.B.; Zhang, W.; Mundel, P.; *et al.* Podocyte-selective deletion of dicer induces proteinuria and glomerulosclerosis. *J. Am. Soc. Nephrol.* **2008**, *19*, 2159–2169.
11. Zhdanova, O.; Srivastava, S.; Di, L.; Li, Z.; Tchelebi, L.; Dworkin, S.; Johnstone, D.B.; Zavadil, J.; Chong, M.M.; Littman, D.R.; *et al.* The inducible deletion of Drosha and microRNAs in mature podocytes results in a collapsing glomerulopathy. *Kidney Int.* **2011**, *80*, 719–730.
12. Sequeira-Lopez, M.L.; Weatherford, E.T.; Borges, G.R.; Monteagudo, M.C.; Pentz, E.S.; Harfe, B.D.; Carretero, O.; Sigmund, C.D.; Gomez, R.A. The microRNA-processing enzyme dicer maintains juxtaglomerular cells. *J. Am. Soc. Nephrol.* **2010**, *21*, 460–467.
13. Wei, Q.; Bhatt, K.; He, H.Z.; Mi, Q.S.; Haase, V.H.; Dong, Z. Targeted deletion of Dicer from proximal tubules protects against renal ischemia reperfusion injury. *J. Am. Soc. Nephrol.* **2010**, *21*, 756–761.

14. Sun, Y.; Koo, S.; White, N.; Peralta, E.; Esau, C.; Dean, N.M.; Perera, R.J. Development of a micro-array to detect human and mouse microRNAs and characterization of expression in human organs. *Nucleic Acids Res.* **2004**, *32*, e188.
15. Tian, Z.; Greene, A.S.; Pietrusz, J.L.; Matus, I.R.; Liang, M. MicroRNA-target pairs in the rat kidney identified by microRNA microarray, proteomic, and bioinformatics analysis. *Genome Res.* **2008**, *18*, 404–411.
16. Kato, M.; Zhang, J.; Wang, M.; Lanting, L.; Yuan, H.; Rossi, J.J.; Natarajan, R. MicroRNA-192 in diabetic kidney glomeruli and its function in TGF-beta-induced collagen expression via inhibition of E-box repressors. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3432–3437.
17. Kato, M.; Putta, S.; Wang, M.; Yuan, H.; Lanting, L.; Nair, I.; Gunn, A.; Nakagawa, Y.; Shimano, H.; Todorov, I.; *et al.* TGF-beta activates Akt kinase through a microRNA-dependent amplifying circuit targeting PTEN. *Nat. Cell Biol.* **2009**, *11*, 881–889.
18. Jiang, B.H.; Liu, L.Z. PI3K/PTEN signaling in angiogenesis and tumorigenesis. *Adv. Cancer Res.* **2009**, *102*, 19–65.
19. Wang, Q.; Wang, Y.; Minto, A.W.; Wang, J.; Shi, Q.; Li, X.; Quigg, R.J. MicroRNA-377 is up-regulated and can lead to increased fibronectin production in diabetic nephropathy. *FASEB J.* **2008**, *22*, 4126–4135.
20. Bracken, C.P.; Gregory, P.A.; Kolesnikoff, N.; Bert, A.G.; Wang, J.; Shannon, M.F.; Goodall, G.J. A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer Res.* **2008**, *68*, 7846–7854.
21. Burk, U.; Schubert, J.; Wellner, U.; Schmalhofer, O.; Vincan, E.; Spaderna, S.; Brabletz, T. A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep.* **2008**, *9*, 582–589.
22. Kato, M.; Arce, L.; Wang, M.; Putta, S.; Lanting, L.; Natarajan, R. A microRNA circuit mediates transforming growth factor-β1 autoregulation in renal glomerular mesangial cells. *Kidney Int.* **2011**, *80*, 358–368.
23. Wang, G.; Kwan, B.C.; Lai, F.M.; Choi, P.C.; Chow, K.M.; Li, P.K.; Szeto, C.C. Intrarenal expression of microRNAs in patients with IgA nephropathy. *Lab. Invest.* **2010**, *90*, 98–103.
24. Gebeshuber, C.A.; Kornauth, C.; Dong, L.; Sierig, R.; Seibler, J.; Reiss, M.; Tauber, S.; Bilban, M.; Wang, S.; Kain, R.; *et al.* Focal segmental glomerulosclerosis is induced by microRNA-193a and its downregulation of WT1. *Nat. Med.* **2013**, *19*, 481–487.
25. Chen, Y.Q.; Wang, X.X.; Yao, X.M.; Zhang, D.L.; Yang, X.F.; Tian, S.F.; Wang, N.S. MicroRNA-195 promotes apoptosis in mouse podocytes via enhanced caspase activity driven by BCL2 insufficiency. *Am. J. Nephrol.* **2011**, *34*, 549–559.
26. Wang, B.; Komers, R.; Carew, R.; Winbanks, C.E.; Xu, B.; Herman-Edelstein, M.; Koh, P.; Thomas, M.; Jandeleit-Dahm, K.; Gregorevic, P.; *et al.* Suppression of microRNA-29 expression by TGF-β1 promotes collagen expression and renal fibrosis. *J. Am. Soc. Nephrol.* **2012**, *23*, 252–265.
27. Schrier, R.W.; Wang, W.; Poole, B.; Mitra, A. Acute renal failure: Definitions, diagnosis, pathogenesis, and therapy. *J. Clin. Invest.* **2004**, *114*, 5–14.

28. Godwin, J.G.; Ge, X.; Stephan, K.; Jurisch, A.; Tullius, S.G.; Iacomini, J. Identification of a microRNA signature of renal ischemia reperfusion injury. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14339–14344.
29. Bhatt, K.; Zhou, L.; Mi, Q.S.; Huang, S.; She, J.X.; Dong, Z. MicroRNA-34a is induced via p53 during cisplatin nephrotoxicity and contributes to cell survival. *Mol. Med.* **2010**, *16*, 409–416.
30. Krupa, A.; Jenkins, R.; Luo, D.D.; Lewis, A.; Phillips, A.; Fraser, D. Loss of microRNA-192 promotes fibrogenesis in diabetic nephropathy. *J. Am. Soc. Nephrol.* **2010**, *21*, 438–447.
31. Pollak, M.R.; Brown, E.M.; Chou, Y.H.; Hebert, S.C.; Marx, S.J.; Steinmann, B.; Levi, T.; Seidman, C.E.; Seidman, J.G. Mutations in the human Ca(2+)-sensing receptor gene cause familial hypocalciuric hypercalcemia and neonatal severe hyperparathyroidism. *Cell* **1993**, *75*, 1297–1303.
32. Simon, D.B.; Lu, Y.; Choate, K.A.; Velazquez, H.; Al-Sabban, E.; Praga, M.; Casari, G.; Bettinelli, A.; Colussi, G.; Rodriguez-Soriano, J.; *et al.* Paracellin-1, a renal tight junction protein required for paracellular Mg^{2+} resorption. *Science* **1999**, *285*, 103–106.
33. Konrad, M.; Schaller, A.; Seelow, D.; Pandey, A.V.; Waldegger, S.; Lesslauer, A.; Vitzthum, H.; Suzuki, Y.; Luk, J.M.; Becker, C.; *et al.* Mutations in the tight-junction gene claudin 19 (CLDN19) are associated with renal magnesium wasting, renal failure, and severe ocular involvement. *Am. J. Hum. Genet.* **2006**, *79*, 949–957.
34. Thorleifsson, G.; Holm, H.; Edvardsson, V.; Walters, G.B.; Styrkarsdottir, U.; Gudbjartsson, D.F.; Sulem, P.; Halldorsson, B.V.; de Vegt, F.; d'Ancona, F.C.; *et al.* Sequence variants in the CLDN14 gene associate with kidney stones and bone mineral density. *Nat. Genet.* **2009**, *41*, 926–930.
35. Hou, J.; Renigunta, A.; Konrad, M.; Gomes, A.S.; Schneeberger, E.E.; Paul, D.L.; Waldegger, S.; Goodenough, D.A. Claudin-16 and claudin-19 interact and form a cation-selective tight junction complex. *J. Clin. Invest.* **2008**, *118*, 619–628.
36. Gong, Y.; Renigunta, V.; Himmerkus, N.; Zhang, J.; Renigunta, A.; Bleich, M.; Hou, J. Claudin-14 regulates renal Ca^{++} transport in response to CaSR signalling via a novel microRNA pathway. *EMBO J.* **2012**, *31*, 1999–2012.
37. Weber, S.; Schneider, L.; Peters, M.; Misselwitz, J.; Rönnefarth, G.; Böswald, M.; Bonzel, K.E.; Seeman, T.; Suláková, T.; Kuwertz-Bröking, E.; *et al.* Novel paracellin-1 mutations in 25 families with familial hypomagnesemia with hypercalciuria and nephrocalcinosis. *J. Am. Soc. Nephrol.* **2001**, *12*, 1872–1881.
38. Hou, J.; Shan, Q.; Wang, T.; Gomes, A.S.; Yan, Q.; Paul, D.L.; Bleich, M.; Goodenough, D.A. Transgenic RNAi depletion of claudin-16 and the renal handling of magnesium. *J. Biol. Chem.* **2007**, *282*, 17114–17122.
39. Hou, J.; Renigunta, A.; Gomes, A.S.; Hou, M.; Paul, D.L.; Waldegger, S.; Goodenough, D.A. Claudin-16 and claudin-19 interaction is required for their assembly into tight junctions and for renal reabsorption of magnesium. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15350–15355.
40. Ma, L.; Young, J.; Prabhala, H.; Pan, E.; Mestdagh, P.; Muth, D.; Teruya-Feldstein, J.; Reinhardt, F.; Onder, T.T.; Valastyan, S.; *et al.* miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. *Nat. Cell Biol.* **2010**, *12*, 247–256.
41. Hu, H.; Du, L.; Nagabayashi, G.; Seeger, R.C.; Gatti, R.A. ATM is down-regulated by *N*-Myc-regulated microRNA-421. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1506–1511.

42. Brennan, S.C.; Thiem, U.; Roth, S.; Aggarwal, A.; Fetahu, I.S.; Tennakoon, S.; Gomes, A.R.; Brandi, M.L.; Bruggeman, F.; Mentaverri, R.; *et al.* Calcium sensing receptor signalling in physiology and cancer. *Biochim. Biophys. Acta* **2012**, *1833*, 1732–1744.
43. Buchert, M.; Papin, M.; Bonnans, C.; Darido, C.; Raye, W.S.; Garambois, V.; Pélegrin, A.; Bourgaux, J.F.; Pannequin, J.; Joubert, D.; *et al.* Symplekin promotes tumorigenicity by up-regulating claudin-2 expression. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 2628–2633.
44. Nübel, T.; Preobraschenski, J.; Tuncay, H.; Weiss, T.; Kuhn, S.; Ladwein, M.; Langbein, L.; Zöller, M. Claudin-7 regulates EpCAM-mediated functions in tumor progression. *Mol. Cancer Res.* **2009**, *7*, 285–299.
45. Mladinov, D.; Liu, Y.; Mattson, D.L.; Liang, M. MicroRNAs contribute to the maintenance of cell-type-specific physiological characteristics: miR-192 targets Na^+/K^+-ATPase β1. *Nucleic Acids Res.* **2013**, *41*, 1273–1283.
46. Guyton, A.C. Blood pressure control—Special role of the kidneys and body fluids. *Science* **1991**, *252*, 1813–1816.
47. Elvira-Matelot, E.; Zhou, X.O.; Farman, N.; Beaurain, G.; Henrion-Caude, A.; Hadchouel, J.; Jeunemaitre, X. Regulation of WNK1 expression by miR-192 and aldosterone. *J. Am. Soc. Nephrol.* **2010**, *21*, 1724–1731.
48. McCormick, J.A.; Yang, C.L.; Ellison, D.H. WNK kinases and renal sodium transport in health and disease: An integrated view. *Hypertension* **2008**, *51*, 588–596.
49. Wilson, F.H.; Disse-Nicodeme, S.; Choate, K.A.; Ishikawa, K.; Nelson-Williams, C.; Desitter, I.; Gunel, M.; Milford, D.V.; Lipkin, G.W.; Achard, J.M.; *et al.* Human hypertension caused by mutations in WNK kinases. *Science* **2001**, *293*, 1107–1112.
50. Lin, D.H.; Yue, P.; Pan, C.; Sun, P.; Wang, W.H. MicroRNA 802 stimulates ROMK channels by suppressing caveolin-1. *J. Am. Soc. Nephrol.* **2011**, *22*, 1087–1098.
51. Huang, W.; Liu, H.; Wang, T.; Zhang, T.; Kuang, J.; Luo, Y.; Chung, S.S.; Yuan, L.; Yang, J.Y. Tonicity-responsive microRNAs contribute to the maximal induction of osmoregulatory transcription factor OREBP in response to high-NaCl hypertonicity. *Nucleic Acids Res.* **2011**, *39*, 475–485.
52. Burg, M.B.; Ferraris, J.D.; Dimitrieva, N.I. Cellular response to hyperosmotic stresses. *Physiol. Rev.* **2007**, *87*, 1441–1474.
53. Krützfeldt, J.; Rajewsky, N.; Braich, R.; Rajeev, K.G.; Tuschl, T.; Manoharan, M.; Stoffel, M. Silencing of microRNAs *in vivo* with "antagomirs". *Nature* **2005**, *438*, 685–689.
54. Obad, S.; dos Santos, C.O.; Petri, A.; Heidenblad, M.; Broom, O.; Ruse, C.; Fu, C.; Lindow, M.; Stenvang, J.; Straarup, E.M.; *et al.* Silencing of microRNA families by seed-targeting tiny LNAs. *Nat. Genet.* **2011**, *43*, 371–378.
55. Castoldi, M.; Vujic Spasic, M.; Altamura, S.; Elmén, J.; Lindow, M.; Kiss, J.; Stolte, J.; Sparla, R.; D'Alessandro, L.A.; Klingmüller, U.; *et al.* The liver-specific microRNA miR-122 controls systemic iron homeostasis in mice. *J. Clin. Invest.* **2011**, *121*, 1386–1396.
56. Yi, R.; Poy, M.N.; Stoffel, M.; Fuchs, E. A skin microRNA promotes differentiation by repressing "stemness". *Nature* **2008**, *452*, 225–229.

57. Thum, T.; Gross, C.; Fiedler, J.; Fischer, T.; Kissler, S.; Bussen, M.; Galuppo, P.; Just, S.; Rottbauer, W.; Frantz, S.; *et al.* MicroRNA-21 contributes to myocardial disease by stimulating MAP kinase signalling in fibroblasts. *Nature* **2008**, *456*, 980–984.
58. Liu, G.; Friggeri, A.; Yang, Y.; Milosevic, J.; Ding, Q.; Thannickal, V.J.; Kaminski, N.; Abraham, E. miR-21 mediates fibrogenic activation of pulmonary fibroblasts and lung fibrosis. *J. Exp. Med.* **2010**, *207*, 1589–1597.
59. Jopling, C.L.; Yi, M.; Lancaster, A.M.; Lemon, S.M.; Sarnow, P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* **2005**, *309*, 1577–1581.
60. Lanford, R.E.; Hildebrandt-Eriksen, E.S.; Petri, A.; Persson, R.; Lindow, M.; Munk, M.E.; Kauppinen, S.; Ørum, H. Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. *Science* **2010**, *327*, 198–201.
61. Putta, S.; Lanting, L.; Sun, G.; Lawson, G.; Kato, M.; Natarajan, R. Inhibiting microRNA-192 ameliorates renal fibrosis in diabetic nephropathy. *J. Am. Soc. Nephrol.* **2012**, *23*, 458–469.
62. Ebert, M.S.; Neilson, J.R.; Sharp, P.A. MicroRNA sponges: Competitive inhibitors of small RNAs in mammalian cells. *Nat. Methods* **2007**, *4*, 721–726.
63. Krützfeldt, J.; Kuwajima, S.; Braich, R.; Rajeev, K.G.; Pena, J.; Tuschl, T.; Manoharan, M.; Stoffel, M. Specificity, duplex degradation and subcellular localization of antagomirs. *Nucleic Acids Res.* **2007**, *35*, 2885–2892.
64. Van Rooij, E.; Sutherland, L.B.; Thatcher, J.E.; DiMaio, J.M.; Naseem, R.H.; Marshall, W.S.; Hill, J.A.; Olson, E.N. Dysregulation of microRNAs after myocardial infarction reveals a role of miR-29 in cardiac fibrosis. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 13027–13032.
65. Lorenzen, J.M.; Thum, T. Circulating and urinary microRNAs in kidney disease. *Clin. J. Am. Soc. Nephrol.* **2012**, *7*, 1528–1533.

Reprinted from *IJMS*. Cite as: Martini, P.; Sales, G.; Calura, E.; Brugiolo, M.; Lanfranchi, G.; Romualdi, C.; Cagnin, S. Systems Biology Approach to the Dissection of the Complexity of Regulatory Networks in the *S. scrofa* Cardiocirculatory System. *Int. J. Mol. Sci.* **2013**, *14*, 23160-23187.

Article

Systems Biology Approach to the Dissection of the Complexity of Regulatory Networks in the *S. scrofa* Cardiocirculatory System

Paolo Martini [1], Gabriele Sales [1], Enrica Calura [1], Mattia Brugiolo [2], Gerolamo Lanfranchi [1,2], Chiara Romualdi [1,*] and Stefano Cagnin [1,2,*]

[1] Department of Biology, University of Padova, Via G. Colombo 3, Padova 35121, Italy; E-Mails: paolo.martini@unipd.it (P.M.); gabriele.sales@unipd.it (G.S.); enrica.calura@unipd.it (E.C.); gerolamo.lanfranchi@unipd.it (G.L.)

[2] C.R.I.B.I. Biotechnology Centre, University of Padova, Via U. Bassi 58/B, Padova 35121, Italy; E-Mail: brugiolo@mpi-cbg.de

***** Authors to whom correspondence should be addressed; E-Mails: chiara.romualdi@unipd.it (C.R.); stefano.cagnin@unipd.it (S.C.); Tel.: +39-049-827-7401 (C.R.); +39-049-827-6162 (S.C.); Fax: +39-049-827-6159 (C.R. & S.C.).

Received: 9 August 2013; in revised form: 23 October 2013 / Accepted: 2 November 2013 / Published: 21 November 2013

Abstract: Genome-wide experiments are routinely used to increase the understanding of the biological processes involved in the development and maintenance of a variety of pathologies. Although the technical feasibility of this type of experiment has improved in recent years, data analysis remains challenging. In this context, gene set analysis has emerged as a fundamental tool for the interpretation of the results. Here, we review strategies used in the gene set approach, and using datasets for the pig cardiocirculatory system as a case study, we demonstrate how the use of a combination of these strategies can enhance the interpretation of results. Gene set analyses are able to distinguish vessels from the heart and arteries from veins in a manner that is consistent with the different cellular composition of smooth muscle cells. By integrating microRNA elements in the regulatory circuits identified, we find that vessel specificity is maintained through specific miRNAs, such as miR-133a and miR-143, which show anti-correlated expression with their mRNA targets.

Keywords: pathway analysis; miRNA; cardiocirculatory; network reconstruction; integrative analysis; pig; artery; vein; vessel

1. Introduction

Genome-wide experiments on RNA expression typically provide lists of differentially expressed genes (DEGs) [1,2] that represent the starting point of a highly challenging process of result interpretation in which the gene-by-gene approach is often used. The lists obtained are highly dependent on the statistical tests adopted and on the threshold used to declare a gene significant. This variability has raised substantial criticism concerning the reproducibility of array experiments. Several studies have demonstrated greater consistency of array results using gene set approaches, rather than single gene approaches [3], indicating that there is greater reproducibility of the main biological themes than of their single elements. A gene set is defined as a set of genes that are functionally related. Gene sets are usually identified based on a priori biological knowledge (see, for example, Gene Ontology "GO" (http://www.geneontology.org/ (accessed on 13 November 2013)) and the Kyoto Encyclopedia of Genes and Genomes "KEGG" (http://www.genome.jp/kegg/ (accessed on 13 November 2013))). In this regard, several new bioinformatics tools have been developed that allow the integration of information such as gene location [4–6], ontological annotations [7–10], or sequence features [11]. These methods can be broadly divided into supervised and unsupervised approaches. Supervised methods use *a priori* information on the functional relationships among genes to identify the processes involved in an experimental condition, while unsupervised approaches attempt to reconstruct functional associations among genes without relying on external information. In the following, we will briefly review these strategies, focusing specifically on their pros and cons; in addition, we will apply these strategies to a case study.

1.1. Supervised Approaches: Pathway Analysis

The integration of gene expression profiles with additional information on pathway annotations is called pathway analysis. The pathway analysis approach evaluates gene expression profiles among related genes, looking for coordinated changes in their expression levels. Several implementations of pathway analysis are now available, from the widely used algorithm developed by Subramanian and colleagues (Gene Set Enrichment Analysis; GSEA) [9], with its improvements [10,12], to more sophisticated implementations that exploit the topology of the pathway [13,14] (for a comprehensive review of existing methods, see [15]). Pathway analysis methods can be divided into (i) methods based on enrichment analysis and performed on a list of genes selected through a gene-level test; and (ii) methods based on global and multivariate approaches that define a model based on the whole gene set. With the first class of methods, the primary concerns are the assumption that genes are independent and the use of a threshold value for the selection of differentially expressed genes. Due to the latter, many genes with moderate but meaningful expression changes are discarded based on the strict cut-off value, leading to a reduction in statistical power. On the other hand, global and multivariate approaches relax the assumption of independence among genes belonging to the same

gene sets and identify moderate but coordinated expression changes that cannot be detected by the enrichment analysis approach [16].

From this perspective, we recently developed three novel algorithms that can be used to perform gene set and pathway analysis. Graphite, a Bioconductor package [17], is a computational framework that can be used to manage, interpret, and convert pathway annotations to gene-gene networks, while STEPath [18] integrates expression levels and chromosome positioning to identify regional gene activation and CliPPER [14,19] explores the topology of a pathway, highlighting the portions most involved in its deregulation. We have implemented most of these analyses in a new web tool called GraphiteWeb [20].

One of the major drawbacks associated with these approaches is the limitation of pathway annotation. Pathway annotation is a highly challenging procedure that exploits the efforts of many researchers, who manually curate each single pathway based on information available in the literature. Pathways are often thought of as the elementary functional and evolutionary building blocks of the complete metabolic network, with each pathway representing a "self-contained" elementary biochemical process. To partition the reaction network of an organism into a set of (possibly overlapping) metabolic pathways requires arbitrary decisions as to where such partitions should be made and how pathway variants should be described [21]. For these reasons, only a portion (in humans, approximately one-third) of known genes are currently annotated in at least one pathway.

In KEGG [22], the metabolic pathways—called "maps"—are subparts of the overall reaction graph. Reactions within a map are connected by their constituent metabolites, which also provide links to reactions in other maps. KEGG metabolic maps are described without reference to a particular species, and each map includes the reactions belonging to all known variants of a particular pathway. MetaCyc is a database of non-redundant, experimentally elucidated metabolic pathways that are found in many species [23] while, in the smaller Reactome database [24], the human database is used as the reference for predicting reactions and pathways in other organisms.

1.2. Unsupervised Approaches: Reverse Engineering Approach

A different approach to dealing with biological networks is the *ab initio* strategy: using genome-wide expression values, these algorithms try to infer the best network of interactions satisfying specific conditions. Unlike the pathway analysis approach, here, all known genes can be taken into consideration. Several methods have been proposed for the reconstruction of gene regulatory networks (GRNs) from experimental data; these include Bayesian Networks (BN) [25], Relevance Networks (RN) [26], and Graphical Gaussian Models (GGM) [27,28]. While BN and GGM distinguish between direct and indirect edges, RN does not. It is worth noting that although BN and GGM are able to infer edge direction this does not necessarily imply an ability to identify biological causality.

BN and GGM function poorly in cases involving thousands of genes and a small number of replicates, while RN has the ability to address such cases. RN uses association measures between two expression profiles, such as correlation and mutual information, to rank gene-gene interactions according to their strengths; the higher the association measure, the greater the probability of a functional interaction between the two genes. All of these approaches produce a large number of false

positives (false interactions). The seminal paper of Basso *et al.*, 2005 [29], extends RN, introducing an algorithm based on Data Processing Inequality (DPI) for removing indirect edges. Their approach, called ARACNE (Algorithm for the Reconstruction of Accurate Cellular Networks) [30], has been successfully used to reconstruct the sub-network of the MYC gene in human B cells.

In this context, we developed a new R package, *parmigene*, that performs network inference by implementing an unbiased estimation of the mutual information between expression profiles, thus yielding more precise results than existing software at strikingly less computational cost [31].

Apart from their low specificity, a significant issue raised by the last network inference challenge (DREAM 5) is that no single network inference method performs optimally across all data sets. In contrast, integration of predictions from multiple inference methods through a consensus network shows robust and high performance across diverse data sets [32].

Apart from the algorithm used, once the whole network has been inferred, the classical approach to dealing with large amounts of interactions is identifying small-connected components as a means of testing their enrichment in specific biological processes.

1.3. The Missing Element: MicroRNAs (miRNAs)

Although highly innovative, the supervised and unsupervised approaches described so far do not take miRNAs into consideration. Many efforts have been made to predict miRNA/mRNA interactions, first by developing various target prediction algorithms and then by introducing new experimental techniques to isolate miRNA/mRNA complexes [33–36]. Computational target prediction is still widely used, although it is characterized by many false positives. For exhaustive reviews on miRNA discovery algorithms and *in silico* target prediction [37,38].

The integration of target predictions with miRNA and gene expression profiles has recently been proposed as a means of computationally improving and refining miRNA-target predictions. As miRNAs act predominantly through target degradation, the expression profiles of miRNAs and those of their target genes are expected to be inversely correlated [39,40].

Although the key role of miRNA in post-transcriptional regulation is universally recognized, few attempts have been made to use combinations of miRNA elements in developing gene set approaches. The only such attempt was described by Nam and colleagues [41], who performed GSEA on the mRNA targets of de-regulated miRNAs.

1.4. Case Study: The Pig as a Model Organism

Considering the advantages and disadvantages of the approaches described above, here we propose a consensus strategy based on the integration of pathway analysis, relevance networks and miRNA expression using as a model organism the pig and its cardiocirculatory system.

The size of organs, as well as various anatomical features, general physiology, and features of organ development, are very similar in pigs and humans. This permits the use of the pig as a model in the study of a number of pathologies, such as those affecting eyes [42], muscle [43], organ transplantation [44,45], and the gastrointestinal [46], nervous [47], and cardiovascular [48] systems. The coronary artery distribution in the pig is more similar to that of humans than is that of other animals. In addition, pigs present very similar cardiac output to humans; they possess a vaso vasorum in the aorta, and the left

azygous vein empties into the coronary sinus instead of into the precava. Blood pressure (145–160/105 BP), heart rate (100–150 BPM) and pulmonary pressure are higher in pigs than in humans.

Despite the medical importance of the pig as a species for study, our knowledge of the genome organization, gene expression regulation, and the molecular mechanisms underlying the pathophysiological processes of the pig is far less than the knowledge we have acquired of the mouse and rat. More than 90% of the porcine genome has been sequenced by the Swine Genome Sequencing Consortium [49]. The availability of detailed information on the porcine genome, together with emerging transgenic technologies, will enhance our ability to create specific and useful pig models. Recently, an atlas of DNA methylomes in porcine adipose and muscle tissues was published [50], and a great effort was made to combine genome sequence information with our knowledge of gene expression. Many of these studies focused on the swine immune system [51–54], while a genome-wide expression analysis in different tissues was described in Freeman's paper [55]. Recently, using sequencing approaches, a compendium of small non-coding RNAs was identified in various pig tissues (e.g., skeletal muscle [56–62], kidney [63], tooth [64], intestinal tract [65], brain [66], testis, ovary, sperm, and embryo [67–71] and pituitary gland [72]). Li and colleagues demonstrated that a complex regulatory network of porcine subcutaneous fat development is reflected in a great diversity of miRNA composition and expression between muscle and adipose tissue [73].

Here, we generate new custom mRNA and miRNA platforms that can be used to dissect the transcriptomic changes and regulatory circuits that are involved in the maintenance of veins and arteries in the pig. An integrative approach, combining pathway analysis and *de novo* network reconstruction, was used to expand our current knowledge of these regulatory circuits and to integrate miRNA activity into these circuits demonstrating their role in vessel specification. We show that vessel specificity can be maintained through different miRNAs (e.g., miR-133a and miR-143), the expression of which is inversely correlated with that of their mRNA targets.

2. Results and Discussion

The integration and analysis of gene and miRNA expression profiles across different tissues is fundamental to our understanding of tissue-specific processes. Here, we focus our analysis on differences in gene and miRNA expression among different tracts of the circulatory system: the two largest veins of the body (superior and inferior vena cava), the aorta (ascending and descending), the pulmonary artery, and the coronary artery. To achieve this goal, we created mRNA and miRNA [74] platforms, the latter based on the RAKE (RNA primed–array-based Klenow enzyme assay) method [75,76], to quantify coding and non-coding gene expression in pig tissues. After quantifying miRNA and mRNA expression, we used a combination of supervised and unsupervised approaches to detect transcriptional and post-transcriptional differences among different tracts of the circulatory system.

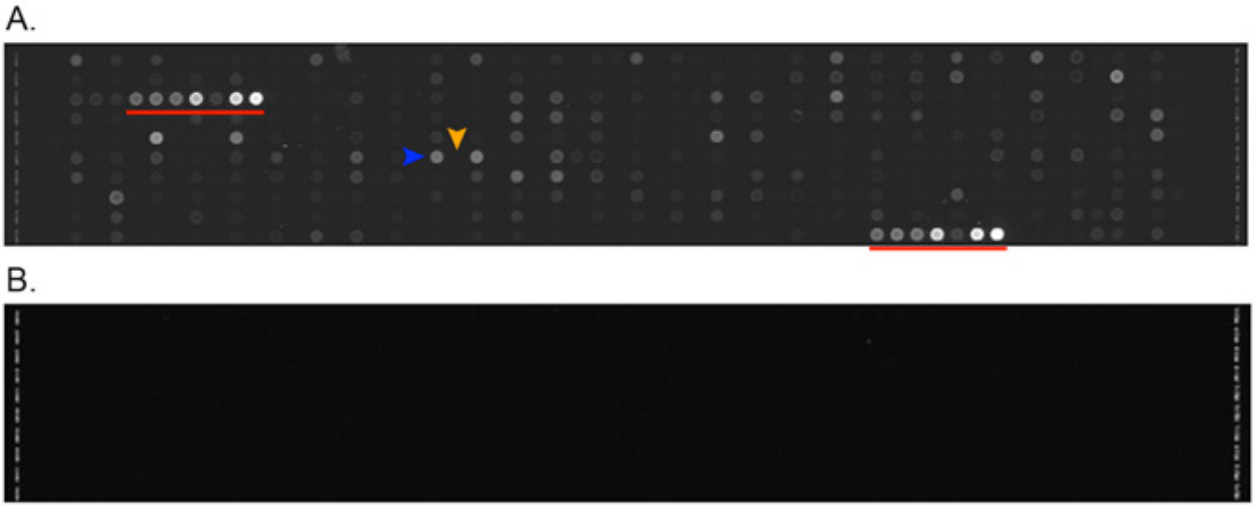
Ensembl transcripts (Ver. 56; EMBL-EBI, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, UK) and UniGene (Ver. 38; National Center for Biotechnology Information, U.S. National Library of Medicine, Bethesda, MD, USA) pig sequences were used to produce a dedicated microarray platform for monitoring mRNA expression. On the basis of sequence similarity, UniGene

features that overlapped more than 40% with an Ensembl transcript were discarded. After this filter, we obtained 40,267 UniGene clusters and 19,603 Ensembl transcripts (protein coding + pseudogenes + retrotransposed elements). For this selected collection of sequences, we designed microarray probes with different specificities and located at different distances from the 3' ends of specific transcripts using six different algorithms. The two best probes for each sequence, as determined by the reliability of the prediction algorithm and by the probe's vicinity to the 3'-end, were experimentally tested in a hybridization trial performed with a pool of mRNA populations independently prepared from 20 pig tissues (GEO: GSE28636). For each transcript with a replicated probe, we selected the probe that was the most responsive and specific on the basis of the intensity of fluorescence in the hybridization test, as suggested by Kronick [77]. The resulting pig whole-genome microarray, which was used in the gene expression analysis, is composed of: (i) 17,048 replicated probes and 963 single probes specific for the Ensembl transcripts; (ii) 11,363 replicated probes specific for the UniGene clusters of lengths between 778 nt and 1348 nt; and (iii) 28,790 single probes specific for the remaining UniGene clusters. Our analysis was not able to identify specific probes for 114 UniGene clusters and 1592 Ensembl transcripts. A limitation we faced in working with gene expression in pig was the poor gene annotation available. The number of annotated features on the array was increased by mining description and protein annotations to associate gene names with our probe symbols. Basically, for genes for which the HUGO (Human Genome Organisation) symbol was not present, we mined the description available from the Unigene database and retrieved additional gene or protein IDs, if present. All IDs were manually curated (ArrayExpress ID: A-MEXP-2351).

Recently, a new microarray platform based on 52,355 expressed sequences comprising miRNAs in miRBase Ver. 15 (Wellcome Trust Sanger Institute, Cambridge, UK) for pigs, cows, humans, and mice was described [55]. Unlike this new platform, which was constructed by spanning 22 probes along the transcripts, the platform we developed detects the 3'-UTR of each transcript; therefore, we are able to distinguish mRNA isoforms. This feature is fundamental because the activity of miRNAs is predominantly based on their interactions with the 3'-UTR region of mRNAs.

The identification of miRNAs was described in [74]. Briefly, bioinformatic analyses were performed on the pig genome for the identification of putative pre-miRNAs. These were experimentally tested using six independent RAKE experiments to identify 5' and 3' miRNA boundaries. After this experimental confirmation, all the pre-miRNAs identified as responsive (1235 hairpins) were tested for the presence of mature miRNA through RNA sequencing experiments. RNA sequencing experiments identified 343 hairpins coding for miRNAs. However, using PCR we were able to validate several miRNAs that were not confirmed by RNA sequencing. Therefore, we decided to produce an miRNA microarray platform (Array Express ID: A-MEXP-2348) containing all miRNAs detected by RAKE experiments. In the following analysis, we will discuss only miRNAs that were confirmed in sequencing experiments. Each specific probe is flanked by a background probe that was used to subtract the corresponding background fluorescence signal in the analysis (Figure 1).

Figure 1. Explicative scan portion of miRNA microarray after the RAKE and labeling reactions (**A**) and before hybridization (**B**). Spike-in spots are indicated by red lines; the blue arrow indicates a specific probe, and the orange arrow indicates its background probe. Each background probe was positioned to the right of its probe.



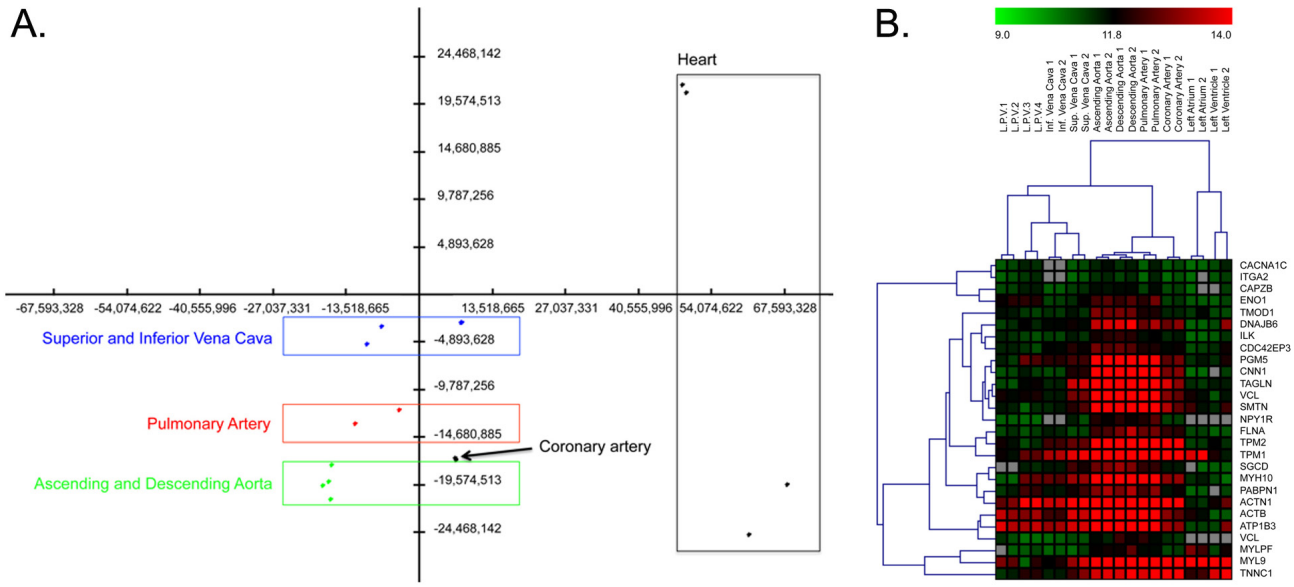
The short length of miRNAs makes complementary probe selection and the identification of optimized PCR primers a challenging task. While miRNA microarrays permit massive parallel and accurate relative measurement of all known miRNAs, they have been less useful for absolute quantification. We developed a new method that integrates the hybridization of miRNAs with an enzymatic elongation reaction that can take place only following a perfect match between the miRNA and the probe. Moreover, we introduced oligonucleotide spikes into the hybridization-enzymatic reaction, permitting the quantification of miRNAs over the linear dynamic range of 10^{-18} moles to 10^{-14} moles and avoiding biases related to sequence, labeling, or hybridization [74].

2.1. Differences between Arteries and Veins

We compared different tracts of the circulatory system: the two largest veins (the superior and inferior vena cava), the aorta (ascending and descending tracts), the pulmonary artery, and the coronary artery. As expected, the ascending and descending aorta and the coronary artery display similar gene expression profiles that are distinct from those of the superior and inferior vena cava (Figure 2A), while the pulmonary artery has an intermediate expression profile (Figure 2A). Arteries and veins are structurally different in terms of their relationship to the heart. Arteries receive blood directly from the heart and are therefore characterized by high pressure; in contrast, veins receive blood from peripheral body regions, and low pressure characterizes them. For this reason, some of the blood in the veins may not return to the heart but instead may back up or collect in these vessels. Veins transport de-oxygenated blood, while arteries transport oxygenated blood (with the exception of the pulmonary artery, which transports de-oxygenated blood to the lungs for oxygenation). The difference in blood pressure in arteries and veins is reflected in the different structures of these vessels. Arteries and arterioles have thicker walls than veins and venules; specifically, they possess an increased amount of smooth muscle that provides extra strength and elasticity to withstand surges of blood from the heart. Moreover, the thinner the vessel, the lower its innervation.

In accordance with the increased number of smooth muscle cells in arteries, the aorta expresses more smooth muscle-specific transcripts than the vena cava (Figure 2B). Genes that are up-regulated in the aorta include genes related to biological structures such as adherence junctions and processes such as nerve function and blood circulation (Table S1). This is consistent with the significantly higher level of innervation of arteries than of veins. Up-regulated genes in the vena cava are enriched in genes coding for proteins involved in the formation of the extracellular matrix (Table S1). These findings may be associated with the differences in elasticity between veins and arteries (veins have less elastic tissue than arteries).

Figure 2. (**A**) Principal component analysis (PCA). The first three components account for 62.8% of the observed variance. The green rectangle identifies the group of ascending and descending aorta samples (green dots); the coronary artery is indicated by a black dot, the red rectangle highlights pulmonary artery samples (red dots), and the blue rectangle surrounds superior and inferior vena cava samples (blue dots). On the right, separated from other samples, are heart samples; (**B**) Heat map of muscle transcripts. Transcripts coding for muscle proteins are up-regulated in arteries with respect to veins. The red squares indicate up-regulated genes, and the green squares indicate down-regulated genes. The grey squares indicate genes for which no expression was detected. L.P.V. = leaflet of pulmonary valve; Inf. Vena Cava = inferior vena cava; Sup. Vena Cava = superior vena cava. The numbers following the sample names indicate the number of experimental replicates.



A major component of the vessel walls of large arteries and veins is the extracellular matrix (ECM), which consists of collagens, elastin, and proteoglycans. The smooth muscle cells of the aorta and vena cava synthesize different amounts of collagen. As expected, our data show that collagen synthesis is four-fold higher in venous than in arterial [78]; collagen type I (COL1A2) is the most highly expressed extracellular matrix component.

Procollagen C-endopeptidase enhancer 2 (*PCOLCE2*) and *P4HA1 prolyl 4-hydroxylase, α polypeptide I* (*P4H4*) genes were found to be up-regulated in the vena cava. PCOLCE2 binds to the *C*-terminal propeptide of type I and II procollagens and may enhance the cleavage of their propeptides,

while P4H4 is a key enzyme in collagen synthesis. Moreover, we found *type VIII collagen* (*COL8A1*), which is typical of the endothelium lining vessels, and *type VI collagen* (*COL6A3*), a subendothelial constituent [79], to be highly expressed in the vena cava.

2.2. Pathway Analysis

Using multivariate pathway analysis methods such as GSEA, we overcame the major limitation of the classical enrichment approach, cut-off-based gene selection, focusing instead on coordinated changes in gene expression. Using this method, we were able to identify gene pathways that are specifically expressed in arteries and veins (Table 1). Among the activated pathways in arteries are those associated with smooth muscle contraction, calcium-calmodulin-dependent events, genome stability and regulation of intracellular signaling cascades. This finding is consistent with the presence of a thicker smooth muscle ring in arteries than in veins. Among the activated pathways in veins, we find the complement cascade, arachidonic acid metabolism, cell surface interactions at the vascular wall, and extracellular matrix metabolism (glycosaminoglycan metabolism and keratin/keratan sulphate metabolism). Arachidonic acid metabolism is involved in the control of various processes within the cardiocirculatory system, including vasoconstriction [80] and vasodilation [81,82]. The two most highly expressed genes related to arachidonic acid metabolism were *prostaglandin-endoperoxide synthase 2* (*PTGS2 or COX-2*) and *γ-glutamyltransferase 5* (*GGT5*). *COX-2* and *endothelial nitric oxide synthase* (*eNOS*) are primarily expressed in endothelial cells and are considered important regulators of vascular function. Under normal conditions, laminar flow induces COX-2 expression and synthesis of PGI_2, which in turn stimulates eNOS activity [83]. GGT expression was also localized in the endothelium [84]. As blood normally flows more slowly through veins than through arteries, thromboses are more common in veins than in arteries. This could be the reason for the control of vasodilation and vasoconstriction through metabolites of arachidonic acid.

In support of the up-regulation of elements of the complement cascade in veins, it is known that inflammation is more readily induced in venous than in arterial epithelium due to the conditions of the venous circulation. We checked for the presence of an inflammatory process by analyzing the expression of complement components in 19 tissues (Figure 3). We find that not all complement components are up-regulated in veins, while most are highly expressed in lymph nodes, spleen, and liver. This is in accordance with complement system synthesis and laundering. The complement system consists of a dozen circulating proteins, most of which are synthesized by the liver, that have the ability to bind to cellular membranes. The spleen and the liver are able to remove immune complexes composed of complement elements linked to erythrocyte membranes [85].

Finally, it is worth noting that pathways describing mucopolysaccharidosis syndromes such as Hurler, Sanfilippo, and Morquio syndromes were found to be significantly expressed in veins. Altered glycosaminoglycan metabolism is a key feature of these pathologies. Glycosaminoglycans are proteoglycans that bind to a varying degree water, electrolytes and macromolecules, such as collagen, within the connective tissue. The lining of veins and arteries comprises a substantial amount of the body's connective tissue. The outer layer of vessels (tunica adventitia) consists chiefly of connective tissue and is the thickest layer of the vein.

Table 1. Summary of Gene Set Enrichment Analysis (GSEA) analysis based on the Reactome database (http://www.reactome.org/ (accessed on 13 Novembre 2013)). Set size refers to the dimension of the pathway, and NTK (Normalized *T*-test of the kth gene set) is the observed value of the statistic as defined in the Graphite web tool [20]. Negative NTK values indicate pathways activated in veins, while positive values indicate pathways activated in arteries. It is worth noting that GSEA is known to have low statistical power; the suggested *Q*-value cut-off for identification of significant pathways is 0.25.

Pathway	Set size	NTk	*Q*-Value
Complement cascade	18	−5.29	0
Arachidonic acid metabolism	11	−3.09	0.044912281
Glycosaminoglycan metabolism	54	−3.09	0.044912281
MPS I—Hurler syndrome	54	−3.09	0.044912281
MPS II—Hunter syndrome	54	−3.09	0.044912281
MPS IIIA—Sanfilippo syndrome A	54	−3.09	0.044912281
MPS IIIB—Sanfilippo syndrome B	54	−3.09	0.044912281
MPS IIIC—Sanfilippo syndrome C	54	−3.09	0.044912281
MPS IIID—Sanfilippo syndrome D	54	−3.09	0.044912281
MPS IV—Morquio syndrome A	54	−3.09	0.044912281
MPS IV—Morquio syndrome B	54	−3.09	0.044912281
Biological oxidations	56	−2.75	0.106666667
Cell surface interactions at the vascular wall	54	−2.75	0.106666667
Keratan sulfate/keratin metabolism	20	−2.46	0.205977011
G α (12/13) signaling events	35	−2.37	0.24
Antigen presentation: Folding, assembly and peptide loading of class I MHC	11	−2.33	0.250980392
Golgi associated vesicle biogenesis	29	−2.29	0.247017544
Glutathione conjugation	10	−2.26	0.249756098
Phase II conjugation	23	−2.26	0.249756098
EGFR interacts with phospholipase C-γ	17	2.12	0.273710692
Ca-dependent events	14	2.14	0.262564103
Calmodulin induced events	14	2.14	0.262564103
CaM pathway	14	2.14	0.262564103
Cell-extracellular matrix interactions	15	2.2	0.254184397
PLCG1 events in ERBB2 signaling	18	2.23	0.252121212
DARPP-32 events	12	2.26	0.249756098
DAG and IP3 signaling	15	2.29	0.247017544
PLC-γ1 signaling	15	2.29	0.247017544
Amyloids	18	2.33	0.250980392
Telomere Maintenance	31	2.46	0.192688172
RNA polymerase I promoter opening	18	2.65	0.131282051
Chromosome maintenance	53	2.75	0.1024
Meiotic synapsis	24	2.88	0.077575758
Deposition of new CENPA-containing nucleosomes at the centromere	21	2.88	0.077575758
Nucleosome assembly	21	2.88	0.077575758
Packaging of telomere ends	12	3.09	0.044912281
Striated muscle contraction	21	4.76	0
Smooth muscle contraction	19	6.13	0
Muscle contraction	36	7.25	0

Figure 3. Expression of genes involved in the complement response. The numbers represent gene expression levels normalized to the average expression of the same gene across all tissues. Down-regulated genes are shown in green, and up-regulated genes are shown in red. Most of the up-regulated genes are expressed in the liver, which is responsible for the synthesis of most of the proteins of the complement system, in the spleen and in lymph nodes (lymphoid organs). NA = Expression not detected; L.P.V. = leaflet pulmonary valve; WBC.A = white blood cells from arterial blood; WBC.V = white blood cells from venous blood.

	Ascending Aorta	Descending Aorta	Conoray Artery	Pulmonary Artery	Inf. Vena Cava	Sup. Vena Cava	Left Atrium	Left Ventr	L.P.V. (animal 1)	L.P.V. (animal 2)	Liver	WBC.A	WBC.V	Lymph node	Spleen	Tongue	Skeletal Muscle	Lung	Kidney	Stomach	Adipose tissue	Skin
C1QA	0.36	0.48	0.45	0.62	0.61	1.06	0.49	0.64	0.82	0.73	0.52	0.39	0.44	7.08	3.31	0.72	0.37	0.55	0.45	0.65	0.58	0.70
C1QB	0.40	0.70	0.42	1.11	0.76	1.84	0.36	0.40	1.45	1.83	0.40	NA	0.27	4.92	2.95	0.47	0.26	0.67	0.29	0.43	0.62	0.44
C1QBP	0.68	0.85	0.93	1.03	1.09	0.88	0.88	2.47	0.98	0.67	0.88	0.69	1.08	0.79	0.89	0.93	0.75	0.75	1.21	0.93	1.75	0.87
C1QC	0.43	0.52	0.72	0.74	0.81	1.21	0.51	0.45	1.81	1.09	0.58	0.41	0.45	5.68	2.51	0.69	0.40	0.63	0.44	0.65	0.63	0.64
C1R	0.67	0.81	0.76	0.98	0.72	1.01	0.88	0.95	0.69	0.71	2.24	1.17	0.81	0.83	2.28	1.86	0.57	0.64	0.65	0.98	0.73	1.03
C1S	0.84	1.01	1.07	1.03	1.50	1.50	0.56	1.15	0.59	0.83	1.07	0.44	1.01	0.94	0.67	1.80	0.45	1.10	0.92	1.01	1.01	1.48
C3	0.50	0.67	0.70	0.69	1.05	1.11	0.66	0.70	1.92	0.90	3.83	0.51	0.50	2.64	0.69	0.65	0.48	1.12	0.50	0.70	0.88	0.61
C3AR1	0.83	1.02	1.12	0.95	1.19	0.95	1.08	0.90	1.14	0.97	0.99	1.07	0.94	0.96	0.83	0.84	1.07	1.23	0.92	1.03	0.89	1.10
C3P1	0.78	0.79	0.96	0.98	0.96	0.92	0.99	1.47	0.98	0.71	1.41	0.87	1.02	0.87	1.15	1.26	0.98	0.85	0.91	1.08	1.04	1.01
C4	0.86	0.92	1.17	0.99	1.11	0.73	0.58	0.98	1.04	0.90	1.29	1.17	1.23	1.68	1.05	0.70	0.60	0.88	0.97	1.02	1.45	0.70
C4BPA	0.38	0.43	0.97	0.71	2.35	1.62	0.53	0.45	1.54	1.77	0.51	0.27	0.31	3.16	1.81	0.47	0.30	0.70	0.33	0.34	1.76	1.29
C5	0.84	1.01	0.85	0.82	0.86	0.88	0.91	1.08	0.85	0.95	2.58	0.86	0.86	0.74	1.01	0.78	1.00	1.16	1.08	1.07	0.90	0.89
C5AR1	0.85	1.00	1.09	0.89	1.07	0.90	1.07	0.98	1.00	1.02	1.00	1.08	0.96	0.94	1.03	0.83	1.14	1.16	0.99	1.06	0.94	1.00
C6	0.92	1.01	0.94	1.12	1.03	0.97	1.01	0.88	0.89	1.16	1.09	0.97	0.99	1.03	0.97	NA	1.08	1.02	1.04	NA	0.93	0.94
C7	0.70	0.67	1.75	0.88	0.98	0.80	0.48	1.00	1.63	1.31	0.94	0.60	1.31	0.66	1.38	0.54	0.39	1.49	1.83	0.55	1.33	0.76
C8A	0.86	0.83	0.77	0.81	NA	0.91	NA	0.82	0.81	0.92	2.85	NA	0.79	NA	0.88	0.89	NA	NA	0.97	NA	0.90	NA
C8B	0.99	0.98	1.02	1.02	0.80	1.15	NA	0.97	1.05	1.14	1.21	0.93	0.96	1.14	1.03	1.04	0.79	0.96	1.15	0.85	1.05	0.78
C8G	0.89	0.84	1.08	0.96	1.03	0.95	1.05	1.07	1.08	0.69	1.92	1.01	0.90	1.03	0.87	1.01	0.86	0.87	0.88	0.96	0.93	1.11
C9	0.62	0.68	0.70	0.70	0.77	0.69	0.75	0.71	0.70	0.74	5.45	0.70	0.76	0.75	0.72	0.70	1.13	1.32	0.93	0.72	1.00	0.74
CD46	0.77	0.92	0.83	0.84	0.79	NA	0.97	0.75	NA	0.89	0.83	0.96	0.84	0.90	3.05	NA	1.08	0.85	0.78	0.96	NA	NA
CD55	1.21	1.03	0.99	1.06	1.00	1.26	NA	0.89	1.06	1.06	0.87	1.01	0.96	1.00	0.99	0.92	NA	0.99	0.98	0.97	0.91	0.83
CD59	1.06	1.56	1.20	1.67	2.19	1.84	0.46	0.84	0.52	1.16	0.28	0.28	0.35	0.60	0.32	0.87	0.39	0.45	0.71	1.97	2.78	0.50
CFAB	0.72	0.68	0.58	0.63	0.66	0.69	0.53	0.70	0.60	0.86	6.73	0.57	0.58	0.71	0.76	0.67	0.59	1.67	0.77	NA	0.72	0.59
CFD	0.91	NA	0.90	1.06	0.88	0.91	1.12	1.05	0.93	NA	1.04	1.06	0.88	1.27	1.00	1.00	0.87	0.91	0.91	1.39	0.86	1.03
CFH	0.70	0.91	1.09	0.95	0.84	0.92	0.72	1.10	1.05	1.06	1.07	0.68	1.11	1.01	0.93	0.83	NA	1.40	1.63	0.68	1.30	1.01
CFI	0.89	0.94	0.85	0.94	0.92	1.29	NA	0.78	0.85	1.92	2.27	0.76	0.67	0.94	1.01	0.70	0.65	1.12	1.20	0.94	0.80	0.59
CFP	1.15	0.69	2.25	1.16	1.90	0.63	0.30	0.67	2.81	0.83	0.33	0.73	0.85	1.24	0.89	0.34	0.35	0.96	1.76	0.51	1.05	0.60
CR1L	0.98	0.96	0.85	0.92	0.97	0.93	1.11	1.04	0.95	0.92	1.41	0.93	0.91	0.96	0.88	1.14	1.33	0.97	0.83	1.04	0.89	1.08
CR2	1.06	1.03	0.91	0.96	0.98	0.98	0.91	0.98	1.07	1.07	0.92	1.01	0.99	1.03	0.98	0.92	1.08	0.99	1.05	1.04	1.05	0.99
DF	0.21	0.37	1.23	0.83	1.94	1.56	0.33	0.42	0.57	0.82	0.24	0.95	0.71	4.58	0.49	0.60	0.33	0.63	0.22	0.24	3.78	0.94
ERCC2	NA	1.03	0.96	0.96	0.98	NA	0.99	0.90	1.34	0.97	0.95	0.97	1.00	0.83	0.88	NA	1.16	0.99	0.96	1.18	0.92	1.01
ITGB2	0.45	0.43	0.31	0.50	0.47	0.73	0.34	0.34	1.53	0.69	0.44	5.18	1.52	3.03	2.14	0.45	0.33	1.56	0.31	0.41	0.47	0.38

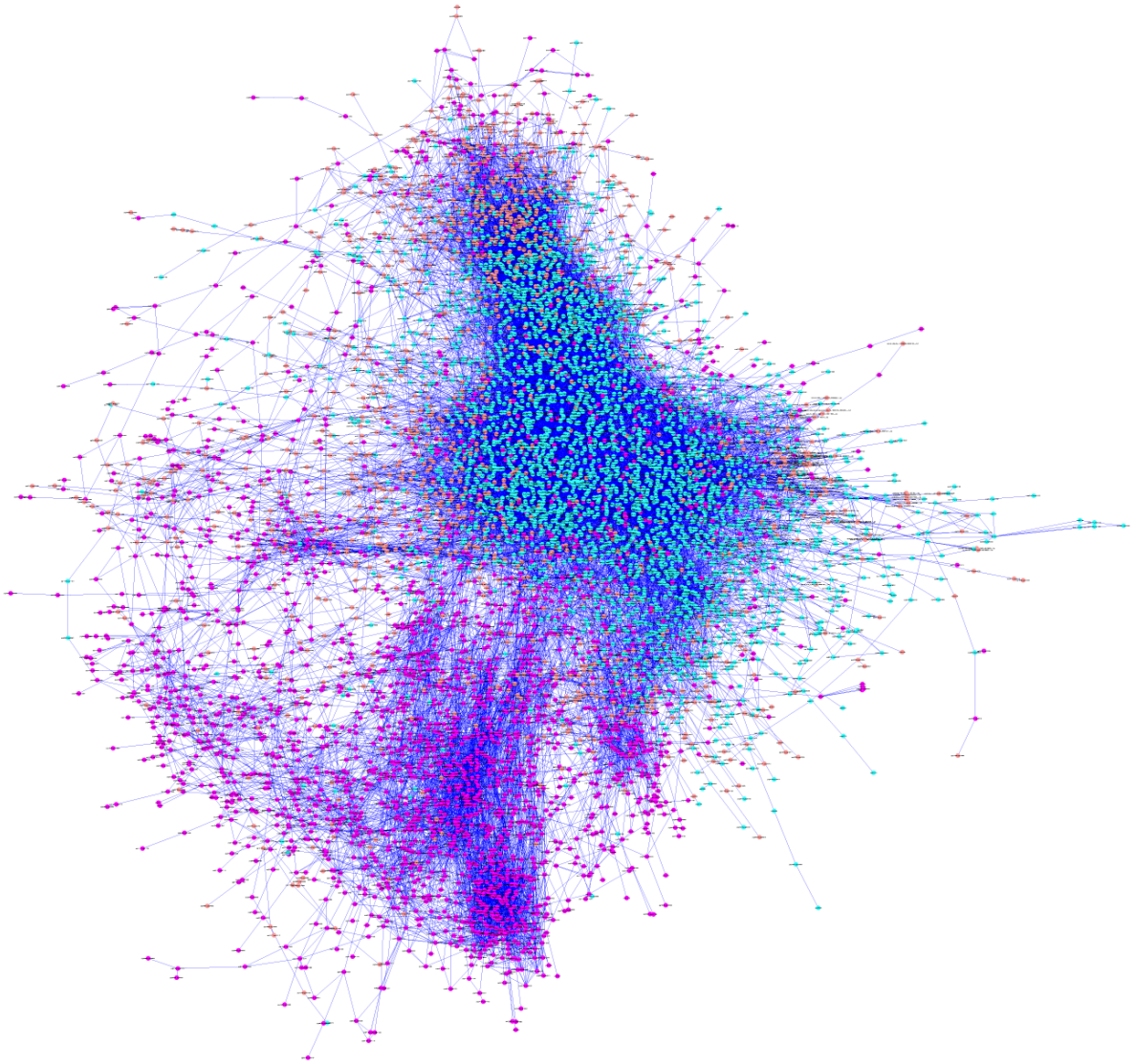
2.3. De Novo *Pathway Reconstruction: Topological Parameters*

Pathway analysis fails to consider many known genes and miRNAs that are not annotated in any pathway. To fill these gaps, we used *de novo* network reconstruction using both mRNA and miRNA profiles. Using a correlation measure with a permutation-based threshold of 0.9 of mutual information (0.9 was the maximum value of mutual information of the network generated by the permuted expression matrix), we generated a network with 7762 nodes (7647 genes and 115 miRNAs) and 44,092 edges (Figure 4). The global architecture of the network is characterized by two large clusters,

which are shown as the blue and violet nodes in Figure 4. As expected (Figure 2A), these two clusters are composed of genes prevalently expressed in heart (the most different tissue) and in blood vessels (Figure S1). Thus, we separated these two clusters to create a vessel-specific and a heart-specific network.

To gain insight into the structure of complex networks of this type, various topological parameters were calculated (Table 2). The heart network is sparser and less connected than the vessel network. This is reflected by a larger number of connected components, a higher diameter and a smaller number of neighborhood genes of the heart network.

Figure 4. Regulatory network reconstructed using mutual information. The edges of the network are colored according to their prevalent expression. Heart-specific genes are shown in violet, vessel-specific genes are shown in blue, and genes without tissue-specific expression are shown in pink.



The degree of a node, also referred to as its connectivity, is the number of edges connected to the node. Based on this definition, the nodes with the highest connectivities are called hubs. In general, hub genes are master regulators and play important roles in the biology of the cell. In our networks, we

define as hubs the top 5% of genes in the connectivity distribution. We found 162 and 128 hubs in the vessel and heart networks, respectively. The hub genes of the vessel network encode proteins that participate in two main processes: RNA processing and the regulation of apoptotic events (Table S2). During normal development as well as in pathology, the formation of new vessels and the regression of pre-existing ones depend on the balance between endothelial cell proliferation and endothelial cell apoptosis. In mature vessels, endothelial cell turnover is also under the control of these tightly regulated phenomena. Among the hubs of the heart network, we identified genes involved in cell membrane structure and signal transduction through MAPK activity as well as genes encoding various ion transporters (e.g., Na^{2+}, K^{+}) (Table S2). The members of the MAPK family are involved in the regulation of many cellular processes, including cell growth, differentiation, development, the cell cycle, death, and survival. Activation of genes in the MAPK family plays a key role in the pathogenesis of various processes in the heart, including myocardial hypertrophy and its transition to heart failure, ischemic and reperfusion injury, and cardioprotection conferred by ischemia- or drug-induced preconditioning [86].

Table 2. Summary of the principal topological parameters estimated for the *de novo* reconstructed network.

Topological parameters	Heart network	Vessels network
Average clustering coefficient	0.195	0.234
Connected components	237	86
Avg. number of neighbors	6.329	15.611
Network radius	1	1
Network diameter	36	16
Network centralization	0.020	0.036
Network density	0.002	0.005
Network heterogeneity	1.198	1.183

The *de novo* reconstructed network (Figure 4) is characterized by the presence of different miRNAs (Table S3) that are responsible for the regulation of vessel specificity. Figure 5 represents the sub-network of the neighboring genes of miRNAs. Interestingly, the central part of the network (the densely connected portion of the sub-network) is characterized by genes involved in smooth muscle contraction (Table S4) that show differential expression in arteries and veins (Figure 6). As discussed previously, a thicker ring of smooth muscle is present in arteries than in veins (see Section 2.2). Our results suggest that this difference may be regulated by specific miRNAs that display anti-correlated expression with their putative targets (Figure 6).

Specifically, the *α 2-actin* (ACTA2) smooth muscle gene in aorta (ENSSSCG00000010447) is regulated by a specific miRNA (prediction_15_14390446_14390503_-_3p) that is down-regulated in the aorta and up-regulated in venous tissue (Figure 6). Defects in ACTA2 are the cause of aortic aneurysm familial thoracic type 6 (AAT6) [MIM:611788]. AATs are characterized by permanent dilation of the thoracic aorta, usually due to degenerative changes in the aortic wall. RHOB (Ssc#S35170885), an important gene involved in vasoconstriction, is also regulated by miR-133a (Figure 6). RHO gene family is involved in vascular morphogenesis [87], and miR-133a contributes to the phenotypic state of smooth muscle cells both *in vitro* and *in vivo*, suggesting a potential for

therapeutic application of this miRNA in vascular disease [88]. In fact, miR-133a, in association with miR143/145, is fundamental for the maintenance of the contractile smooth muscle cell phenotype [88]. The expression of miRNAs prediction_15_14390446_14390503_-_3p and miR-133 and their targets ACAT2 and RHOB was confirmed by qRT-PCR (Figure 6C).

Figure 5. Gene and miRNA interaction sub-network describing vessel specificity. Triangles represent miRNAs; circles represent mRNAs. Gene expression in the ascending aorta according to $\log_2$ (gene expression/average gene expression) is represented by color; green indicates down-regulation, red indicates up-regulation. Under each node, histograms representing $\log_2$ (gene expression/average gene expression) in the ascending aorta, descending aorta, inferior vena cava, and superior vena cava (reading from left to right) are shown. The area highlighted by the circle indicates the densely connected portion of the sub-network (an enlarged view of this area is available in Figure 6).

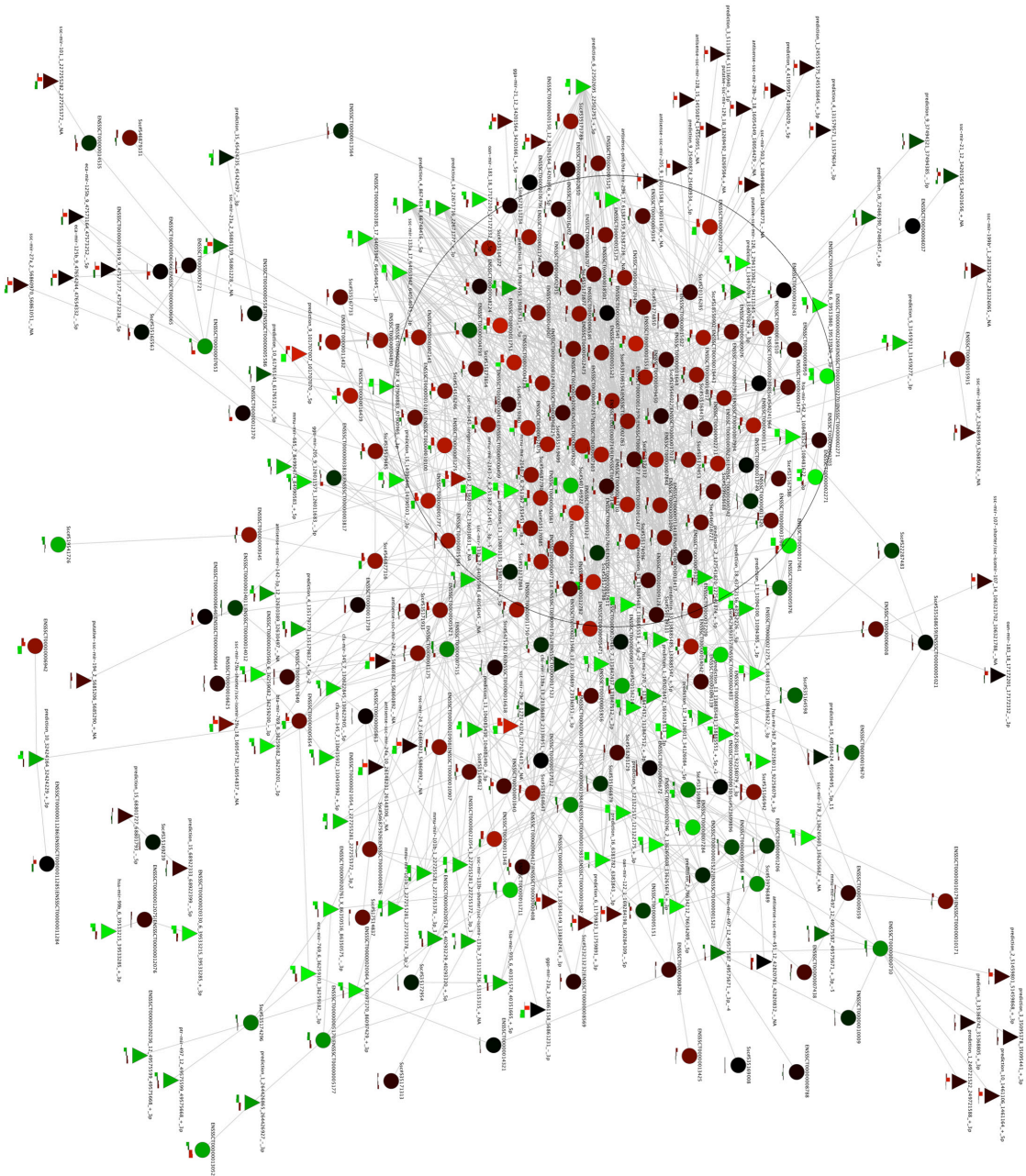
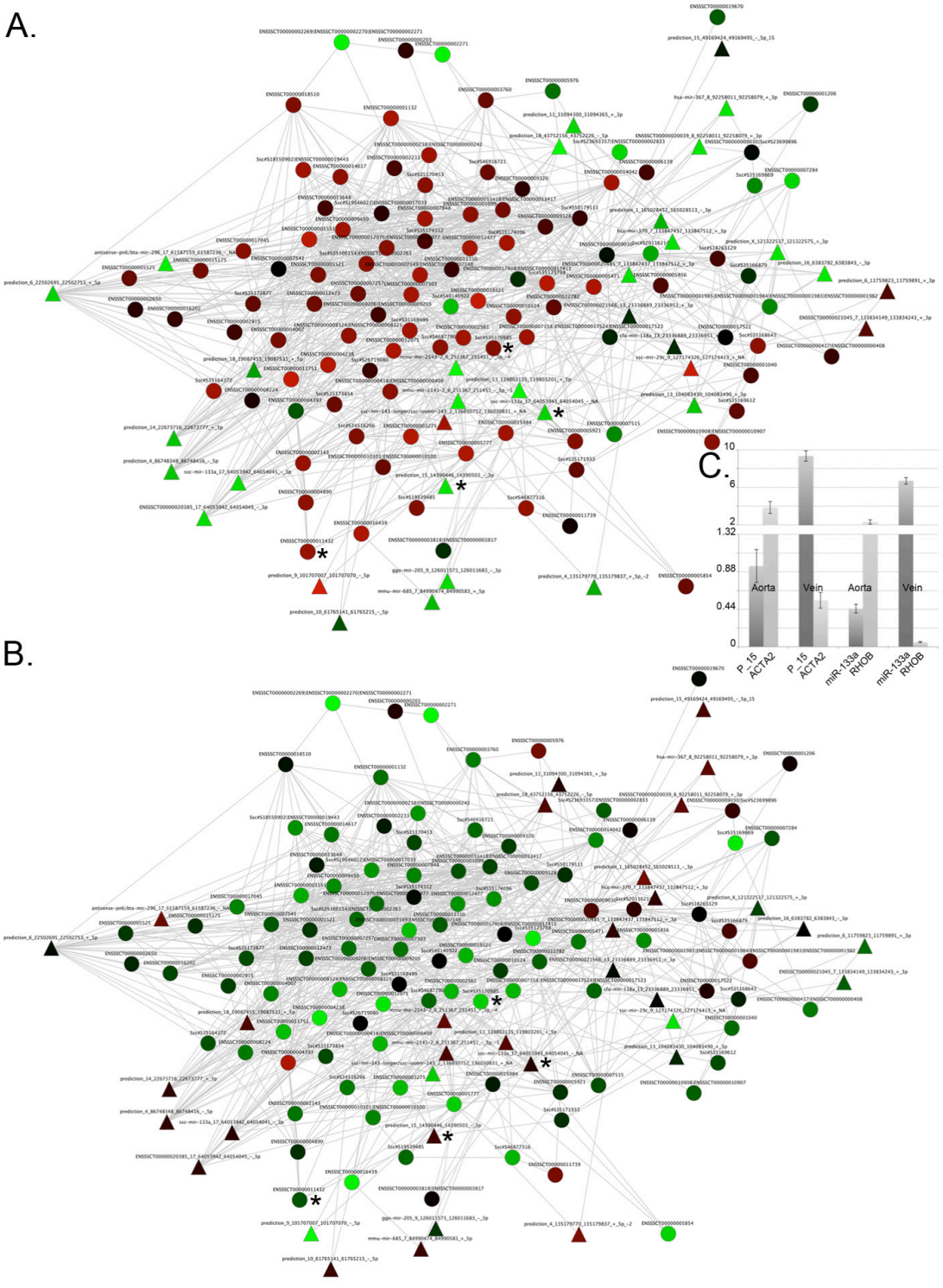


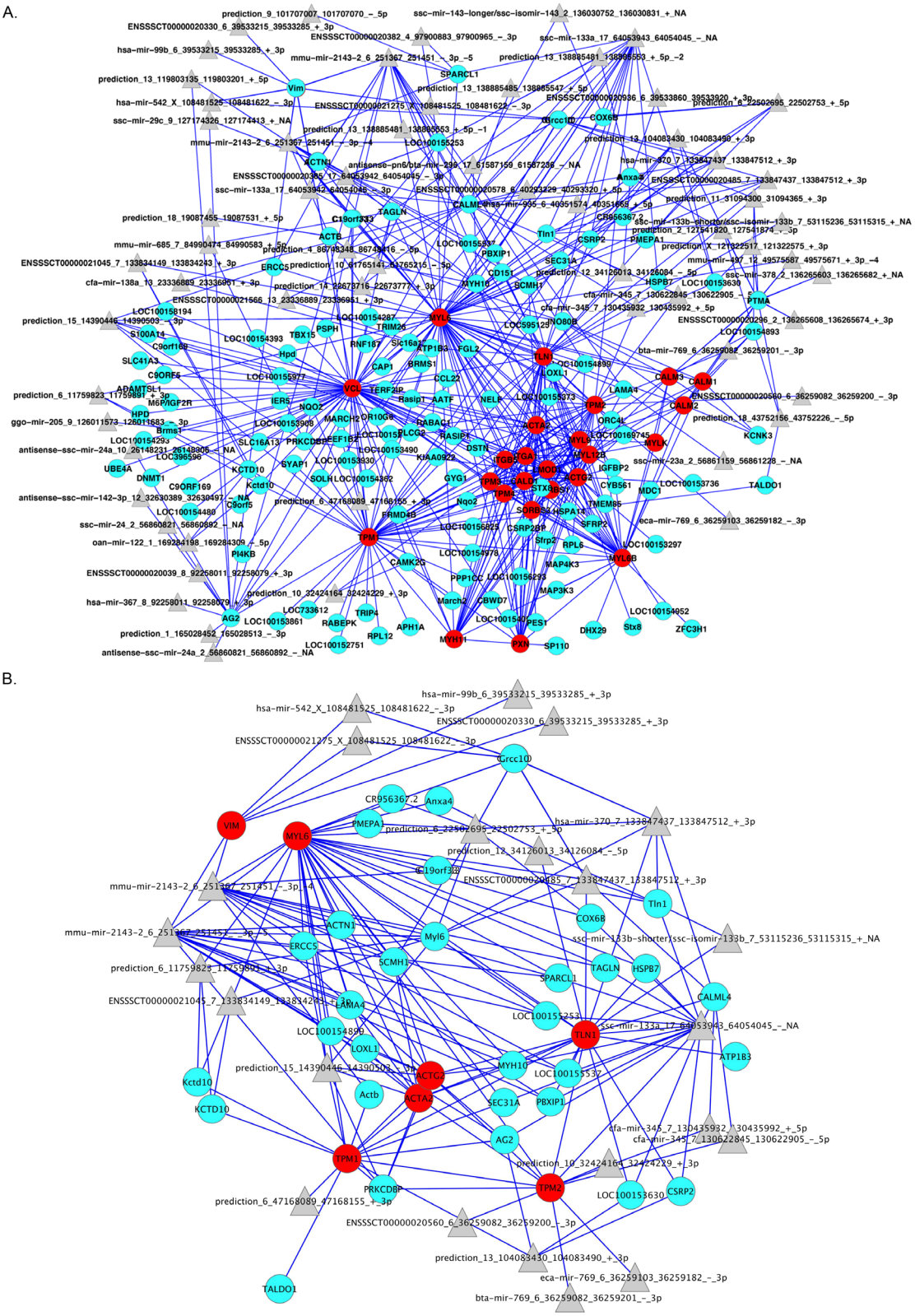
Figure 6. Enlarged view of the densely connected area of Figure 5. (**A**) The colors indicate expression in the aorta; (**B**) The colors indicate expression in veins. The triangles represent miRNAs; circles represent mRNAs. Up-regulated = red; down-regulated = green; * = nodes discussed in the text; (**C**) qRT-PCR results confirm that there is an inverse relationship between miRNAs and their targets. P_15 is for prediction_15_14390446_14390503_-_3p. In *Y* axis the original expression level related to H3. Bars are for standard deviation between three replicates.



2.4. Integration of Supervised and Unsupervised Approaches

Supervised and unsupervised approaches gave similar results in terms of biological processes involved in tissue specificity. However, their complementary behavior might be better exploited through the use of an integrative approach. Specifically, our aim is to combine the topology of the discovered pathways with that of the *de novo* reconstructed network. The advantage of combining the topologies obtained in sections 2.2 and 2.3 is two-fold: (i) it allows the expansion of pathway definitions to include genes currently without pathway annotation; and (ii) it permits the inclusion of miRNAs. Using the topological structure of the pathway as a backbone, we include new genes in the pathway, following two rules: (i) a gene/miRNA is added only if it presents an edge in the *de novo* network with at least one gene in the pathway; and (ii) additional miRNAs are included if they share an edge with previously added non-annotated genes. Here, we will use this strategy to discuss one of the most interesting pathways significantly activated in arteries: the smooth muscle contraction pathway (Figure 7A). The genes used to expand this pathway (the γ isoform of the catalytic subunit of *protein phosphatase 1* (*PPP1CC*), *transgelin* (*TAGLN*), and smooth muscle and non-muscle *myosin light chain 6* (*Myl6*), among others) are primarily involved in membrane and actin filament organization, actomyosin function and responses to specific stimuli (NF-κB binding and response to unfolded protein) (Table S5), reflecting their functional congruence with the smooth muscle contraction pathway. Indeed, the membrane organization category includes the organismation of the sarcoplasmic reticulum, which is involved in the regulation of intracellular Ca^{2+} concentration (Figure 7). All of these genes are prevalently expressed in smooth muscle; in particular, TAGLN was purified from bovine aorta [89]. Moreover, we added 61 miRNAs that putatively regulate genes involved, directly or indirectly in the smooth muscle contraction pathway (Figure 7A). Interestingly, 23 miRNAs are involved in the regulation of the original genes of the pathway (core genes). Among these miRNAs, miR-542 (ENSSSCT00000021275), which was shown in a previous work to be involved in the epithelial-mesenchymal transition [90], was found to be associated with *vimentin* (*VIM*) regulation (Figure 7B). Finally, it is worth noting that many other miRNAs important for vascular remodeling and smooth muscle phenotypic control, such as miR-133 [88], miR-143 [91], miR-99b [92], miR-23a [93], miR-138 (ENSSSCT00000021566) [94], miR-29c [95], miR-125a (ENSSSCT00000020936) [95], and miR-24 [96]), are included in this network.

Figure 7. (**A**) Combination of pathway topology and *ab initio* reconstructed network. Nodes corresponding to the Reactome pathway (core nodes) are shown in red; additional genes in the first neighborhood of the core nodes obtained from the *ab initio* network are shown in light blue, and miRNAs are shown in grey; (**B**) Portion of (**A**) representing the miRNAs regulating the core nodes.



3. Experimental Section

3.1. Sample Preparation

RNA samples (total RNA and small RNAs) were extracted from the analyzed tissues of three non-inbred pigs and kept at −80 °C until use. Before the experiments were performed, the three samples from the same tissues were pooled, and miRNA was selected using a flashPAGE instrument (Ambion, Carlsbad, CA, USA). RNA extraction was performed using TRIzol (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. The PureLink Isolation Kit (Invitrogen, Carlsbad, CA, USA) was used to separate long RNA from short (<200 nt, after use in the flashPAGE instrument). All samples were quantitated using a NanoDrop ND-1000 spectrophotometer; RNA quality was then analyzed using the Agilent Bioanalyser 2100 (Agilent, Santa Clara, CA, USA) (Agilent RNA 6000 nano kit; RIN at least 7 accepted) and for the presence of miRNA using the Agilent small RNA kit.

3.2. Microarray Platforms

For this study, we synthesized two different types of microarray platforms: (a) 4 × 2 K Combimatrix microarrays for miRNA expression profiling (ArrayExpress ID: A-MEXP-2348); (b) 90 K Combimatrix microarrays (ArrayExpress ID: A-MEXP-2351) for mRNA expression profiling. All microarrays were synthesized using the Combimatrix oligonucleotide synthesizer station (Combimatrix, Mukilteo, WA, USA), which allows *in situ* synthesis of oligonucleotide probes through phosphoramidite chemistry. All synthesized microarray platforms were tested for uniformity of the probes as suggested by the manufacturer.

The 4 × 2 K microarrays contain specific probes for miRNAs. Each specific probe is flanked by a background probe that is used in the analysis to subtract the corresponding background fluorescence signal (Figure 1). The background probes were derived from a previous RAKE experiment aimed at the identification of specific ends of miRNAs in which a tiling microarray was used for the scope (Figure S2) [74].

3.3. Microarray mRNA and miRNA Gene Expression and qRT-PCR

3.3.1. mRNA

Pooled RNA (1 μg; three samples from the same tissue) was linearly amplified and labeled by the addition of biotinylated nucleotides according to the procedure described in the Ambion MessageAmp™ II aRNA Amplification kit (Ambion, Carlsbad, CA, USA). The procedure includes reverse transcription with an oligo-dT primer carrying a T7 promoter to produce the first-strand cDNA. After second-strand synthesis and clean-up, the cDNA is used as template in an *in vitro* transcription reaction to generate a large quantity of antisense RNA (aRNA). Biotinylated UTPs were incorporated into the aRNA during the *in vitro* transcription reaction. Following purification, 18 μg of aRNA was fragmented using the Ambion Fragmentation Kit (Ambion, Carlsbad, CA, USA). Intact and fragmented aRNAs were tested on an Agilent Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA) using the RNA 6000 Nano LabChip (Agilent, Santa Clara, CA, USA). The size of intact aRNAs

ranged from 300 to 4000 nucleotides, while that of fragmented aRNAs ranged from 50 to 250 nucleotides. Fragmented aRNA was hybridized to pre-hybridized 90 K Combimatrix microarrays. The pre-hybridization step was performed for 2 h at 42 °C in a solution containing 5× Denhardt's solution, 100 ng/µL salmon sperm DNA and 0.05% SDS in 1× Hyb solution prepared as suggested by Combimatrix. Hybridizations were carried out with 4.8 µg of fragmented aRNA in 25% DI formamide, 100 ng/µL salmon sperm DNA and 0.04% SDS in 1× Hybridization solution at 42 °C for 18 h with constant mixing. After hybridization, the microarray platforms were washed with the following:

- 6× SSPET (SSPE added with 0.05% of Tween-20) preheated at 42 °C for 5 min;
- 3× SSPET for 1 min at room temperature;
- 0.5× SSPET for 1 min at room temperature; and
- PBST for 1 min at room temperature.

The microarray chamber was then filled with biotin blocking solution (0.1% Tween-20 and 10 mg/mL BSA in 2× PBS) and incubated at room temperature for 1 h. Labeling was performed by incubating the microarray with dye labeling solution (0.1% Tween-20, 10 mg/mL BSA and 1.6 ng of Cy3-streptavidin (Amersham, Little Chalfont, UK) in 2× PBS) for 1 h at room temperature. After the washing steps (PBST for 1 min at room temperature two times; PBS for 1 min at room temperature), microarrays were scanned at 3 µm resolution with the VersArray ChiprRaderTM (BioRad, Hercules, CA, USA) (ArrayExpress ID: E-MTAB-1941).

3.3.2. miRNA

A sample of the miRNA pool (350 ng) was hybridized for 20 h at 37 °C in a static hybridization oven in hybridization buffer consisting of 6× SSPE, 8 mg/mL BSA, 700 ng of small RNAs and spike-in. After hybridization, the microarrays were washed with the following stringent procedure:

- 1 min at room temperature with 6× SSPET (SSPE containing 0.05% Tween-20);
- 1 min at room temperature with 3× SSPET;
- 1 min at room temperature with 2× PBS;
- 1 min at room temperature with 1× Buffer 2 (the buffer for the Klenow enzyme).

The RAKE reaction was performed at 36.5 °C by incubating the microarray for 1.5 h in 1× Buffer 2 containing 16 µM biotin-14-dATP (Invitrogen, Carlsbad, CA, USA) and 0.25 U/µL Klenow fragment (3'→5' exo−) (NEB, Ipswich, MA, USA). The microarrays were washed two times in 1× Buffer 2 and incubated in biotin blocking solution for 1 h at room temperature. Extended miRNAs (primers) were labeled by incubating the microarray in the dye labeling solution for 1 h at room temperature. The microarrays were rinsed in PBST (0.1% Tween-20 in 2× PBS) for 1 min at room temperature and in 2× PBS for 1 min at room temperature and scanned (ArrayExpress ID: E-MTAB-1938).

qRT-PCR was used to validate the expression of miRNAs and mRNAs. For mRNA, the SYBR green approach was used in association with the *Power* SYBR® Green PCR Master Mix (Applied Biosystems, Carlsbad, CA, USA); for miRNA, the NCode™ SYBR® Green miRNA qRT-PCR Kit (Life Technologies, Carlsbad, CA, USA) was used according to the manufacturer's specifications. The primers used were GCATGCAGAAGGAGATCACA (left) and GCTGGAAGGTGGACAGA

GAG (right) for ACTA2, TATGTGCTTCTCGGTGGACA (left) and CGAGGTAGTCGTA GGCTTGG (right) for RHOB, and GGTTCCCAGGCTAGGGGTCG (specific) for prediction_15_14390446_14390503_-_3p and CAGCTGGTTGAAGGGGACCA for miR-133a. The reference genes used were GAPDH for mRNA and snU6 for miRNA. The results shown are normalized to the expression of histone H3.

3.4. Data Analysis

Images of hybridized mRNA microarrays were quantitated using the Combimatrix imaging software. The raw data were normalized using the quantile method. The goal of the quantile method is to normalize the distribution of probe intensities across a set of microarrays. After normalization, the fluorescence intensities of probe spots presenting values lower than the average of the medians of all negative control probes were set as missing values (NA). The negative control probes were used to calculate the background value (filter). Probe spots presenting NA in more than six experiments were excluded from data analysis. Before performing the analysis, the intensity values of the replicated probes were averaged. Differentially expressed genes were identified using the MeV suite [97] and applying PCA (Principal Components Analysis) [98] and SAM (Significance Analysis of Microarrays) [2] analysis. COA (Correspondence Analysis) analysis [99] was used to determine the specificity of the *de novo* reconstructed network. Gene enrichment was performed using the DAVID web application [100]; pathway analysis was performed using GraphiteWeb [20].

miRNA data were pre-processed as previously described except that cyclic lowess normalization was applied [101]. After inter-array normalization, the fluorescence intensity of the specific miRNA probe was subtracted from the corresponding background fluorescence and used to extrapolate the miRNA concentration from the spike-in-derived curve. The spike-in curve was extrapolated using spline interpolation [102].

Pig gene symbols from Ensembl were converted to human gene symbols using the Ensembl orthologous database through the BioMart service. For UniGene clusters, we extracted the most similar protein or gene curated by NCBI (http://www.ncbi.nlm.nih.gov/ (accessed on 13 November 2013)) based on sequence similarity and then used the NCBI HomoloGene database to translate the protein or gene to its human homolog. This method is commonly used to map genes to pathways in non-model organisms or to map genes that are poorly annotated in model organisms [103]; it is also common to use the well-curated human pathways to extrapolate pathways for non-model organisms. GSEA [10] was then performed using the GraphiteWeb web tool [20].

Mutual information (MI) between all pairs of genes and miRNAs was estimated using the parmigene Bioconductor package [31]; miRNA-miRNA interactions have been removed. To assess MI significance, we estimated the null distribution using a permutational approach. The expression profiles of miRNAs and mRNAs were randomly shuffled, and MI was then estimated on the shuffled matrices. To generate the global network, we included only interactions with MI that were greater than the maximum MI value obtained from the null distribution, which was 0.9 (corresponding to quantile 0.999 in the empirical distribution).

The Cytoscape tool [104] with the Networkanalyser [105] plugin was used to estimate the topological properties of heart and vessel networks.

The topologies of the most interesting pathways derived from pathway annotation (graphite Bioconductor package) were integrated with the topology of the *de novo* reconstructed network. The combination was performed using the pathway topology as backbone; new genes/miRNAs were then added based on fulfilment of one of the following criteria: (i) if the new gene/miRNA shares an edge in the *de novo* network with at least one gene in the pathway; and (ii) if an miRNA shares an edge in the *de novo* network with at least one previously added gene.

4. Conclusions

Gene set analyses have been shown to provide better insights and more robust results in array experiments than classical gene-by-gene approaches. Here, we reviewed various strategies used in gene set analysis and showed how to address their integration. We combined genome and pathway information with expression data and applied this approach to a case study, the analysis of the pig cardiocirculatory system. Two new platforms for pig transcriptome analysis (mRNA and miRNA) were presented and applied to the study of tissue specificity. Different expression patterns were identified in heart and vessels; within these, arteries show distinct profiles from those of veins. These findings seem to be associated with the functional and structural composition of the vessels. In agreement with histochemical evidence, pathway analysis revealed the greater importance of smooth muscle in arteries than in veins. We showed that miRNAs participate in the definition of arterial and venous pathways; specifically, for smooth muscle, our data indicate the importance of miR-133a in regulating the *RHOB* gene. The use of a combination of supervised and unsupervised approaches allowed us to expand the compositions of known pathways to include new genes involved in membrane and actin filament organization, actomyosin function and response to stimuli and new miRNAs, most of which are known to be associated with vascular remodeling and control of the smooth muscle phenotype. These results demonstrate the feasibility and usefulness of combining these two approaches in identifying new candidate genes whose expression is associated with specific experimental conditions.

Acknowledgments

The authors acknowledge the CARIPARO Foundation (Project for Excellence 2012: “Role of coding and non-coding RNA in chronic myeloproliferative neoplasms: from bioinformatics to translational research”) and the CRIBI Center for high-performance computing resources funded by the Regione Veneto (RISIB project SMUPR n. 4145). The authors wish to thank the University of Padova for support of this work (CPDR075919 and CPDA119031 to Chiara Romualdi; CPDR070805 to Gabriele Sales) and MicroCribi service for microarray synthesis support (http://microcribi.cribi.unipd.it).

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Smyth, G.K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **2004**, *3*, 1–28.
2. Tusher, V.G.; Tibshirani, R.; Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 5116–5121.
3. Shen, K.; Tseng, G.C. Meta-analysis for pathway enrichment analysis when combining multiple genomic studies. *Bioinformatics* **2010**, *26*, 1316–1323.
4. Callegaro, A.; Basso, D.; Bicciato, S. A locally adaptive statistical procedure (LAP) to identify differentially expressed chromosomal regions. *Bioinformatics* **2006**, *22*, 2658–2666.
5. Toedling, J.; Schmeier, S.; Heinig, M.; Georgi, B.; Roepcke, S. MACAT—Microarray chromosome analysis tool. *Bioinformatics* **2005**, *21*, 2112–2113.
6. Turkheimer, F.E.; Roncaroli, F.; Hennuy, B.; Herens, C.; Nguyen, M.; Martin, D.; Evrard, A.; Bours, V.; Boniver, J.; Deprez, M. Chromosomal patterns of gene expression from microarray data: Methodology, validation and clinical relevance in gliomas. *BMC Bioinform.* **2006**, *7*, 526.
7. Barry, W.T.; Nobel, A.B.; Wright, F.A. Significance analysis of functional categories in gene expression studies: A structured permutation approach. *Bioinformatics* **2005**, *21*, 1943–1949.
8. Goeman, J.J.; van de Geer, S.A.; de Kort, F.; van Houwelingen, H.C. A global test for groups of genes: Testing association with a clinical outcome. *Bioinformatics* **2004**, *20*, 93–99.
9. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550.
10. Tian, L.; Greenberg, S.A.; Kong, S.W.; Altschuler, J.; Kohane, I.S.; Park, P.J. Discovering statistically significant pathways in expression profiling studies. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13544–13549.
11. Levin, A.M.; Ghosh, D.; Cho, K.R.; Kardia, S.L. A model-based scan statistic for identifying extreme chromosomal regions of gene expression in human tumors. *Bioinformatics* **2005**, *21*, 2867–2874.
12. Efron, B.; Tibshirani, R. On testing the significance of sets of genes. *Ann. Appl. Stat.* **2007**, *1*, 107–129.
13. Tarca, A.L.; Draghici, S.; Khatri, P.; Hassan, S.S.; Mittal, P.; Kim, J.S.; Kim, C.J.; Kusanovic, J.P.; Romero, R. A novel signaling pathway impact analysis. *Bioinformatics* **2009**, *25*, 75–82.
14. Martini, P.; Sales, G.; Massa, M.S.; Chiogna, M.; Romualdi, C. Along signal paths: An empirical gene set approach exploiting pathway topology. *Nucleic Acids Res.* **2013**, *41*, e19.
15. Ackermann, M.; Strimmer, K. A general modular framework for gene set enrichment analysis. *BMC Bioinform.* **2009**, *10*, 47.
16. Nam, D.; Kim, S.Y. Gene-set approach for expression pattern analysis. *Brief. Bioinform.* **2008**, *9*, 189–197.
17. Sales, G.; Calura, E.; Cavalieri, D.; Romualdi, C. Graphite—A Bioconductor package to convert pathway topology to gene network. *BMC Bioinform.* **2012**, *13*, 20.

18. Martini, P.; Risso, D.; Sales, G.; Romualdi, C.; Lanfranchi, G.; Cagnin, S. Statistical Test of Expression Pattern (STEPath): A new strategy to integrate gene expression data with genomic information in individual and meta-analysis studies. *BMC Bioinform.* **2011**, *12*, 92.
19. Massa, M.S.; Chiogna, M.; Romualdi, C. Gene set analysis exploiting the topology of a pathway. *BMC Syst. Biol.* **2010**, *4*, 121.
20. Sales, G.; Calura, E.; Martini, P.; Romualdi, C. Graphite Web: Web tool for gene set analysis exploiting pathway topology. *Nucleic Acids Res.* **2013**, *41*, W89–W97.
21. Morgat, A.; Coissac, E.; Coudert, E.; Axelsen, K.B.; Keller, G.; Bairoch, A.; Bridge, A.; Bougueleret, L.; Xenarios, I.; Viari, A. UniPathway: A resource for the exploration and annotation of metabolic pathways. *Nucleic Acids Res.* **2012**, *40*, D761–D769.
22. Kanehisa, M.; Goto, S.; Furumichi, M.; Tanabe, M.; Hirakawa, M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* **2010**, *38*, D355–D360.
23. Caspi, R.; Foerster, H.; Fulcher, C.A.; Kaipa, P.; Krummenacker, M.; Latendresse, M.; Paley, S.; Rhee, S.Y.; Shearer, A.G.; Tissier, C.; *et al.* The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res.* **2008**, *36*, D623–D631.
24. Joshi-Tope, G.; Gillespie, M.; Vastrik, I.; D'Eustachio, P.; Schmidt, E.; de Bono, B.; Jassal, B.; Gopinath, G.R.; Wu, G.R.; Matthews, L.; *et al.* Reactome: A knowledgebase of biological pathways. *Nucleic Acids Res.* **2005**, *33*, D428–D432.
25. Friedman, N. Inferring cellular networks using probabilistic graphical models. *Science* **2004**, *303*, 799–805.
26. Butte, A.J.; Tamayo, P.; Slonim, D.; Golub, T.R.; Kohane, I.S. Discovering functional relationships between RNA expression and chemotherapeutic susceptibility using relevance networks. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 12182–12186.
27. Schafer, J.; Strimmer, K. An empirical Bayes approach to inferring large-scale gene association networks. *Bioinformatics* **2005**, *21*, 754–764.
28. Markowetz, F.; Spang, R. Inferring cellular networks—A review. *BMC Bioinform.* **2007**, *8*, S5.
29. Basso, K.; Margolin, A.A.; Stolovitzky, G.; Klein, U.; Dalla-Favera, R.; Califano, A. Reverse engineering of regulatory networks in human B cells. *Nat. Genet.* **2005**, *37*, 382–390.
30. Margolin, A.A.; Nemenman, I.; Basso, K.; Wiggins, C.; Stolovitzky, G.; Dalla Favera, R.; Califano, A. ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinform.* **2006**, *7*, S7.
31. Sales, G.; Romualdi, C. Parmigene—A parallel R package for mutual information estimation and gene network reconstruction. *Bioinformatics* **2011**, *27*, 1876–1877.
32. Marbach, D.; Costello, J.C.; Kuffner, R.; Vega, N.M.; Prill, R.J.; Camacho, D.M.; Allison, K.R.; Consortium, D.; Kellis, M.; Collins, J.J.; *et al.* Wisdom of crowds for robust gene network inference. *Nat. Methods* **2012**, *9*, 796–804.
33. Macias, S.; Plass, M.; Stajuda, A.; Michlewski, G.; Eyras, E.; Caceres, J.F. DGCR8 HITS-CLIP reveals novel functions for the Microprocessor. *Nat. Struct. Mol. Biol.* **2012**, *19*, 760–766.
34. Thomson, D.W.; Bracken, C.P.; Goodall, G.J. Experimental strategies for microRNA target identification. *Nucleic Acids Res.* **2011**, *39*, 6845–6853.

35. Hafner, M.; Landthaler, M.; Burger, L.; Khorshid, M.; Hausser, J.; Berninger, P.; Rothballer, A.; Ascano, M., Jr.; Jungkamp, A.C.; Munschauer, M.; *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **2010**, *141*, 129–141.
36. Chi, S.W.; Zang, J.B.; Mele, A.; Darnell, R.B. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* **2009**, *460*, 479–486.
37. Yousef, M.; Showe, L.; Showe, M. A study of microRNAs *in silico* and *in vivo*: Bioinformatics approaches to microRNA discovery and target identification. *FEBS J.* **2009**, *276*, 2150–2156.
38. Witkos, T.M.; Koscianska, E.; Krzyzosiak, W.J. Practical aspects of microRNA target prediction. *Curr. Mol. Med.* **2011**, *11*, 93–109.
39. Sales, G.; Coppe, A.; Bisognin, A.; Biasiolo, M.; Bortoluzzi, S.; Romualdi, C. MAGIA, a web-based tool for miRNA and Genes Integrated Analysis. *Nucleic Acids Res.* **2010**, *38*, W352–W359.
40. Bisognin, A.; Sales, G.; Coppe, A.; Bortoluzzi, S.; Romualdi, C. MAGIA^2: From miRNA and genes expression data integrative analysis to microRNA-transcription factor mixed regulatory circuits (2012 update). *Nucleic Acids Res.* **2012**, *40*, W13–W21.
41. Nam, S.; Li, M.; Choi, K.; Balch, C.; Kim, S.; Nephew, K.P. MicroRNA and mRNA integrated analysis (MMIA): A web tool for examining biological functions of microRNA expression. *Nucleic Acids Res.* **2009**, *37*, W356–W362.
42. Ross, J.W.; Fernandez de Castro, J.P.; Zhao, J.; Samuel, M.; Walters, E.; Rios, C.; Bray-Ward, P.; Jones, B.W.; Marc, R.E.; Wang, W.; *et al.* Generation of an inbred miniature pig model of retinitis pigmentosa. *Investig. Ophthalmol. Vis. Sci.* **2012**, *53*, 501–507.
43. Maxmen, A. Model pigs face messy path. *Nature* **2012**, *486*, 453.
44. Sandrin, M.S.; Loveland, B.E.; McKenzie, I.F. Genetic engineering for xenotransplantation. *J. Card. Surg.* **2001**, *16*, 448–457.
45. Ekser, B.; Rigotti, P.; Gridelli, B.; Cooper, D.K. Xenotransplantation of solid organs in the pig-to-primate model. *Transpl. Immunol.* **2009**, *21*, 87–92.
46. Zhang, Q.; Widmer, G.; Tzipori, S. A pig model of the human gastrointestinal tract. *Gut Microbes* **2013**, *4*, 193–200.
47. Kragh, P.M.; Nielsen, A.L.; Li, J.; Du, Y.; Lin, L.; Schmidt, M.; Bogh, I.B.; Holm, I.E.; Jakobsen, J.E.; Johansen, M.G.; *et al.* Hemizygous minipigs produced by random gene insertion and handmade cloning express the Alzheimer's disease-causing dominant mutation APPsw. *Transgenic Res.* **2009**, *18*, 545–558.
48. Granada, J.F.; Kaluza, G.L.; Wilensky, R.L.; Biedermann, B.C.; Schwartz, R.S.; Falk, E. Porcine models of coronary atherosclerosis and vulnerable plaque for imaging and interventional research. *EuroIntervention* **2009**, *5*, 140–148.
49. Groenen, M.A.; Archibald, A.L.; Uenishi, H.; Tuggle, C.K.; Takeuchi, Y.; Rothschild, M.F.; Rogel-Gaillard, C.; Park, C.; Milan, D.; Megens, H.J.; *et al.* Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* **2012**, *491*, 393–398.
50. Li, M.; Wu, H.; Luo, Z.; Xia, Y.; Guan, J.; Wang, T.; Gu, Y.; Chen, L.; Zhang, K.; Ma, J.; *et al.* An atlas of DNA methylomes in porcine adipose and muscle tissues. *Nat. Commun.* **2012**, *3*, 850.
51. Fairbairn, L.; Kapetanovic, R.; Beraldi, D.; Sester, D.P.; Tuggle, C.K.; Archibald, A.L.; Hume, D.A. Comparative analysis of monocyte subsets in the pig. *J. Immunol.* **2013**, *190*, 6389–6396.

52. Martins, R.P.; Lorenzi, V.; Arce, C.; Lucena, C.; Carvajal, A.; Garrido, J.J. Innate and adaptive immune mechanisms are effectively induced in ileal Peyer's patches of Salmonella typhimurium infected pigs. *Dev. Comp. Immunol.* **2013**, *41*, 100–104.
53. Hulst, M.; Smits, M.; Vastenhouw, S.; de Wit, A.; Niewold, T.; van der Meulen, J. Transcription networks responsible for early regulation of Salmonella-induced inflammation in the jejunum of pigs. *J. Inflamm.* **2013**, *10*, 18.
54. Adler, M.; Murani, E.; Brunner, R.; Ponsuksili, S.; Wimmers, K. Transcriptomic response of porcine PBMCs to vaccination with tetanus toxoid as a model antigen. *PLoS One* **2013**, *8*, e58306.
55. Freeman, T.C.; Ivens, A.; Baillie, J.K.; Beraldi, D.; Barnett, M.W.; Dorward, D.; Downing, A.; Fairbairn, L.; Kapetanovic, R.; Raza, S.; *et al.* A gene expression atlas of the domestic pig. *BMC Biol.* **2012**, *10*, 90.
56. McDaneld, T.G.; Smith, T.P.; Harhay, G.P.; Wiedmann, R.T. Next-generation sequencing of the porcine skeletal muscle transcriptome for computational prediction of microRNA gene targets. *PLoS One* **2012**, *7*, e42039.
57. Zhou, B.; Liu, H.L.; Shi, F.X.; Wang, J.Y. MicroRNA expression profiles of porcine skeletal muscle. *Anim. Genet.* **2010**, *41*, 499–508.
58. Liu, Y.; Li, M.; Ma, J.; Zhang, J.; Zhou, C.; Wang, T.; Gao, X.; Li, X. Identification of differences in microRNA transcriptomes between porcine oxidative and glycolytic skeletal muscles. *BMC Mol. Biol.* **2013**, *14*, 7.
59. Siengdee, P.; Trakooljul, N.; Murani, E.; Schwerin, M.; Wimmers, K.; Ponsuksili, S. Transcriptional profiling and miRNA-dependent regulatory network analysis of longissimus dorsi muscle during prenatal and adult stages in two distinct pig breeds. *Anim. Genet.* **2013**, *44*, 398–407.
60. McDaneld, T.G.; Smith, T.P.; Doumit, M.E.; Miles, J.R.; Coutinho, L.L.; Sonstegard, T.S.; Matukumalli, L.K.; Nonneman, D.J.; Wiedmann, R.T. MicroRNA transcriptome profiles during swine skeletal muscle development. *BMC Genomics* **2009**, *10*, 77.
61. Huang, T.H.; Zhu, M.J.; Li, X.Y.; Zhao, S.H. Discovery of porcine microRNAs and profiling from skeletal muscle tissues during development. *PLoS One* **2008**, *3*, e3225.
62. Shen, H.; Liu, T.; Fu, L.; Zhao, S.; Fan, B.; Cao, J.; Li, X. Identification of microRNAs involved in dexamethasone-induced muscle atrophy. *Mol. Cell. Biochem.* **2013**, *381*, 105–113.
63. Timoneda, O.; Balcells, I.; Nunez, J.I.; Egea, R.; Vera, G.; Castello, A.; Tomas, A.; Sanchez, A. miRNA expression profile analysis in kidney of different porcine breeds. *PLoS One* **2013**, *8*, e55402.
64. Li, A.; Song, T.; Wang, F.; Liu, D.; Fan, Z.; Zhang, C.; He, J.; Wang, S. MicroRNAome and expression profile of developing tooth germ in miniature pigs. *PLoS One* **2012**, *7*, e52256.
65. Sharbati, S.; Friedlander, M.R.; Sharbati, J.; Hoeke, L.; Chen, W.; Keller, A.; Stahler, P.F.; Rajewsky, N.; Einspanier, R. Deciphering the porcine intestinal microRNA transcriptome. *BMC Genomics* **2010**, *11*, 275.
66. Podolska, A.; Kaczkowski, B.; Kamp Busk, P.; Sokilde, R.; Litman, T.; Fredholm, M.; Cirera, S. MicroRNA expression profiling of the porcine developing brain. *PLoS One* **2011**, *6*, e14494.

67. Zhou, Y.; Tang, X.; Song, Q.; Ji, Y.; Wang, H.; Jiao, H.; Ouyang, H.; Pang, D. Identification and characterization of pig embryo microRNAs by Solexa sequencing. *Reprod. Domest. Anim.* **2013**, *48*, 112–120.
68. Lian, C.; Sun, B.; Niu, S.; Yang, R.; Liu, B.; Lu, C.; Meng, J.; Qiu, Z.; Zhang, L.; Zhao, Z. A comparative profile of the microRNA transcriptome in immature and mature porcine testes using Solexa deep sequencing. *FEBS J.* **2012**, *279*, 964–975.
69. Li, M.; Liu, Y.; Wang, T.; Guan, J.; Luo, Z.; Chen, H.; Wang, X.; Chen, L.; Ma, J.; Mu, Z.; *et al.* Repertoire of porcine microRNAs in adult ovary and testis by deep sequencing. *Int. J. Biol. Sci.* **2011**, *7*, 1045–1055.
70. Curry, E.; Safranski, T.J.; Pratt, S.L. Differential expression of porcine sperm microRNAs and their association with sperm morphology and motility. *Theriogenology* **2011**, *76*, 1532–1539.
71. Luo, L.; Ye, L.; Liu, G.; Shao, G.; Zheng, R.; Ren, Z.; Zuo, B.; Xu, D.; Lei, M.; Jiang, S.; *et al.* Microarray-based approach identifies differentially expressed microRNAs in porcine sexually immature and mature testes. *PLoS One* **2010**, *5*, e11744.
72. Li, H.; Xi, Q.; Xiong, Y.; Cheng, X.; Qi, Q.; Yang, L.; Shu, G.; Wang, S.; Wang, L.; Gao, P.; *et al.* A comprehensive expression profile of microRNAs in porcine pituitary. *PLoS One* **2011**, *6*, e24883.
73. Li, H.Y.; Xi, Q.Y.; Xiong, Y.Y.; Liu, X.L.; Cheng, X.; Shu, G.; Wang, S.B.; Wang, L.N.; Gao, P.; Zhu, X.T.; *et al.* Identification and comparison of microRNAs from skeletal muscle and adipose tissues from two porcine breeds. *Anim. Genet.* **2012**, *43*, 704–713.
74. Martini, P.; Sales, G.; Brugiolo, M.; Gandaglia, A.; Naso, F.; De Pitta', C.; Spina, M.; Gerosa, G.; Romualdi, C.; Cagnin, S.; *et al.* Tissue-specific expression and regulatory networks of pig microRNAome. *PLoS One* **2013**, unpublished work.
75. Nelson, P.T.; Baldwin, D.A.; Kloosterman, W.P.; Kauppinen, S.; Plasterk, R.H.; Mourelatos, Z. RAKE and LNA-ISH reveal microRNA expression and localization in archival human brain. *RNA* **2006**, *12*, 187–191.
76. Nelson, P.T.; Baldwin, D.A.; Scearce, L.M.; Oberholtzer, J.C.; Tobias, J.W.; Mourelatos, Z. Microarray-based, high-throughput gene expression profiling of microRNAs. *Nat. Methods* **2004**, *1*, 155–161.
77. Kronick, M.N. Creation of the whole human genome microarray. *Expert Rev. Proteomics* **2004**, *1*, 19–28.
78. Wong, A.P.; Nili, N.; Strauss, B.H. *In vitro* differences between venous and arterial-derived smooth muscle cells: Potential modulatory role of decorin. *Cardiovasc. Res.* **2005**, *65*, 702–710.
79. Ross, J.M.; McIntire, L.V.; Moake, J.L.; Rand, J.H. Platelet adhesion and aggregation on human type VI collagen surfaces under physiological flow conditions. *Blood* **1995**, *85*, 1826–1835.
80. Smedegard, G.; Hedqvist, P.; Dahlen, S.E.; Revenas, B.; Hammarstrom, S.; Samuelsson, B. Leukotriene C4 affects pulmonary and cardiovascular dynamics in monkey. *Nature* **1982**, *295*, 327–329.
81. Pawloski, J.R.; Chapnick, B.M. Antagonism of LTD4-evoked relaxation in canine renal artery and vein. *Am. J. Physiol.* **1993**, *265*, H980–H985.

82. Brink, C.; Dahlen, S.E.; Drazen, J.; Evans, J.F.; Hay, D.W.; Nicosia, S.; Serhan, C.N.; Shimizu, T.; Yokomizo, T. International Union of Pharmacology XXXVII. Nomenclature for leukotriene and lipoxin receptors. *Pharmacol. Rev.* **2003**, *55*, 195–227.
83. Inoue, H.; Taba, Y.; Miwa, Y.; Yokota, C.; Miyagi, M.; Sasaguri, T. Transcriptional and posttranscriptional regulation of cyclooxygenase-2 expression by fluid shear stress in vascular endothelial cells. *Arterioscler. Thromb. Vasc. Biol.* **2002**, *22*, 1415–1420.
84. Dahboul, F.; Leroy, P.; Maguin Gate, K.; Boudier, A.; Gaucher, C.; Liminana, P.; Lartaud, I.; Pompella, A.; Perrin-Sarrado, C. Endothelial γ-glutamyltransferase contributes to the vasorelaxant effect of *S*-nitrosoglutathione in rat aorta. *PLoS One* **2012**, *7*, e43190.
85. Yousaf, N.; Howard, J.C.; Williams, B.D. Studies in the rat of antibody-coated and *N*-ethylmaleimide-treated erythrocyte clearance by the spleen. I. Effects of *in vivo* complement activation. *Immunology* **1986**, *59*, 75–79.
86. Ravingerova, T.; Barancik, M.; Strniskova, M. Mitogen-activated protein kinases: A new therapeutic target in cardiac pathology. *Mol. Cell. Biochem.* **2003**, *247*, 127–138.
87. Howe, G.A.; Addison, C.L. RhoB controls endothelial cell morphogenesis in part via negative regulation of RhoA. *Vasc. Cell* **2012**, *4*, 1.
88. Torella, D.; Iaconetti, C.; Catalucci, D.; Ellison, G.M.; Leone, A.; Waring, C.D.; Bochicchio, A.; Vicinanza, C.; Aquila, I.; Curcio, A.; *et al.* MicroRNA-133 controls vascular smooth muscle cell phenotypic switch *in vitro* and vascular remodeling *in vivo*. *Circ. Res.* **2011**, *109*, 880–893.
89. Kobayashi, R.; Kubota, T.; Hidaka, H. Purification, characterization, and partial sequence analysis of a new 25-kDa actin-binding protein from bovine aorta: A SM22 homolog. *Biochem. Biophys. Res. Commun.* **1994**, *198*, 1275–1280.
90. Rhodes, L.V.; Tilghman, S.L.; Boue, S.M.; Wang, S.; Khalili, H.; Muir, S.E.; Bratton, M.R.; Zhang, Q.; Wang, G.; Burow, M.E.; *et al.* Glyceollins as novel targeted therapeutic for the treatment of triple-negative breast cancer. *Oncol. Lett.* **2012**, *3*, 163–171.
91. Cordes, K.R.; Sheehy, N.T.; White, M.P.; Berry, E.C.; Morton, S.U.; Muth, A.N.; Lee, T.H.; Miano, J.M.; Ivey, K.N.; Srivastava, D. miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature* **2009**, *460*, 705–710.
92. Ikeda, S.; Pu, W.T. Expression and function of microRNAs in heart disease. *Curr. Drug Targets* **2010**, *11*, 913–925.
93. Chhabra, R.; Dubey, R.; Saini, N. Cooperative and individualistic functions of the microRNAs in the miR-23a~27a~24-2 cluster and its implication in human diseases. *Mol. Cancer* **2010**, *9*, 232.
94. Li, S.; Ran, Y.; Zhang, D.; Chen, J.; Zhu, D. MicroRNA-138 plays a role in hypoxic pulmonary vascular remodelling by targeting Mst1. *Biochem. J.* **2013**, *452*, 281–291.
95. Park, C.; Yan, W.; Ward, S.M.; Hwang, S.J.; Wu, Q.; Hatton, W.J.; Park, J.K.; Sanders, K.M.; Ro, S. MicroRNAs dynamically remodel gastrointestinal smooth muscle cells. *PLoS One* **2011**, *6*, e18628.
96. Talasila, A.; Yu, H.; Ackers-Johnson, M.; Bot, M.; van Berkel, T.; Bennett, M.; Bot, I.; Sinha, S. Myocardin regulates vascular response to injury through miR-24/-29a and platelet-derived growth factor recepto-β. *Arterioscler. Thromb. Vasc. Biol.* **2013**, *33*, 2355–2365.

97. Saeed, A.I.; Sharov, V.; White, J.; Li, J.; Liang, W.; Bhagabati, N.; Braisted, J.; Klapa, M.; Currier, T.; Thiagarajan, M.; *et al.* TM4: A free, open-source system for microarray data management and analysis. *Biotechniques* **2003**, *34*, 374–378.
98. Raychaudhuri, S.; Stuart, J.M.; Altman, R.B. Principal components analysis to summarize microarray experiments: Application to sporulation time series. *Pac. Symp. Biocomput.* **2000**, *5*, 455–466.
99. Fellenberg, K.; Hauser, N.C.; Brors, B.; Neutzner, A.; Hoheisel, J.D.; Vingron, M. Correspondence analysis applied to microarray data. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10781–10786.
100. Huang da, W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4*, 44–57.
101. Risso, D.; Massa, M.S.; Chiogna, M.; Romualdi, C. A modified LOESS normalization applied to microRNA arrays: A comparative evaluation. *Bioinformatics* **2009**, *25*, 2685–2691.
102. Spath, H. *Two Dimensional Spline Interpolation Algorithms*; A K Peters/CRC Press: Wellesley, MA, USA, 1995.
103. Kanehisa, M.; Goto, S.; Kawashima, S.; Okuno, Y.; Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **2004**, *32*, D277–D280.
104. Saito, R.; Smoot, M.E.; Ono, K.; Ruscheinski, J.; Wang, P.L.; Lotia, S.; Pico, A.R.; Bader, G.D.; Ideker, T. A travel guide to Cytoscape plugins. *Nat. Methods* **2012**, *9*, 1069–1076.
105. Doncheva, N.T.; Assenov, Y.; Domingues, F.S.; Albrecht, M. Topological analysis and interactive visualization of biological networks and protein structures. *Nat. Protoc.* **2012**, *7*, 670–685.

Reprinted from *IJMS*. Cite as: Iaconetti, C.; Gareri, C.; Polimeni, A.; Indolfi, C. Non-Coding RNAs: The "Dark Matter" of Cardiovascular Pathophysiology. *Int. J. Mol. Sci.* **2013**, *14*, 19987-20018.

Review

Non-Coding RNAs: The "Dark Matter" of Cardiovascular Pathophysiology

Claudio Iaconetti, Clarice Gareri, Alberto Polimeni and Ciro Indolfi *

Division of Cardiology, Magna Graecia University, URT Consiglio Nazionale delle Ricerche (CNR), Catanzaro 88100, Italy; E-Mails: iaconetticlaudio@unicz.it (C.I.); c.gareri@unicz.it (C.G.); polimeni@unicz.it (A.P.)

* Author to whom correspondence should be addressed; E-Mail: indolfi@unicz.it; Tel.: +39-961-3694231; Fax: +39-961-6394073.

Received: 4 July 2013; in revised form: 12 September 2013 / Accepted: 16 September 2013 / Published: 9 October 2013

Abstract: Large-scale analyses of mammalian transcriptomes have identified a significant number of different RNA molecules that are not translated into protein. In fact, the use of new sequencing technologies has identified that most of the genome is transcribed, producing a heterogeneous population of RNAs which do not encode for proteins (ncRNAs). Emerging data suggest that these transcripts influence the development of cardiovascular disease. The best characterized non-coding RNA family is represented by short highly conserved RNA molecules, termed microRNAs (miRNAs), which mediate a process of mRNA silencing through transcript degradation or translational repression. These microRNAs (miRNAs) are expressed in cardiovascular tissues and play key roles in many cardiovascular pathologies, such as coronary artery disease (CAD) and heart failure (HF). Potential links between other ncRNAs, like long non-coding RNA, and cardiovascular disease are intriguing but the functions of these transcripts are largely unknown. Thus, the functional characterization of ncRNAs is essential to improve the overall understanding of cellular processes involved in cardiovascular diseases in order to define new therapeutic strategies. This review outlines the current knowledge of the different ncRNA classes and summarizes their role in cardiovascular development and disease.

Keywords: non-coding RNA; microRNA; long non-coding RNA; vascular development; vascular disease; heart pathophysiology

---

1. Introduction

Many studies have recently focused on understanding RNA metabolism and its implication in development and disease processes. Genomic tiling arrays and RNA-Sequencing have showed that the human genome is dynamically transcribed and leads to the production of a complex world of RNA molecules of which only a small fraction is translated into proteins [1]. In fact, application of high-throughput sequencing technologies in the analysis of mammalian transcriptomes, revealed a wide spectrum of RNA molecules that do not encode protein, termed non-coding RNAs (ncRNAs) [2]. For many years the role of these molecules remained unknown, so ncRNAs were called the "Dark Matter" of biology. To date many studies have been carried out on these molecules, especially on microRNAs, partially clarifying their roles. However many mechanisms and functions of different classes of ncRNA still remain unknown. Emerging evidence indicates that the non-coding portion of the genome is critical in the regulation of multiple biological processes, such as differentiation, development, post-transcriptional regulation of gene expression and epigenetic regulation [3–5]. Recently, many classes of ncRNA have been described to be associated with human disease [6]. Cardiovascular disease is a major cause of mortality and hospitalization worldwide [7], and the work of multiple research groups has been devoted to determine the molecular mechanism underlying heart and vascular disease. Recent studies indicate that altered ncRNA expression and function have been strongly implicated in cardiovascular disease such as myocardial infarction, cardiac hypertrophy and coronary artery disease [8–10]. The transcriptome of a cell contains different types of ncRNA that can be divided into two principal classes (Table 1): structural and regulatory ncRNAs. Structural ncRNAs include RNA molecules that are usually constitutively expressed such as ribosomal and transfer RNAs. Regulatory ncRNAs can be classified into three major classes based on transcript size: small (small ncRNAs), medium and long non-coding RNAs (lncRNAs) [6]. The most studied class of small ncRNAs in cardiovascular research is the microRNAs (miRNAs). MiRNAs are endogenous, single-stranded molecules consisting of approximately 20–22 nucleotides that regulate their target genes by reducing mRNA stability and/or translation [11]. Changes in microRNA expression lead to changes in gene function. This dysregulation of miRNA expression appears to play a significant role in the onset and progression of cardiovascular diseases [12]. Despite the progress in defining the role of microRNAs in cardiac and vascular biology, the complex network of ncRNAs and their interaction with different states of cardiovascular development and disease is still unknown. This is related to the multiple diversity of biogenesis, expression and functional properties of different classes of ncRNAs. Among these, the long non-coding RNAs (lncRNAs) are apparently the most numerous and functionally different [13]. LncRNAs are broadly classified as transcripts longer than 200 nucleotides and some of them are preferentially expressed in specific tissues [14]. Thus it is becoming increasingly clear that lncRNAs can regulate numerous molecular mechanisms. Recently, lncRNAs have emerged as new players in cardiovascular development and disease demonstrating potential roles in different cellular processes [15,16]. However, the characteristics and functions of the overwhelming majority of these lncRNAs are currently unknown. Accordingly, the functional characterization of lncRNAs is essential to advance our comprehensive understanding of cellular processes underlying cardiovascular development and disease. In the present review, emerging roles of ncRNAs in cardiovascular pathophysiology are discussed. Particular focus will be on the evaluation of biological roles of

microRNAs and lncRNAs in vascular as well as cardiac disorders. Moreover, the focus of this review is to provide an overview of the current state of knowledge of molecular processes implicated in differentiation and cardiovascular development, which are related to the function of ncRNAs.

Table 1. Classes of non-coding RNAs (ncRNAs).

Non-coding RNAs	Symbol	Functions
Structural ncRNAs		
Transfer RNA	tRNA	mRNA translation
Ribosomal RNA	rRNA	mRNA translation
Regulatory ncRNA		
Short ncRNA		
Micro RNAs	miRNA	post-transcriptional regulators
PIWI-interacting RNA	piRNA	DNA methylation, transposon repression
Short interfering RNA	siRNA	RNA interference
Medium ncRNA		
Small nucleolar RNAs	snoRNA	RNA modification, rRNA processing
Promoter upstream transcripts	PROMPTs	Associated with chromatin changes
Transcription initiation RNAs	tiRNAs	Epigenetic regulation
Long ncRNAs		
Long intergenic ncRNA	lincRNAs	Epigenetic regulators of transcription
Enhancer-like ncRNA	eRNA	Transcriptional gene activation
Transcribed ultraconserved regions	T-UCRs	Regulation of miRNA and mRNA levels
Natural antisense transcripts	NATs	mRNA stability
Promoter-associated long RNAs	PALRs	chromatin changes
Pseudogenes	None	microRNA decoys

2. An Overview on the Main Methods to Analyze the ncRNAs Expression

Each ncRNA has expression levels that are tissue- or stage-specific. In recent years several methods have been developed to study ncRNA expression. A common approach is Real-time PCR, which is employed mainly to analyze microRNA expression levels but can be used also for studies on long ncRNA [17]. Also many approaches based on immunoprecipitation assays have been developed in recent years (e.g., RNA immunoprecipitation or RIP, Cross-linking and immunoprecipitation or CLIP, RNA-chromatin immunoprecipitation or RNA-ChIP) [18–20]. RNA-IP was developed to identify ncRNAs, especially ncRNA, that interact with a specific protein. The basic principal behind all immunoprecipitation approaches is the same. Using a specific antibody it is possible to isolate a ncRNA-protein complex, then a cDNA library is constructed and the ncRNA is sequenced. Unfortunately, for any of these immunoprecipitation-based approaches the results are influenced by the specificity and affinity of the antibodies. Moreover, these methods (Real-time PCR or IP) allow evaluation of the expression of a few specific molecules but do not permit the discovery of new ncRNAs or provide an overview of all ncRNAs. Recently, advances in technology enabled the development of new genome-wide screening methods to study ncRNAs and their targets. Among these the most commonly used are microarray analysis and RNA sequencing. These technologies are very accurate and permit large-scale analysis of ncRNAs. In particular, the microarray [21–23] approach

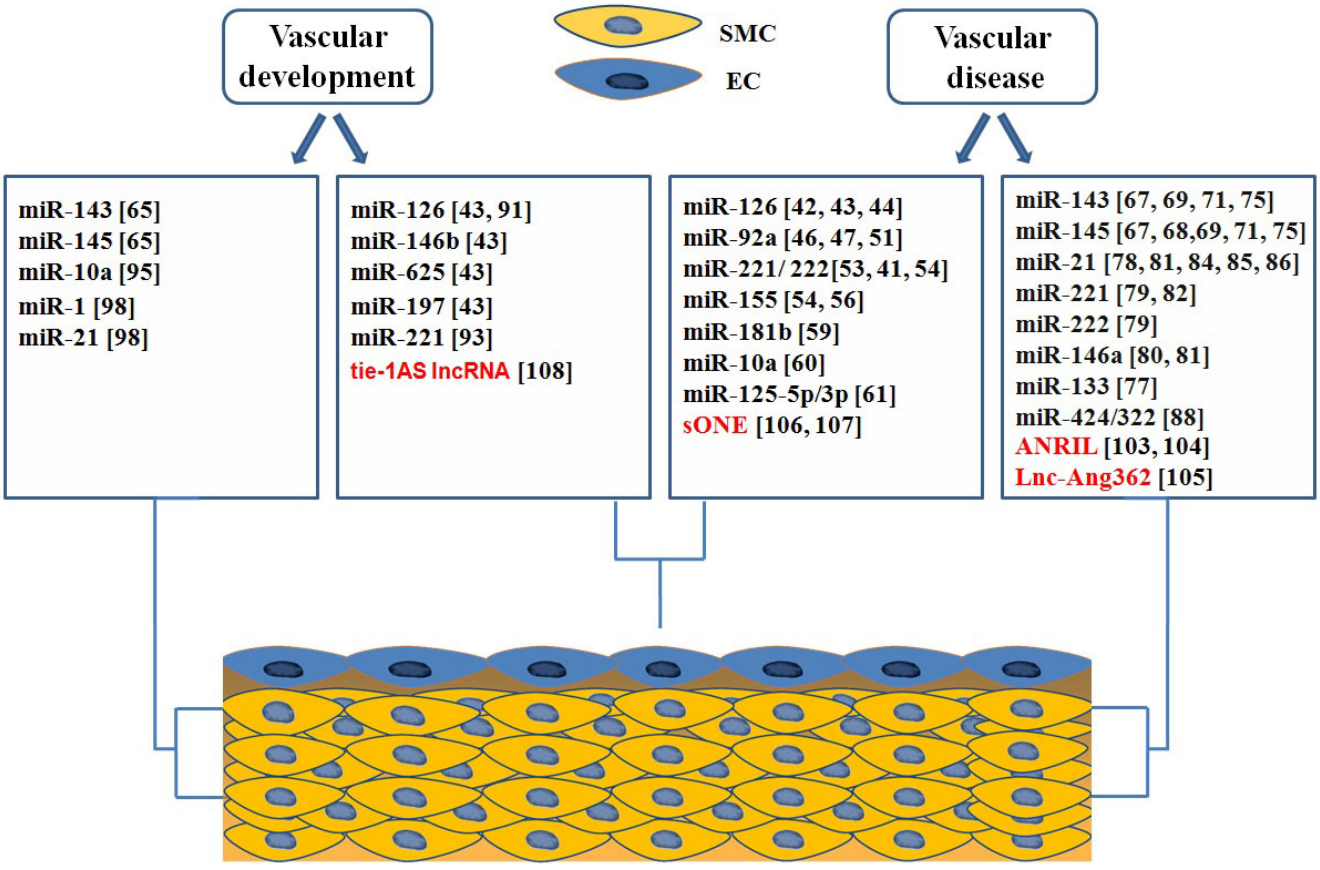
offers various platforms allowing the study of microRNAs and mRNAs targets, although to date there are only a few chips to analyze long non-codingRNA. Analysis with traditional microarrays is limited to detecting the presence or the absence of known ncRNAs and it is incapable of identifying new molecules or revealing different splicing variants. To get around this problem a new approach has been defined: tiling arrays. Unlike traditional microarrays, these platforms permit identification of new ncRNAs in a selected DNA region without prior knowledge of their precise location. For instance, Rinn *et al*. used this approach to study lncRNAs expressed in the region of HOX genes in humans [24]. The RNA sequencing or RNAseq [25] refers to the use of high-throughput sequencing technologies to get information about a sample's RNA content. This approach permits information to be obtained on differential expression of the interest gene, microRNA or long ncRNA. RNAseq is very sensitive in detecting less-abundant transcript and it can reveal alternatively spliced isoforms. Moreover, sequencing the entire transcriptome has been widely used to discover new non coding molecules. However, given the time and the cost related to the downstream analysis of the data generated by RNA sequencing, microarrays remain the first choice in many applications.

3. Functions of Non Coding RNAs

Although the function of most lncRNAs remains unknown, it has become clear that these molecules are intimately involved in many biological processes. LncRNAs can regulate gene expression programs through a variety of mechanisms, such as epigenetic modifications of DNA, alternative splicing, post-transcriptional gene regulation and mRNA stability and translation [5,26,27]. Given their established roles in transcriptional regulation, lncRNAs play a key role in several cellular events including proliferation, migration, apoptosis and development [21,28]. LncRNAs are now known to regulate the expression of protein-coding genes: they can positively or negatively control the expression of their target genes. Several lncRNAs are involved with *in cis* inactivation of larger genomic regions by epigenetic mechanisms. Kcnq1ot1 is a regulatory non coding antisense RNA that regulates epigenetic gene silencing in an imprinted gene cluster *in cis* [29]. This lncRNA specifically interacts with nearby genes in embryonic tissues causing transcriptional gene silencing. More recently, it was found that lncRNAs can act *in cis* to regulate expression of neighboring genes during cardiomyocyte differentiation [30]. Notably, many lncRNAs are now known to regulate the expression of genes by a *trans* mechanism. One example of a lncRNA that acts *in trans* is AK143260, termed Braveheart (Bvht) that specifically promotes activation of a core gene regulatory network to direct cardiovascular lineage commitment [15]. So far, several other functions have been attributed to lncRNAs. These molecules can act as scaffolds bringing together multiple proteins to form ribonucleoprotein complexes. For example, Miao-Chih Tsai *et al*. showed that a long non coding transcript, termed HOTAIR (HOX Antisense Intergenic RNA), acts as a scaffold for Polycomb Repressive Complex 2 (PRC2) and LSD1/CoREST/REST complex [31]. In addition to their role in chromatin regulation, lncRNAs can also function as molecular "decoys" of transcription factors and other regulatory proteins. PANDA (P21 associated ncRNA DNA damage activated) is an example of a lncRNA with decoy functionality. In fact, PANDA interacts with the transcription factor NF-YA to limit expression of pro-apoptotic genes [32]. Finally, the presence of a complex network of interactions between lncRNAs and miRNAs is becoming increasingly clear. In fact, lncRNAs may exert their biological activity through their ability to act as endogenous decoys for miRNAs. For

example, a muscle-specific long noncoding RNA, linc-MD1, could interact with two specific miRNAs, miR-133 and miR-135, and promote muscle differentiation by acting as a competing endogenous RNA (ceRNA) in mouse and human myoblasts [33]. Another lncRNA which has been identified in association with microRNAs is the pseudogene PTENP1 [34]. Similar to Linc-MD1, PTENP1 mRNA acts as a decoy for miRNAs that directly target the tumor suppressor protein PTEN. Accordingly, PTENP1 reduces down-regulation of PTEN messenger RNA. Recent reports also show that stable circular lncRNAs (circRNAs) can act as molecular decoys of microRNAs [35,36]. Taken together, these observations suggest that lncRNAs could have profound effects on several molecular mechanisms. Nevertheless, lncRNAs are poorly conserved among species resulting in an additional degree of complexity in the definition of their functions. Despite rapid progress in lncRNA discovery, evidence of physiologic function for lncRNAs remains poor and further investigation is necessary.

Figure 1. Role of non-coding RNAs in Vascular Development and Disease.



4. Roles of ncRNAs in Vascular Biology and Disease

The vessel wall is composed of endothelial cells (ECs) and smooth muscle cells (SMCs) that play central roles in vascular biology and disease. In fact, these cells can undergo profound changes in phenotype during vascular injury and remodeling; these changes are correlated with pathologies such as atherosclerosis and proliferative thickening of the vessel known as restenosis. Atherosclerosis is a chronic inflammatory disease of the arterial wall and is the major cause of death in western countries [37]. It is a complex process involving multiple cell types and the interactions of many different molecular pathways. The events that lead to the formation of atherosclerotic lesions include

modification of endothelial cell function, monocyte adherence and entry into vessel wall, phenotypic modulation of smooth muscle cell, and platelet adhesion and aggregation [38]. Phenotypic modulation of smooth muscle cells is, also, crucial in the neointimal lesion formation after stent implantation [39]. Numerous ncRNAs, especially microRNAs, have been shown to govern these processes during vascular disease. In fact, miRNA control endothelial cell and vascular smooth muscle cell biology, and thereby regulate the progression of vascular disease, such as atherosclerosis and restenosis. Current evidence also suggests that other ncRNA classes, such as lnc-RNA molecules play a critical role in endothelial and smooth muscle cell function. Figure 1 summarizes the role of ncRNA classes in different cells of the vessel wall.

4.1. microRNAs in Endothelial Biology and Dysfunction

In endothelial cells (ECs) the action of specific miRNAs is important for vascular signaling and function. Different studies indicate that the major miRNA-regulating enzymes, Dicer and Drosha, are essential for angiogenic functions of endothelial cells [40,41]. The endothelial-specific miR-126 is the most abundant miRNA found in adult ECs and it is involved in endothelial dysfunction and inflammation [42]. It is interesting to observe how miR-126 regulates the response of ECs to VEGF by inhibiting sprout-related protein SPRED1, a negative inhibitor of VEGF signaling [43]. Another group demonstrated that VCAM-1 is a direct target of miR-126 [44]. In the early phase of atherosclerotic disease, inflammatory cytokines increase a series of adhesion molecules, such as VCAM-1, on the surface of ECs. Inhibition of miR-126 increases leukocyte adherence in TNFα-stimulated ECs. Endothelial cell functions are critically regulated by other microRNAs: the miR-17-92 cluster, a polycistronic miRNA gene that produces six mature miRNAs: miR-17, miR-18a, miR-19a, miR-19b-1, miR-20a, and miR-92a [45]. Individual members of the miR-17-92 cluster, function as negative regulators of angiogenesis. In particular, miR-92a inhibited angiogenesis by targeting several functional genes, including integrin α5 (ITGa5) [46]. In addition, miR-92a negatively regulates KLF2 and KLF4 expression in athero-susceptible endothelium [47]. Given that both endothelial KLF4 and KLF2 are implicated in protection against atherogenesis [48–50], miR-92a may be important in arterial disease. Moreover, we have recently analyzed the effect of miR-92a in endothelial cell by loss-of-function studies [51]. Our group demonstrated that systemic administration of a complementary oligonucleotide (antagomiR-92a) significantly enhanced re-endothelialization in carotid arteries after balloon injury or arterial stenting. Our group and others [46,51] showed the relationship between miR-92a and endothelial nitric oxide synthase (eNOS) expression. Nitric oxide (NO) limits the formation of neointimal hyperplasia in animal models of arterial injury to a large part by inhibiting vascular smooth muscle cell proliferation [52]. Accordingly, the functional consequences of the miR-92a inhibition are an increase in NO bioavailability and an antiproliferative effect on SMCs [51]. A further example of negative correlation between microRNAs and eNOS activity is represented by miR-221 and miR-222. These microRNAs are highly expressed in ECs and exhibit anti-angiogenic effects [53]. Notably, over-expression of miR-221 and miR-222 indirectly reduces the expression of eNOS [33]. miR221 and miR-222 directly target c-kit, the receptor for stem cell factor (SCF), which plays a key role in endothelial cell migration [53]. Recently it has been shown that the miR-221 and miR-222 are negatively correlated with the expression of Ets-1 [54] that regulates the expression of several inflammatory molecules in the endothelial cell during vascular inflammation [55]. Another miRNA

which has been identified in endothelial cells is miR-155. Similar to miR-221 and miR-222, miR-155 directly targets ETS-1 in ECs [54]. Also, miR-155 down-regulates eNOS expression through decreasing eNOS mRNA stability by binding its 3'-UTR [56]. Given their role in regulating endothelial cell biology, miR-221, miR 222 and miR-155 represent possible therapeutic targets in the inflammatory response of endothelial cells during the initial stage of atherosclerosis. Several other groups provide additional examples of the intersection between microRNAs and endothelial cell activation and dysfunction. In response to inflammatory stimuli, the nuclear factor-KappaB (NF-κB) signaling pathway is activated leading to the expression of multiple pro-inflammatory genes in ECs [57]. In fact, in Apolipoprotein E (ApoE)-deficient mice, endothelial cell-specific inhibition of NF-κB resulted in reduced development of atherosclerosis [58]. Two endothelial-specific microRNAs, miR-10a and miR-181b, inhibit the activation of the NF-κB signaling pathway in ECs. Recently, miR-181b has been identified as a key player in vascular inflammatory disease. miR-181b expression is reduced in response to TNF-a in the vascular endothelium, whereas its over-expression inhibits TNF-α-induced NF-κB-responsive targets gene such as VCAM-1 and E-selectin [59]. Moreover, miR-181b targets importin-α3, a critical protein in NF-κB nuclear translocation and activation. miR-10a directly inhibits mitogen-activated kinase kinase kinase 7 (MAP3K7) and beta-transducin repeat-containing gene (β-TRC) [60]. These molecules are essential in promoting IκBα degradation, an inhibitor of NF-κB activation. Inhibition of miR-10a enhances the NF-κB-dependent expression of adhesion molecules in ECs. Other specific microRNAs that regulate endothelial cell function have been described. For example, miR-125a-5p and miR-125b-5p have been identified as negative regulators of ET-1 [61], a potent vasoconstrictive and mitogen peptide that plays multiple roles in the progression of vascular disorder [62]. Taken together, the results described above indicate that several microRNAs play an essential role in endothelial pathophysiology. Accordingly, the identification of specific microRNAs involved in biological processes, such as angiogenesis and inflammation, could lead to the definition of new strategies to treat vascular diseases. Given that the same microRNAs may have opposite effects in different biological contexts, further studies are necessary to clarify their roles in endothelial dysfunction. For example, identification of signaling pathways which modulate the activity of microRNAs is critical for development of microRNA-based therapeutic strategies.

4.2. microRNAs in Phenotypic Switching of VSMCs

SMCs within adult animals retain remarkable plasticity and can undergo profound and reversible changes in phenotype, a process referred to as phenotypic switching [63]. SMCs play a role during all phases of the atherogenic process as well as in proliferative disease [64–66]. Several microRNAs are implicated in VSMC phenotypic switching in response to vascular injury or atherosclerotic disease.

4.2.1. miR-143 and miR-145 Play a Role in the Regulation of Phenotype of VSMCs in Response to Injury

miRNA-143 and -145 are considered the master regulators of contractile phenotype by promoting contractile protein expression [67]. The expression levels of miR-143/145 are down-regulated in injured or atherosclerotic vessels and are associated with the phenotypic switch from a contractile/quiescent to a synthetic/proliferative phenotype. A recent study reported that adenovirus-mediated over-expression of

miR-145 could partially restore down-regulation of SMC marker genes and neointima formation following balloon injury of the rat carotid artery [68]. In addition, miR-143/145-knockout mice present morphological changes in the aorta, due to an incomplete differentiation of VSMCs [69]. Several growth factors promote phenotypic switching of VSMCs during vascular disease [63]. The growth factor, PDGF-BB, is a critical regulator of VSMCs phenotype in vessel injury. Indeed, PDGF stimulation increases migration and proliferation of SMCs *in vitro* and *in vivo* [70]. It has been shown that PDGF can reduce miR-145 and miR-143 expression through Src and p53 activity [71] and promote formation of podosomes. miR-143 and miR-145 directly target key regulators of podosome formation, such as PDGF receptor α (PDGF-Rα), protein kinase Cε (PKCε) and fascin. Another cytochines that participate in the phenotypic control of VSMCs are Transforming Growth Factor beta (TGFβ) and bone morphogenetic proteins (BMPs). Unlike PDGF, the TGF-family of growth factors has been shown to promote contractile phenotype [72–74]. Induction of miR-143 and miR-145 by TGF-β or BMP4 leads to down-regulation of KLF4 expression and activation of contractile genes [75]. It is interesting to note that the expression of miR-143 and miR-145 is regulated by multiple growth factor signaling pathways promoting phenotypic modulation of SMC. Therefore, because miR-143 and mir-145 may be important modulators of vascular disorder, further studies are necessary to define the regulatory mechanisms of their expression in response to vascular injury or atherosclerotic disease.

4.2.2. miR-133 Is a Negative Regulator of SP-1 and Promotes Contractile Phenotype of VSMCs

Similarly to miR-143 and miR-145, a number of additional microRNAs play a role in smooth muscle cell phenotypic switching and in vascular disease. In a recent study from our laboratory, we found that miR-133 has a potent inhibitory role on VSMC phenotypic switching. miR-133 specifically suppresses Sp-1 expression *in vitro* and *in vivo* and participates in a complex network with Serum response Factor (SRF) in regulating smooth muscle gene expression. Sp-1 is a key regulator of KLF4 expression in phenotypically modulated SMCs [76]. KLF-4 is up-regulated in VSMCs following vascular injury and inhibits myocardin-induced SMC marker gene expression. Following balloon injury of the rat carotid artery there is an increase of Sp1 expression in the neointima related to the acquisition of proliferative/synthetic phenotype of VSMCs. Accordingly, over-expression of miR-133 reduces neointima formation and SMCs proliferation after balloon injury of the rat carotid artery [77].

4.2.3. microRNAs that Promote De-Differentiated Phenotype of VSMCs

Other miRNAs have been implicated in the phenotypic modulation of SMCs: miR-21, miR-146a, miR-221, and miR-222. Unlike the miR-143/145 cluster and miR-133, these microRNAs were found to be significantly up-regulated after vascular injury [78–80] and their inhibition reduces neointimal formation following balloon injury of rat carotid arteries *in vivo*. Therefore, expression profiles of microRNAs in human atherosclerotic plaques in comparison to control, demonstrate that miR-21 and miR-146a were up-regulated in human atherosclerotic plaques, whereas several predicted targets of these miRNAs were down-regulated [81]. miR-221 and miR-222 contribute to SMC phenotype by repressing specific targets such as p57Kip2 and p27Kip1 [79]. They are cyclin-dependent kinase inhibitors and have an antiproliferative effect on VSMCs. Interestingly, PDGF induces the expression of miR-221, leading to down-regulation of multiple target genes and promoting the proliferation of

SMCs [82]. Accordingly, inhibition of miR-221 prevents reduction of p27Kip1 in response to PDGF, as well as VSMC proliferation. miR-146a promotes SMCs proliferation *in vitro* and neointimal hyperplasia *in vivo* [80]. Notably, miR-146a inhibits KLF4 expression by targeting its 3'-UTR. Inhibition of miR-146a increases KLF4 expression, while its over-expression induces an opposite effect. It has been reported that KLF4 has a critical role in the regulation of the SMCs phenotype [83] and its expression is correlated to different microRNAs. Particularly, miR-143 and miR-145 directly target KLF4 in SMCs [75]. miR-146a promotes the proliferative phenotype of SMCs through a reduction of KLF4 expression; in contrast, KLF4 regulation of miR-145 and miR-143 promote the contractile phenotype of SMCs. Thus, further studies are necessary to define the regulatory mechanisms of KLF4 in SMCs. miR-21 promotes the cellular response that leads to proliferative thickening of the vessel by directly targeting Phosphatase and Tensin Homolog (PTEN), a critical regulator of SMCs function both *in vivo* and *in vitro* [78]. SRF is a transcription factor that plays a critical role in SMCs biology, influencing both proliferation and differentiation depending on the types of coactivators or repressors present at specific cellular stages. A recent report demonstrated that Serum Response Factor (SRF) regulates PTEN expression through a reduction of miR-21 levels [84]. Regulation SRF-mediated of miR-21 occurs through a miR-143-dependent signaling pathway. Recently, the role of miR-21 in the regulation of SMC phenotype was correlated with abdominal aortic aneurysm (AAA). The expression levels of miR-21 increase during development of AAA in two murine models [85]; moreover Lentivirus-mediated over-expression of miR-21 induced cell proliferation of SMCs, with protective effects on aneurysm expansion. In addition, miR-21 targets several signal molecules associated with SMCs phenotype such as Programmed Cell Death 4 (PDCD4) [86], B-cell leukemia/lymphoma 2 (BCL-2) [78], and Tropomyosin 1 (TPM1) [87]. These observations indicate that miR-21 could play important roles in diverse vascular diseases. Different microRNAs can be involved in the same process; for example recent studies [88] demonstrated that miR-424/322 also play a key role in modulation of SMCs phenotype in response to vascular injury. In particular, ectopic expression of miR-424/322 induces inhibition of proliferation and migration in SMCs and reduces restenosis in injured carotid arteries in rats. miR-424/322 regulates SMCs phenotype suppressing its direct targets cyclin D1 and CA+2 regulating proteins (calumenin). Interestingly, miR-424/322 is significantly up-regulated after vascular injury and this suggest that miR-424/322 is correlated to an adaptive response to counteract proliferation of SMCs. To summarize, these findings suggest a new therapeutic strategy for vascular diseases connected with phenotypic switching of SMCs. Interestingly, chemically modified antisense oligonucleotides, termed “antagomirs”, have been used to decrease miRNA expression and function in different animal models [32,37,64]. Given that many microRNAs are increased after injury, it is possible that their specific inhibition by antagomirs could be considered as potential therapeutic targets for several vascular diseases. Nevertheless, further research is needed regarding the role of these aberrantly expressed microRNAs in SMCs.

4.3. Roles of microRNAs in Vascular Development

The formation of a vascular system requires the creation and remodeling of a continuous series of vessels. They are made mainly by endothelial cells, but smooth muscle cells ensure the correct tone and contractility of the vessels necessary for proper blood flow. After birth, VSMCs retain remarkable plasticity; they can switch between a contractile and proliferative phenotype, a characteristic fundamental in vascular development and remodeling. Both these processes are regulated by numerous factors, including microRNAs. miRNAs have been implicated in endothelial cell differentiation and are involved in the regulation of formation of blood vessels during vascular development. Knockout of DICER, the enzyme responsible for the maturation of microRNAs, reduces postnatal angiogenesis in response to several stimuli, such as exogenous VEGF, tumors, limb ischemia, and wound healing [89]. Specific silencing Dicer using siRNA, increases activation of the eNOS pathway but reduces proliferation and cord formation of human endothelial cell *in vitro* [41].

4.3.1. miRNAs Involved in Endothelial Development

The first miR shown to be essential for vessel formation and integrity is miR-126. It is a positive regulator of angiogenic signaling in endothelial cells and also of vascular integrity *in vivo*. miR-126-deficient endothelial cells failed to respond to various angiogenic factors, including VEGF, EGF and bFGF [43,90,91]. Studies on zebrafish show that down-regulation of this microRNA reduces vascular integrity and induces hemorrhages [43]; furthermore, studies in mice demonstrate that the deletion of miR-126 causes defects in endothelial cell proliferation, migration and angiogenesis [91]. However, both these studies demonstrate that miR-126 affects endothelial cell function but it is not essential for cell differentiation or embryonic vessel formation. Fish *et al*. show additionally that miR-146b, miR-625 and miR-197 appear up-regulated in mouse ESC-derived endothelial cells, however the role of these miRNAs in the vascular system is still unknown [43]. Kane *et al*., in order to study the endothelial differentiation, carried out a Taqman low-density array (TLDA) analysis of miRNA levels in pluripotent hESCs and in hESCs differentiated for 10 days. They found that up-regulation of miR-126 causes an increase in levels of further microRNAs such as let-7 family, miR-210, miR-130a, miR-196, miR-133a [92]. However, despite the fact that all of these microRNAs are modulated during endothelial differentiation, not one of them has been shown to be directly correlated to endothelial cell control. Recently, Nicoli *et al*. investigated the role of miR-221 in endothelial cells in vascular development; they demonstrated that miR-221 promotes endothelial cell proliferation and migration through repression of two targets: cyclin dependent kinase inhibitor 1b (cdkn1b) and phosphoinositide-3-kinase regulatory subunit 1 (pik3r1). Also miR-221 expression is inhibited by Notch signaling [93].

4.3.2. microRNAs in Regulation of Vascular Smooth Muscle Cell Differentiation

Several studies have provided compelling evidence that microRNAs play a critical role in the initial specification of vascular smooth muscle cell lineage during development. Indeed, inactivation of Dicer in VSMCs results in late embryonic lethality due to decreased VSMC proliferation and differentiation and due to vascular abnormalities and extensive hemorrhage [94]. Thus, the function of

Dicer-generated miRNAs is essential during development of VSMC. A recent report demonstrated a crucial role for miR-143 and miR-145 in VSMC differentiation [65]. The down-regulation of miR-145 using cholesterol-modified antisense oligonucleotides inhibits myocardin-induced reprogramming of fibroblasts into SMC and represses expression of multiple SMC markers, such as ACTA2, MyH11 and Calponin. Also, miR-145 over-expression is sufficient to induce differentiation of multipotent neural crest stem cells into smooth muscle cells and to inhibit their proliferation. Accordingly, miR-143 and miR-145 target a network of factors, such as KLF4 and ELK-1, in order to promote VSMCs differentiation and repress proliferation. Other miRNAs that may promote VSMC differentiation are miR-10a, miR-1 and miR-21. Recently, it has been shown that miR-10a is involved in the differentiation to smooth muscle cell lineage from mouse ESCs in response to Retinoid acid [95]. Several studies indicate that Retinoid Signaling positively influences the SMC differentiation program from stem cells [96]. Notably, miR-10a is up-regulated during retinoid acid-induced SMC differentiation. Furthermore, miR-10a directly targets histone deacetylase 4 (HDAC4), which is a negative regulator of SMC differentiation [97]. miR-1 plays a critical role in the SMC lineage differentiation in embryonic stem cell-derived SMC cultures[98]. miR-1 expression is highly up-regulated during differentiation of mouse embryonic stem cell (ESC) to SMCs. miR-1 has been implicated in SMC differentiation by directly targeting the 3'UTR of KLF4 and enhancing expression of the smooth muscle-restricted markers gene. miR-21 was shown to promote differentiation of VSMCs in response to transforming growth factor and bone morphogenetic protein stimulation [99]. miR-21 directly targets PDCD4 (programmed cell death 4), which acts as a negative regulator of smooth muscle contractile genes.

4.4. Long Non Coding RNAs in Vascular Development and Disease

Although the heterogeneous group of lncRNAs play a wide range of roles in cellular function, their characterization pertaining to vascular development and disease is limited to only a few examples. Variation on chromosome 9p21 is associated with risk of coronary artery disease (CAD) [100,101]. This genomic region contains a long intergenic noncoding RNA, designated antisense noncoding RNA in the INK4 locus (ANRIL). ANRIL is a long non-coding RNA which is transcribed from the INK/ARF locus. ANRIL is expressed in tissues and cell types that are affected by atherosclerosis such as primary coronary smooth muscle cells, vascular endothelial cells, human monocyte-derived macrophage cells and RNA extracted from carotid and arterectomy [102]. Notably, increased expression of ANRIL transcripts was directly correlated with the severity of atherosclerosis [103]. However, despite the potential importance of ANRIL to vascular disease, the pathophysiology underlying the link between ANRIL and coronary artery disease remains currently unknown. ANRIL has been associated with epigenetic silencing of the INK4B-ARF-INK4A locus on chromosome 9p21.3 [104]. In fact, ANRIL binds the p15(INK4b) transcript and recruits the Polycomb Repressor Complex (PRC) to repress the transcription of genes at this locus. Therefore, it is possible that the increased expression of ANRIL is correlated with altered expression of p15INK4B leading to coronary artery disease. Future studies will be necessary to define the role of ANRIL in vascular disease. Using whole transcriptome sequencing, a recent publication revealed the expression profile of lncRNAs in VSMCs in response to Ang II [105]. In this paper, the authors showed that two miRNAs, miR-221 and

miR-222, are co-transcribed with a specific lnc-RNA, Lnc-Ang362. These microRNAs have been found to play a critical role in smooth muscle cell proliferation and neointimal hyperplasia in response to vascular injury [79,82]. Interestingly, knockdown of Lnc-Ang362 reduces the expression of these miRNAs as well as cell growth. Correlations between the expression of lncRNAs and miRNAs raise the intriguing possibility of complex functional regulatory pathways in which several types of ncRNAs interact and influence the phenotype of VSMCs during vascular disease. Future studies are needed to dissect the exact roles of lncRNAs in phenotypic switching of VSMCs. A further example of a ncRNA correlated with vascular disease is a natural antisense transcript (NAT), termed sONE. A key function of this lncRNAs is the regulation of eNOS expression in a post-transcriptional manner under normoxic and hypoxic conditions [106,107]. Over-expression of sONE in endothelial cells reduces eNOS expression. Alterations of NO production by the vascular endothelium results in endothelial dysfunction, which occurs as a prelude to atherosclerosis. Thus it could be interesting to investigate the role of antisense lncRNA sONE in the post-transcriptional regulation of eNOS in order to define a potential therapeutic target in vascular disease. More recently, Keguo Li *et al*. showed that a new long noncoding antisense transcript, termed tie-1AS lncRNA, is required for the regulation of tyrosine kinase containing immunoglobulin and epidermal growth factor homology domain-1 gene (tie-1) levels *in vivo* and *in vitro* [108]. Analysis of tie-1AS lncRNA and tie-1 revealed that the ratio of tie-1 versus tie-1AS lncRNA is opposite in normal placenta tissue compared with vascular anomaly tissue. Also, the tie-1AS lncRNA selectively binds tie-1 mRNA, resulting in down-regulation of tie-1 protein and thus specific defects in endothelial cell contact junctions. Over-expression of tie-1 AS lncRNAs resulted in defects in endothelial cell junctions and tube formation. For the first time these results identified a lncRNA that plays a functional regulatory role with potential implications in the control of vascular development. In summary, the results described above indicate that lncRNAs are involved in different aspects of development and disease but their role in the cardiovascular system remains to be further investigated.

5. ncRNA in Heart Development and Pathophysiology

The heart is the first organ to form during embryo development, a complex process involving many classes of regulatory molecules. Heart uninterrupted contractility and correct function are essential for life; its alterations are associated with numerous diseases, including atherosclerosis and stroke [7]. Due to its importance, a complex system of transcription factors closely intertwined with a family of different molecules precisely controls multiple aspects of heart development, function and dysfunction. In these mechanisms, a central role is played by ncRNAs.

5.1. microRNA in Heart Development

5.1.1. miRNAs Encoded by MHC Genes

Two miRNAs highly expressed in the heart, miR-1 and miR-133, are implicated in the control of cardiac growth, regulating fundamental aspects of heart development *in vivo*. There are two isoforms of miR-1, miR-1-1 and miR-1-2, whereas miR-133 presents three isoforms, miR-133a-1, miR-133a-2 and miR-133b. These miRNAs are strictly related, in fact miR-1 and miR-133a are encoded by the

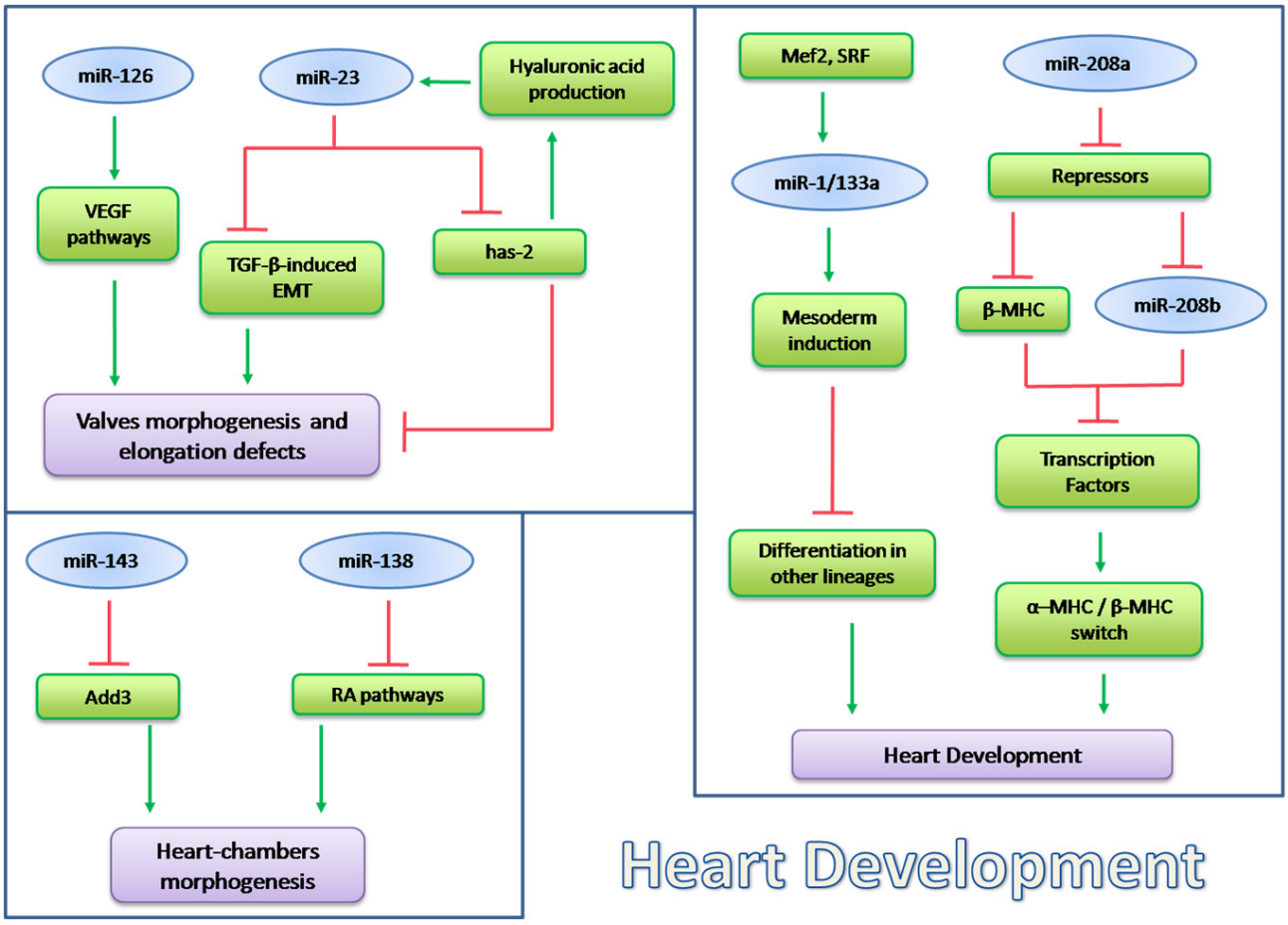
same gene, in particular miR-1-1/miR-133a-2 and miR-1-2/miR133a-1. The expression of both these clusters is under the control of SRF and Mef2. SRF enhances the expression of these miRNAs in ventricular and atrial myocytes through a serum response element. Instead, Mef2 binds an intronic enhancer of these miRNAs to activate their expression in ventricular myocytes [109]. During differentiation from ES cells, miR-1 and miR-133 are expressed to promote mesoderm induction and to suppress the differentiation in other lineages, in mice [110,111]. Furthermore, the over-expression of miR-133 causes the inhibition of cardiomyocyte proliferation. Particularly, the up-regulation of miR-133 results in embryonic lethality due to the thinning of the ventricular walls and VSDs (Ventricular Septal Defects); whereas miR-133a-null mice present an ectopic expression of smooth muscle genes in the developing heart as well as aberrant cardiomyocyte proliferation [112]. The up-regulation of miR-1 causes embryonic lethality due to a deficiency in cardiomyocytes [109]. Furthermore deficient miR-1 mice exhibit an increased number of proliferating cardiomyocytes [113]. Other microRNAs encoded by MHC genes are implicated in heart development and stress responsiveness. α-MHC encodes for miR-208a, a heart-specific miRNA, while a closely related microRNA, miR-208b, is encoded by β-MHC. Both these miRs have the same seed sequence, but differ in their 3'UTR region. miR-208a is highly expressed in adult mouse heart, whereas mir-208b is very abundant in embryonic heart but it is present at low levels in an adult heart. Furthermore, while miR-208b is expressed in the heart and in other tissues like skeletal muscle, miR-208a is only expressed in the heart [114]. Interestingly the expression of MHC genes, essential for cardiac muscle contraction, is not equal during life; in fact α-MHC is the predominant myosin isoform in the adult heart, whereas β-MHC is highly expressed in the developing heart but is down-regulated after birth. Cardiac stress and diseases modulate MHC gene transcription, causing a switch in myosin content in the heart, which has a marked effect on cardiac contractility and function [115]. Clearly miR-208a and miR-208b follow the trend of their host genes, hinting that they have a role in the regulation of the α-MHC to β-MHC switch and consequently in cardiac conduction, in arrhythmias and in other aspects of the stress response. Nevertheless none of these microRNAs are essential for function of the adult heart, but they appear to function primarily to adapt adult cardiac gene expression to physiological and pathological signaling.

5.1.2. microRNAs in Heart Chamber Morphogenesis

The heart is composed of cells with similar origin that develop divergent patterns of gene expression with numerous transcriptional networks that establish chamber or domain-specific gene expression and function. A precise regulation in time and space of gene expression and protein activity is necessary for a correct cardiac patterning. The most important microRNA in this process is miR-138; it was studied in zebrafish which represents an excellent model to study heart development, despite it having a heart containing a single atrium and ventricle [116]. Morton *et al*. show that miR-138 is required for cardiac maturation, in fact knock-down of this miR in zebrafish embryo causes the failure of ventricular cardiomyocytes to fully mature. So, miR-138 is necessary to establish an appropriate chamber-specific gene expression pattern during embryo development [116]. This miRNA is expressed in specific domains of the heart and targets various members of different pathways, particularly of retinoic acid (RA). It establishes discrete temporal and spatial domains of gene

expression during cardiac morphogenesis, ensuring a correct development of the heart. Another microRNA required for proper morphogenesis of heart chambers is miR-143. It directly targets adducin3 (add3), an F-actin capping protein. As reviewed by Taber [117], alterations in cytoskeletal dynamics could drive cardiac morphogenesis by promoting regional changes in cell size and shape. So, the regulation operated by miR-143 on add3-pathway modulates cytoskeletal protein and, consequently, it could influence heart development and particularly chamber formation through active adjustment of myocardial cell morphology [118].

Figure 2. miRNAs involved in Heart Development. Schematic representation of the relevant microRNAs involved in heart development with a subset of their principal targets: miRNAs which regulate heart-chambers morphogenesis (miR-143, miR-138); miRNAs involved in valves morphogenesis and elongation defects (miR-126, miR-23); miRNAs regulating heart differentiation and development.



5.1.3. microRNAs in Valves Development

Defects in cardiac valves are the most common subtype of cardiovascular malformation and, in adults, are a major cause of morbidity and mortality. One gene was found which was involved in the regulation of this process: miR-23. Lagendijk *et al*. studied miR-23 in endothelial cells of mice and demonstrated that in this model miR-23 was able to inhibit a TGF-β-induced endothelial to mesenchymal transition (EMT), a process that normally occurs during heart valve development. They proposed that miR-23, has-2 (hyaluronic acid synthase 2) and hyaluronic acid (HA) create a regulatory

feedback loop that could respond to various signals, including TGF-β [119]. Recently it has been shown in zebrafish that the loss of this microRNA causes endocardial defects, including cushion formation. miR-23 acts by the down-regulation of has-2, an extracellular remodeling enzyme required for endocardial cushion and valve formation. So this microRNA in the embryonic heart is required to restrict endocardial cushion formation by inhibiting has-2 expression and extracellular hyaluronic acid production [119,120]. A recent study shows that miR-126 is also implicated in valve elongation defects [121]. As shown above, this miR has a role in VEGF signaling in the development of endocardial cells. miR-126 targets a subunit of PI3K and Spred1, two negative regulators of VEGF pathway, so it positively regulates VEGF signaling in heart valve morphogenesis [122]. Despite current knowledge on microRNA functions in cardiovascular development, our complete understanding of their role is far from complete. Figure 2 summarizes the role of microRNAs in heart development.

5.2. microRNA in Heart Pathophysiology

The possibility that microRNAs might participate in heart disease was first suggested by the discovery of distinctive patterns of microRNA expression in the hearts of normal mice *vs*. mice that suffered from heart disease [123]. Recent studies on miRNA expression in heart have identified a subset of miRNAs highly expressed in the normal heart and modulated during cardiovascular disease [124].

5.2.1. miR-195 and miR-98/let-7b in Cardiac Hypertrophy

The first-characterized miRNA involved in inducing hypertrophic growth in the adult heart was miR-195. Adenoviral-mediated over-expression of this microRNA leads to dilated cardiomyopathy and heart dysfunction *in vivo*, it is also sufficient to induce hypertrophy in neonatal rat cardiomyocytes [123]. Chen and colleagues demonstrated the requirement of proper LKB1/STRAD/MO25 complex formation for full activation of AMPK signaling; miR-195 is sufficient to suppress MO25 expression and downstream targets of the LKB1/STRAD/MO25 pathway [125]. They hypothesized that miR-195 targets the LKB1/AMPK signaling axis in hypertrophic cardiomyopathy progression, implicating a functional role of this microRNA in this process.

Conversely, miR-98/let-7b has been demonstrated to mediate the anti-hypertrophic effect of thioredoxin (Trx1), an ubiquitously expressed antioxidant that inhibits NF-κB (nuclear factor kappa-light-chain enhancer of activated B cells), Ras and ASK1 (apoptosis signal-regulating kinase 1). Trx1 negatively regulates the protein kinase cascade known to stimulate hypertrophy. Particularly, Yang *et al*. studied the effects of miR-98 up-regulation or down-regulation on cardiac hypertrophy *in vivo*, at baseline and in response to Ang-II. These studies show that Trx1 negatively regulates Ang-II-induced cardiac hypertrophy through up-regulation of miR-98/let-7b but does not affect heart morphology at baseline [126]. A validated target of miR-98/let-7 is cyclin D2, a cyclin that plays a key role in hypertrophy mediated by this microRNA; the down-regulation of Trx1 causes the up-regulation of miR-98 and the inhibition of hypertrophy.

5.2.2. miR-1 in Heart Physiopathology

The most abundantly expressed microRNA in human heart is miR-1; as described before it is clear that it has a key role in the developing heart. After subjecting the heart of mice to increased pressure overload a down-regulation of miR-1 resulting in an increase in cardiac mass and contractile dysfunction was observed [113,127]. Furthermore, miR-1 is down-regulated in several models of cardiac hypertrophy and heart failure, conversely its over-expression attenuates cardiomyocyte hypertrophy indicating that miR-1 down-regulation has a causative role in the pathogenesis of this disease [113]. It is interesting to note that miR-1 was shown to regulate different pathways implicated in heart hypertrophy. First, it modulates calmodulin and Mef2a, two mediators of calcium signaling, and in addition the transcriptional effectors MEF2A and GATA4, suggesting that miR-1 controls calcium signaling by different modalities simultaneously [113,128]. Second, a dysregulation of insulin-like growth factor (IGF-1) has also been involved in pathological hypertrophy; it is a validated target of miR-1. In exercised trained rats and cardiac-specific Akt transgenic mice, which are models of physiological cardiac hypertrophy, miR-1, as well as miR-133, are down-regulated [110]. There is an inverse correlation between miR-1 and IGF-1: the microRNA controls the expression of IGF-1 and IGF-1 receptor and reciprocally it is down-regulated by IGF-1 stimulation depending on the activation of PI3K/AKT pathway and repression of Foxo3 transcription factor. Accordingly, acromegalic patients, in whom there is an atypical synthesis of IGF-1, display increased cardiac mass and wall thickness [127]. Finally, miR-1 targets twifilin 1 (Twf1), a cytoskeletal regulatory protein that binds to actin monomers preventing their assembly into filaments. The level of Twf1 is inversely correlated with expression of miR-1, so it is expressed at low levels in an adult heart. Moreover down-regulation of miR-1 induced by hypertrophic stimuli results in increased Twf1 expression; likewise Twf1 over-expression is sufficient to induce cardiac hypertrophy in neonatal rat cardiomyocytes, suggesting the therapeutic relevance of modulation of Twf1 expression in attenuating cardiac hypertrophy [129]. Moreover, miR-1 was studied in a model of acute ischemic heart disease (IHD): in cardiomyocytes from ischemia/reperfused (I/R) rats, this microRNA appears up-regulated and inversely correlated with the anti-apoptotic protein Bcl-2 [130], suggesting a potential role of this miRNA in cardiomyocyte apoptosis. In heart failure models levels of miR-1, like miR-133, appear decreased; the same effect is observed in the hearts of patients with hypertrophic cardiomyopathy or atrial dilation [110,131]. The down-regulation of miR-1, and likewise that of miR-133, is associated with the increased levels of two members of the HCN ion channel family, HCN2/HCN4, in hypertrophic hearts. Probably the up-regulation of these channels may contribute to enhanced automaticity and arrhythmias in heart failure [131]. Moreover, a recent study demonstrated that miR-1 directly targets connexin 43 (Cx43), the main cardiac connexin, which has an aberrant increased expression in hypertrophic cardiomyocytes *in vitro* and *in vivo* [132]. Given the numerous processes regulated by miR-1, it could be an important therapeutic target, however alterations of miR-1 levels could alter several mechanisms, so other studies are necessary before its application in the medical field.

5.2.3. miRNA Implicated in Calicineurin/NFATs Pathway

Cellular and *in vivo* models of cardiac hypertrophy induced by transverse aortic constriction and phenylephrine (PE) treatment involve increased activity and expression of calcineurin and decreased expression of miR-133 [133]. NATFc4, a member of calcineurin-activated NFAT family, has two functional binding sites for miR-133. Gain-of-function approaches show that miR-133 decreases NFAT mRNA levels as well as the hypertrophic response to PE-mediated stimulation in primary cardiomyocytes, and miR-133 loss-of-function increases NFATc4 expression and a hypertrophic response [123,134]. Moreover this microRNA decreases cardiac hypertrophy targeting RhoA, Cdc42 and Nelf-A/WHSC2 [110]. Various studies on this microRNA show that the over-expression of miR-133 attenuates agonist-induced hypertrophy [135]; conversely silencing of miR-133 makes the myocardium more sensitive to excessive cardiac growth [136]. Another member of NFAT family, NFATc3, is positively regulated by Myocardin, a transcriptional co-activator that promotes cardiac hypertrophy [137]. Under physiological conditions Myocardin is expressed at low levels but upon hypertrophic stimulation its expression is increased and consequently NFATc3 is up-regulated. Myocardin is a direct target of miR-9, so it was studied like a potential regulator of cardiac hypertrophy. Studies of miR-9 over-expression or inhibition, under hypertrophic stimulation, demonstrate that this microRNA negatively regulates cardiac hypertrophy, *in vivo* and *in vitro*, by targeting Myocardin [137]. Furthermore NFATC3 positively regulates miR-23a, which is up-regulated at transcriptional level by this factor. In fact, the expression of miR-23a is required to mediate hypertrophic growth in response to activation of the calcineurin/NFAT pathway; it directly targets an anti-hypertrophic protein: the muscle-specific ring finger protein 1 MuRF1 [134,138]. Another microRNA implicated in this process is miR-199b: it is a direct target of calcineurin-NFAT signaling, with an increased expression in heart failure. This microRNA modulates calcineurin-NFAT signaling-mediated hypertrophy in a positive feedback loop. Calcineurin induces miR-199b expression through a functional NFAT site upstream of the miR's gene; then the miR targets Dyrk1a, the dual-specificity tyrosine (Y) phosphorylation-regulated kinase 1a, in a process that constitutes a pathogenic feed-forward mechanism affecting calcineurin-responsive gene expression. Mice over-expressing miR-199b exhibit a strong hypertrophic phenotype induced by calcineurin/NFAT signaling, whereas inhibition of miR-199b normalizes Dyrk1a expression, reduces nuclear NFAT activity, and inhibits and even reverses the cardiac hypertrophy and fibrosis in mice models of heart failure [139].

5.2.4. miRNAs Regulated by Thyroid Hormone

miR-208a and miR-499 are implicated in the regulation of myosin gene expression and cardiac stress response. They play a redundant role: *in vivo* deletion of miR-208a resulting in viable animals with normal cardiac size at baseline, but these animals show a decline in cardiac function up to five months of age [140]. However miR-208a is under the thyroid hormone receptor TR, and its over-expression induces hypertrophic growth in mice by suppressing two negative regulators: thyroid hormone-associated protein-1 (Thap1) and myostatin [141]. Consistently elevated levels of miR-499 led to cardiomyopathy and cardiac hypertrophy in a dose-dependent manner [142]. miR-208a is

up-regulated in response to a hemodynamic pressure overload and in heart failure [141]. It is a positive regulator of β-MHC, required for the development of cardiac hypertrophy and myocardial fibrosis [140].

5.2.5. miRNAs Regulated by TGF-β in Heart Physiopathology

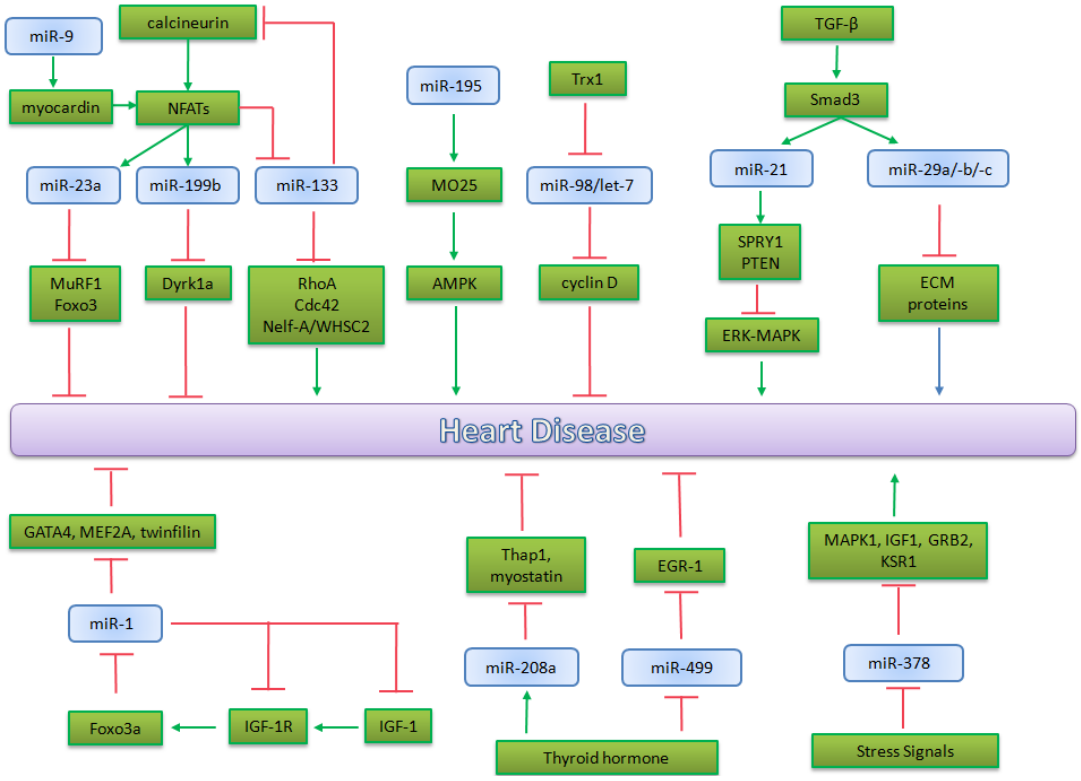
Fibroblasts are the most abundant class of non-cardiomyocyte cells in the heart; they produce ECM proteins as well as paracrine factors that can regulate the function of cardiomyocytes. In particular, in response to TGF-β, fibroblasts produce ECM and reduce collagenase production, leading to an excessive matrix accumulation. In this way, a key role is played by miRNAs regulated by TGF-β, a known agonist in the production and deposition of collagens in the heart, which contribute to cardiac hypertrophy. Numerous miRNAs are dysregulated in excessive fibrosis; including miR-29 and miR-21 [143]. The miR-29 family is composed of three members, 29a-b and -c, which are preferentially expressed in fibroblasts as compared with cardiomyocytes. All miR-29 family members target mRNA encoding multiple collagens, fibrillins and elastins and another multitude of ECM-related proteins involved in fibrosis. Interestingly, these microRNAs are down-regulated after TGF-β stimulation in cardiac fibroblasts, suggesting that they could contribute to TGF-β-induced fibrosis [129]. Another microRNA induced by TGF-β and dysregulated in fibroblasts, including in cardiac fibroblasts, in multiple types of stress, is miR-21 [144,145]. Studies on miR-21 reveal that this microRNA contributes to myocardial remodeling through regulation of ERK-MAPK-signaling, which is crucial in fibroblast survival and activation. The over-expression of miR-21 indirectly enhances the activity of ERK-MAP kinase; in fact it targets directly Sprouty-1 (SPRY1), a negative regulator of this pathway. In this manner miR-21 positively regulates cardiac fibroblast survival and growth factor secretion that eventually controls interstitial fibrosis and cardiac hypertrophy [143]. However there is a disagreement in the literature on the role of this microRNA: miR-21, in fact, induces the expression of matrix metalloproteinase-2 by targeting the phosphatase and tension homolog (PTEN) in fibroblasts [146]. Furthermore miR-21-null mice display fibrosis levels comparable to wild-type littermates, suggesting that this microRNA is not essential for pathological cardiac remodeling [147]. miR-21 has a potential role in the regulation of different mechanism for the decrease in cardiac contractile function in heart failure. Studies in tumor cells demonstrate that it targets the tropomyosin I (TPM1) [148], but it has not yet been studied in human heart failure.

5.2.6. The Role of miR-378 in MAPK Signaling

MAPK signaling is also controlled by miR-378. This microRNA is sufficient to repress cardiomyocyte hypertrophy regulating this pathway by targeting four members: MAPK1 itself, IGF-1, GRB2 and KSR1. Ganesan *et al*. studied miR-378 *in vivo* in a mouse model for chronic pressure overload (TAC); they demonstrate that the restoration of this microRNA in mice in which it was down-regulated, partially prevents cardiomyocyte hypertrophy and also does not trigger apoptosis *in vivo* [149]. So the tissue-specific up-regulation of miR-378 may be the basis of future therapeutic approaches to counteract cardiomyocyte hypertrophy. miRNAs are interesting targets for therapeutic use because they are selective. In fact they act on diseased tissues but they seem to have minimal effects on healthy tissues. Moreover, unlike their use in cardiovascular development, new delivery

methods have been found, catheter-based delivery systems which allow the injury site to be directly targeted, bypassing effects on other tissue [150]. Figure 3 summarizes the main microRNAs involved in heart pathophysiology.

Figure 3. miRNAs involved in Heart Pathophysiology. A schematic overview of the relevant microRNAs implicated in principal heart disease (*i.e.*, cardiac hypertrophy, heart failure and arrhythmias).



5.3. lncRNA in Heart Pathophysiology

miRNA are not the only ncRNAs implicated in regulatory processes. Recently another class of ncRNA, long non-coding RNA, so called because they are longer than 200 nt, has aroused much interest in cardiovascular function and disease (Table 2). lncRNAs are involved in cardiomyocyte differentiation, e.g., a new lncRNA (AK143260) was identified as a regulator of the cardiac lineage *in vitro*; it is required for mediating the transition from mesoderm to multipotent cardiac progenitors, regulating the activation of a network of cardiac differentiation specific genes [151]. Klattenhoff and colleagues named this lncRNA Braveheart because it is highly expressed in the heart, they identified Braveheart as a critical regulator of cardiovascular commitment from nascent mesoderm [15]. Recently, ANRIL, a multi-exonic lncRNA, has been shown to be implicated in epigenetic modulation in cardiac development and adult heart and also it has been associated with a locus implicated in cardiovascular disease [152,153]. Recently, it was shown that Fendrr, a lateral-mesoderm specific ncRNA, is fundamental in heart development in mouse. It mediates the epigenetic modification of target promoters thereby causing attenuation of the expression of transcription factors which are important in lateral mesoderm differentiation. Fendrr acts as chromatin modulator regulating PRC2 and TrxG/MLL, two histone-modifying complexes. PRC2 performs the methylation of histone H3 at

lysine 27 whereas TrxG/MLL catalyzes the methylation at lysine 4; both are essential for embryonic development. Generally, the action of transcription factors is restricted in time and space; they have effects only in the cell in which they are expressed and for a limited time. Conversely, the epigenetic pattern can persist for different stages during differentiation. Fendrr follows this trend and it has a long-term effect [154]. Moreover the regulation operated by ncRNA might confer susceptibility to various diseases: e.g., a myocardial infarction-associated transcript (MIAT), also known as RNCR2 or Gomafu, is a long intergenic non-coding RNA that presents many genetic variants implicated in different processes. A large scale case-control association study regarding cardiovascular disease demonstrates that a MIAT variant (rs2301523) confers susceptibility to myocardial infarction [155]. However it is still unclear how this lincRNA acts. Another abundant class of ncRNA is Natural Antisense Transcript (NATs); they can derive from protein or non protein coding genes. NATs exhibit typical mRNAs properties, such as 3' polyadenylation and 5' cap but present very little sequence conservation [156]. Some NATs essential for heart function are implicated in regulation of the cardiac troponin I (cTNI) and myosin heavy chains (MHCs) and light chains. Various NATS of these genes have been identified in human and rat and all these have the ability to form sense-antisense RNA duplexes at different ratios during a life span, suggesting a role in the regulation of gene expression. However, the exact mechanism of repression used by NATs is to date unknown (reviewed in Luther 2005) [157]. Korostowski *et al*. found Kcnq1ot1, a new long ncRNA, on studying imprinting in the developing heart. Surprisingly the study revealed that Kcnq1 and Kcnq1ot1 lose their imprinted expression at the same time in the heart. However, Kcnq1ot1 regulates Kcnq1 transcription, not by regulating its imprinting, but through modulating chromatin flexibility and access to enhancers [16]. The increasing knowledge of endogenous antisense RNA will help us to understand better the mechanism of gene expression regulation. In summary, given the ever-expanding number of non-coding RNAs, understanding their function represents a formidable task. They can specifically target different genes, often in a one-to-many manner. Fine tuning the level of single ncRNA may therefore affect many pathways in a pleiotropic manner. New therapeutic strategies face the major challenge of developing standardized methods that combine high transfection efficiency with targeted delivery of miRNA to act on specific pathways. The new technologies might provide the ability to translate laboratory potential into clinical practice to prevent or treat cardiovascular diseases.

Table 2. Long-ncRNA in Heart Development and Pathophysiology.

Long-ncRNA	Function	Reference
BraveHeart (AK143260)	Cardiomyocytes differentiation	[15,151]
ANRIL	Epigenetic modulation in cardiac development.	[152,153]
Fendrr	Heart development	[154]
MIAT	Susceptibility to myocardial infarction.	[155]
NATs	Regulation of gene expression	[157]
Kcnq1ot1	Regulation of gene expression	[16]

6. Conclusions

Non coding-RNAs critically affect the main molecular mechanisms involved in cardiovascular development and disease. It is well-established that dysregulation of miRNAs leads to development of

vascular as well as cardiac disorders. Moreover, the signaling pathways regulated by miRNAs and miRNAs themselves could be potential therapeutic targets in cardiovascular diseases. Nevertheless, further investigation is necessary to define how the cross-talk between microRNAs and their targets can affect different physiological and pathological pathways in the cardiovascular system. Recently, increasing interest has been aroused by lncRNAs in cardiovascular research. Despite rapid progress in our understanding of lncRNAs, data on their role in cardiovascular patophysiology is still poor. Given that long ncRNAs are associated with several cellular processes, an improved understanding of the functional roles of long non-coding RNA is needed to identify new therapeutic targets in cardiovascular diseases. For this purpose, it will be interesting to investigate the expression of lncRNA in cardiovascular tissues and subsequently to define the mechanisms in which these molecules are involved.

Acknowledgments

This study was partly supported by a grant from the Italian Ministry of Education, University and Research (MIUR): PON01_02833 "CARDIOTECH".

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Birney, E.; Stamatoyannopoulos, J.A.; Dutta, A.; Guigó, R.; Gingeras, T.R.; Margulies, E.H.; Weng, Z.; Snyder, M.; Dermitzakis, E.T.; Thurman, R.E.; *et al*. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **2007**, *447*, 799–816.
2. Alexander, R.P.; Fang, G.; Rozowsky, J.; Snyder, M.; Gerstein, M.B. Annotating non-coding regions of the genome. *Nat. Rev. Genet*. **2010**, *11*, 559–571.
3. He, L.; Hannon, G.J. MicroRNAs: Small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* **2004**, *5*, 522–531.
4. Mercer, T.R.; Dinger, M.E.; Mattick, J.S. Long non-coding RNAs insight into functions. *Nat. Rev. Genet*. **2009**, *10*, 155–159.
5. Mercer, T.R.; Mattick, J.S. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat. Struct. Mol. Biol*. **2013**, *20*, 300–307.
6. Esteller, M. Non-coding RNAs in human disease. *Nat. Rev. Genet.* **2011**, *12*, 861–874.
7. World Health Organization. *The Global Burden of Disease 2004 Update*. WHO: Geneva, Switzerland, 2008.
8. Small, E.M.; Olson, E.N. Pervasive roles of microRNAs in cardiovascular biology. *Nature* **2011**, *469*, 336–342.
9. Fiedler, J.; Thum, T. MicroRNAs in myocardial infarction. *Arterioscler. Thromb. Vasc. Biol*. **2013**, *33*, 201–205.
10. Holdt, L.M.; Teupser, D. Recent studies of the human chromosome 9p21 locus, which is associated with atherosclerosis in human populations. *Arterioscler. Thromb. Vasc. Biol.* **2012**, *32*, 196–206.

11. Ambros, V. The functions of animal microRNAs. *Nature* **2004**, *431*, 350–355.
12. Small, E.M.; Frost, R.J.; Olson, E.N. MicroRNAs add a new dimension to cardiovascular disease. *Circulation* **2010**, *121*, 1022–1032.
13. Batista, P.J.; Chang, H.Y. Long noncoding RNAs: Cellular address codes in development and disease. *Cell* **2013**, *152*, 1298–1307.
14. Hung, T.; Chang, H.Y. Long noncoding RNA in genome regulation: Prospects and mechanisms. *RNA Biol.* **2010**, *7*, 582–585.
15. Klattenhoff, C.A.; Scheuermann, J.C.; Surface, L.E.; Bradley, R.K.; Fields, P.A.; Steinhauser, M.L.; Ding, H.; Butty, V.L.; Torrey, L.; Haas, S.; *et al*. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell* **2013**, *152*, 570–583.
16. Korostowski, L.; Sedlak, N.; Engel, N. The Kcnq1ot1 long non-coding RNA affects chromatin conformation and expression of Kcnq1, but does not regulate its imprinting in the developing heart. *PLoS Genet.* **2012**, *8*, e1002956.
17. Chen, C.; Ridzon, D.A.; Broomer, A.J.; Zhou, Z.; Lee, D.H.; Nguyen, J.T.; Barbisin, M.; Xu, N.L.; Mahuvakar, V.R.; Andersen, M.R.; *et al*. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res*. **2005**, *33*, doi:10.1093/nar/gni178.
18. Ule, J.; Jensen, K.; Mele, A.; Darnell, R.B. CLIP: A method for identifying protein-RNA interaction sites in living cells. *Methods* **2005**, *37*, 376–386.
19. Zhao, J.; Ohsumi, T.K.; Kung, J.T.; Ogawa ,Y.; Grau, D.J.; Sarma, K.; Song, J.J.; Kingston, R.E.; Borowsky, M.; Lee, J.T. Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol. Cell* **2010**, *40*, 939–953.
20. Bittencourt, D.; Auboeuf, D. Analysis of co-transcriptional RNA processing by RNA-ChIP assay. *Methods Mol. Biol*. **2012**, *809*, 563–577.
21. Hu, W.; Yuan, B.; Flygare, J.; Lodish, H.F. Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. *Genes Dev*. **2011**, *25*, 2573–2578.
22. Michelhaugh, S.K.; Lipovich, L.; Blythe, J.; Jia, H.; Kapatos, G.; Bannon, M.J. Mining Affymetrix microarray data for long non-coding RNAs: Altered expression in the nucleus accumbens of heroin abusers. *J. Neurochem*. **2011**, *116*, 459–466.
23. Bentwich, I.; Avniel, A.; Karov, Y.; Aharonov, R.; Gilad, S.; Barad, O.; Barzilai, A.; Einat, P.; Einav, U.; Meiri, E.; *et al*. Identification of hundreds of conserved and nonconserved human microRNAs. *Nat. Genet*. **2005**, *37*, 766–770.
24. Rinn, J.L.; Kertesz, M.; Wang, J.K.; Squazzo, S.L.; Xu, X.; Brugmann, S.A.; Goodnough, L.H.; Helms, J.A.; Farnham, P.J.; Segal, E.; *et al*. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **2007**, *129*, 1311–1323.
25. Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods*. **2008**, *5*, 621–628.
26. Tripathi, V.; Ellis, J.D.; Shen, Z.; Song, D.Y.; Pan, Q.; Watt, A.T.; Freier, S.M.; Bennett, C.F.; Sharma, A.; Bubulya, P.A.; *et al*. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* **2010**, *39*, 925–938.
27. Yoon, J.H.; Abdelmohsen, K.; Srikantan, S.; Yang, X.; Martindale, J.L.; De, S.; Huarte, M.; Zhan, M.; Becker, K.G.; Gorospe, M. LincRNA-p21 suppresses target mRNA translation. *Mol. Cell* **2012**, *47*, 648–655.

28. Guttman, M.; Donaghey, J.; Carey, B.W.; Garber, M.; Grenier, J.K.; Munson, G.; Young, G.; Lucas, A.B.; Ach, R.; Bruhn, L.; *et al*. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **2011**, *477*, 295–300.
29. Pandey, R.R.; Mondal, T.; Mohammad, F.; Enroth, S.; Redrup, L.; Komorowski, J.; Nagano, T.; Mancini-Dinardo, D.; Kanduri, C. Kcnq1ot1 antisense noncodingRNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell* **2008**, *32*, 232–246.
30. Wamstad, J.A.; Alexander, J.M.; Truty, R.M.; Shrikumar, A.; Li, F.; Eilertson, K.E.; Ding, H.; Wylie, J.N.; Pico, A.R.; Capra, J.A.; *et al*. Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. *Cell* **2012**, *151*, 206–220.
31. Tsai, M.C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J.K.; Lan, F.; Shi, Y.; Segal, E.; Chang, H.Y. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **2010**, *329*, 689–693.
32. Hung, T.; Wang, Y.; Lin, M.F.; Koegel, A.K.; Kotake, Y.; Grant, G.D.; Horlings, H.M.; Shah, N.; Umbricht, C.; Wang, P.; *et al*. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat. Genet*. **2011**, *43*, 621–629.
33. Cesana, M.; Cacchiarelli, D.; Legnini, I.; Santini, T.; Sthandier, O.; Chinappi, M.; Tramontano, A.; Bozzoni, I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **2011**, *147*, 358–369.
34. Poliseno, L.; Salmena, L.; Zhang, J.; Carver, B.; Haveman, W.J.; Pandolfi, P.P. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **2010**, *465*, 1033–1038.
35. Memczak, S.; Jens, M.; Elefsinioti, A.; Torti, F.; Krueger, J.; Rybak, A.; Maier, L.; Mackowiak, S.D.; Gregersen, L.H.; Munschauer, M.; *et al*. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **2013**, *495*, 333–338.
36. Hansen, T.B.; Jensen, T.I.; Clausen, B.H.; Bramsen, J.B.; Finsen, B.; Damgaard, C.K.; Kjems, J. Natural RNA circles function as efficient microRNA sponges. *Nature* **2013**, *495*, 384–388.
37. Hansson, G.K. Inflammation, atherosclerosis, and coronary artery disease. *N. Engl. J. Med.* **2005**, *352*, 1685–1695.
38. Polimeni, A.; de Rosa, S.; Indolfi, C. Vascular miRNAs after balloon angioplasty. *Trends Cardiovasc. Med.* **2013**, *23*, 9–14.
39. Curcio, A.; Torella, D.; Indolfi, C. Mechanisms of smooth muscle cell proliferation and endothelial regeneration after vascular injury and stenting: Approach to therapy. *Circ. J.* **2011**, *75*, 1287–1296.
40. Kuehbacher, A.; Urbich, C.; Zeiher, A.M.; Dimmeler, S. Role of Dicer and Drosha for endothelial microRNA expression and angiogenesis. *Circ. Res.* **2007**, *101*, 59–68.
41. Suárez, Y.; Fernández-Hernando, C.; Pober, J.S.; Sessa, W.C. Dicer dependent microRNAs regulate gene expression and functions in human endothelial cells. *Circ. Res.* **2007**, *100*, 1164–1173.
42. Urbich, C.; Kuehbacher, A.; Dimmeler, S. Role of microRNAs in vascular diseases, inflammation, and angiogenesis. *Cardiovasc. Res.* **2008**, *79*, 581–588.
43. Fish, J.E.; Santoro, M.M.; Morton, S.U.; Yu, S.; Yeh, R.F.; Wythe, J.D.; Ivey, K.N.; Bruneau, B.G.; Stainier, D.Y.; Srivastava, D. miR-126 regulates angiogenic signaling and vascular integrity. *Dev. Cell*. **2008**, *15*, 272–284.

44. Harris, T.A.; Yamakuchi, M.; Ferlito, M.; Mendell, J.T.; Lowenstein, C.J. MicroRNA-126 regulates endothelial expression of vascular cell adhesion molecule 1. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1516–1521.
45. Bonauer, A.; Dimmeler, S. The microRNA-17–92 cluster: Still a miRacle? *Cell Cycle* **2009**, *8*, 3866–3873.
46. Bonauer, A.; Carmona, G.; Iwasaki, M.; Mione, M.; Koyanagi, M.; Fischer, A.; Burchfield, J.; Fox, H.; Doebele, C.; Ohtani, K.; *et al*. MicroRNA-92a controls angiogenesis and functional recovery of ischemic tissues in Mice. *Science* **2009**, *324*, 1710–1713.
47. Fang, Y.; Davies, P.F. Site-specific microRNA-92a regulation of Kruppel-like factors 4 and 2 in atherosusceptible endothelium. *Arterioscler. Thromb. Vasc. Biol.* **2012**, *32*, 979–987.
48. Hamik, A.; Lin, Z.; Kumar, A.; Balcells, M.; Sinha, S.; Katz, J.; Feinberg, M.W.; Gerzsten, R.E.; Edelman, E.R.; Jain, M.K. Kruppel-like factor 4 regulates endothelial inflammation. *J. Biol. Chem.* **2007**, *282*, 13769–13779.
49. SenBanerjee, S.; Lin, Z.; Atkins, G.B.; Greif, D.M.; Rao, R.M.; Kumar, A.; Feinberg, M.W.; Chen, Z.; Simon, D.I.; Luscinskas, F.W.; *et al*. KLF2 is a novel transcriptional regulator of endothelial proinflammatory activation. *J. Exp. Med.* **2004**, *199*, 1305–1315.
50. Lin, Z.; Kumar, A.; SenBanerjee, S.; Staniszewski, K.; Parmar, K.; Vaughan, D.E.; Gimbrone, M.A., Jr; Balasubramanian, V.; García-Cardeña, G.; Jain, M.K. KLF2 Is a novel transcriptional regulator of endothelial proinflammatory activation. *J. Exp. Med.* **2004**, *199*, 1305–1315.
51. Iaconetti, C.; Polimeni, A.; Sorrentino, S.; Sabatino, J.; Pironti, G.; Esposito, G.; Curcio, A.; Indolfi, C. Inhibition of mir-92a increases endothelial proliferation and migration *in vitro* as well as reduces neointimal proliferation *in vivo* after vascular injury. *Basic Res. Cardiol.* **2012**, *107*, doi:10.1007/s00395-012-0296-y.
52. Indolfi, C.; Torella, D.; Coppola, C.; Curcio, A.; Rodriguez, F.; Bilancio, A.; Leccia, A.; Arcucci, O.; Falco, M.; Leosco, D.; *et al*. Physical training increases eNOS vascular expression and activity and reduces restenosis after balloon angioplasty or arterial stenting in rats. *Circ. Res.* **2002**, *91*, 1190–1197.
53. Poliseno, L.; Tuccoli, A.; Mariani, L.; Evangelista, M.; Citti, L.; Woods, K.; Mercatanti, A.; Hammond, S.; Rainaldi, G. MicroRNAs modulate the angiogenic properties of HUVECs. *Blood* **2006**, *108*, 3068–3071.
54. Zhu, N.; Zhang, D.; Chen, S.; Liu, X.; Lin, L.; Huang, X.; Guo, Z.; Liu, J.; Wang, Y.; Yuan, W.; *et al*. Endothelial enriched microRNAs regulate angiotensin II-induced endothelial inflammation and migration. *Atherosclerosis* **2011**, *215*, 286–293.
55. Zhan, Y.; Brown, C.; Maynard, E.; Anshelevich, A.; Ni, W.; Ho, I.C.; Oettgen, P. Ets-1 is a critical regulator of Ang IImediated vascular flammation and remodeling. *J. Clin. Investig.* **2005**, *115*, 2508–2516.
56. Sun, H.X.; Zeng, D.Y.; Li, R.T.; Pang, R.P.; Yang, H.; Hu, Y.L.; Zhang, Q.; Jiang, Y.; Huang, L.Y.; Tang, Y.B.; *et al*. Essential role of microRNA-155 in regulating endothelium-dependent vasorelaxation by targeting endothelial nitric oxide synthase. *Hypertension* **2012**, *60*, 1407–1414.
57. Hansson, G.K.; Libby, P. The immune response in atherosclerosis: A double-edged sword. *Nat. Rev. Immunol*. **2006**, *6*, 508–519.

58. Gareus, R.; Kotsaki, E.; Xanthoulea, S.; van der Made, I.; Gijbels, M.J.; Kardakaris, R.; Polykratis, A.; Kollias, G.; de Winther, M.P.; Pasparakis, M. Endothelial cell-specific NF-kappaB inhibition protects mice from atherosclerosis. *Cell Metab.* **2008**, *8*, 372–383.
59. Sun, X,; Icli, B.; Wara, A.K.; Belkin, N.; He, S.; Kobzik, L.; Hunninghake, G.M.; Vera, M.P.; MICU Registry; Blackwell, T.S.; *et al*. MicroRNA-181b regulates NF-κB-mediated vascular inflammation. *J. Clin. Investig.* **2012**, *122*, 1973–1990.
60. Fang, Y.; Shi, C.; Manduchi, E.; Civelek, M.; Davies, P.F. MicroRNA-10a regulation of proinflammatory phenotype in athero-susceptible endothelium *in vivo* and *in vitro*. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 13450–13455.
61. Li, D.; Yang, P., Xiong, Q.; Song, X.; Yang, X.; Liu. L.; Yuan, W.; Rui, Y.C. MicroRNA-125a/b-5p inhibits endothelin-1 expression in vascular endothelial cells. *J. Hypertens.* **2010**, *28*, 1646–1654.
62. Ohkita, M.; Tawa, M.; Kitada, K.; Matsumura, Y. Pathophysiological roles of endothelin receptors in cardiovascular diseases. *J. Pharmacol. Sci.* **2012**, *119*, 302–313.
63. Alexander, M.R.; Owens, G.K. Epigenetic control of smooth muscle cell differentiation and phenotypic switching in vascular development and disease. *Annu. Rev. Physiol*. **2012**, *74*, 13–40.
64. Gomez, D.; Owens, G.K. Smooth muscle cell phenotypic switching in atherosclerosis. *Cardiovasc. Res.* **2012**, *95*, 156–164.
65. Indolfi, C.; Stabile, E.; Coppola, C.; Gallo, A.; Perrino, C.; Allevato, G.; Cavuto, L.; Torella, D.; Di Lorenzo, E.; Troncone, G.; *et al*. Membrane-bound protein kinase A inhibits smooth muscle cell proliferation *in vitro* and *in vivo* by amplifying cAMP-protein kinase A signals. *Circ. Res.* **2001**, *88*, 319–324.
66. Indolfi, C.; Avvedimento, E.V.; Rapacciuolo, A.; Di Lorenzo, E.; Esposito, G.; Stabile, E.; Feliciello, A.; Mele, E.; Giuliano, P.; Condorelli, G.; *et al*. Inhibition of cellular ras prevents smooth muscle cell proliferation after vascular injury *in vivo*. *Nat. Med.* **1995**, *1*, 541–545.
67. Cordes, K.R.; Sheehy, N.T.; White, M.P.; Berry, E.C.; Morton, S.U.; Muth, A.N.; Lee, T.H.; Miano, J.M.; Ivey, K.N.; Srivastava, D. miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature* **2009**, *460*, 705–710.
68. Cheng, Y.; Liu, X.; Yang, J.; Lin, Y.; Xu, D.Z.; Lu, Q.; Deitch, E.A.; Huo, Y.; Delphin, E.S.; Zhang, C. MicroRNA-145, a novel smooth muscle cell phenotypic marker and modulator, controls vascular neointimal lesion formation. *Circ. Res.* **2009**, *105*, 158–166.
69. Elia, L.; Quintavalle, M.; Zhang, J.; Contu, R.; Cossu, L.; Latronico, M.V.; Peterson, K.L.; Indolfi, C.; Catalucci, D.; Chen, J.; *et al*. The knockout of miR-143 and -145 alters smooth muscle cell maintenance and vascular homeostasis in mice: Correlates with human disease. *Cell Death Differ.* **2009**, *16*, 1590–1598.
70. Owens, G.K. Molecular control of vascular smooth muscle cell differentiation and phenotypic plasticity. *Novartis Found Symp.* **2007**, *283*, 174–191.
71. Quintavalle, M.; Elia, L.; Condorelli, G.; Courtneidge, S.A. MicroRNA control of podosome formation in vascular smooth muscle cells *in vivo* and *in vitro*. *J. Cell Biol.* **2010**, *189*, 13–22.
72. Lagna, G.; Ku, M.M.; Nguyen, P.H.; Neuman, N.A.; Davis, B.N.; Hata, A. Control of phenotypic plasticity of smooth muscle cells by bone morphogenetic protein signaling through the myocardin-related transcription factors. *J. Biol. Chem.* **2007**, *282*, 37244–37255.

73. Ten Dijke, P.; Arthur, H.M. Extracellular control of TGFβ signalling in vascular development and disease. *Nat. Rev. Mol. Cell Biol*. **2007**, *8*, 857–869.
74. King, K.E.; Iyemere, V.P.; Weissberg, P.L.; Shanahan, C.M. Krüppel-like factor 4 (KLF4/GKLF) is a target of bone morphogenetic proteins and transforming growth factor beta 1 in the regulation of vascular smooth muscle cell phenotype. *J. Biol. Chem.* **2003**, *278*, 11661–11669.
75. Davis-Dusenbery, B.N.; Chan, M.C.; Reno, K.E.; Weisman, A.S.; Layne, M.D.; Lagna, G.; Hata, A. Down-regulation of Kruppel-like factor-4 (KLF4) by microRNA-143/145 is critical for modulation of vascular smooth muscle cell phenotype by transforming growth factor-beta and bone morphogenetic protein 4. *J. Biol. Chem.* **2011**, *286*, 28097–28110.
76. Deaton, R.A.; Gan, Q.; Owens, G.K. Sp1-dependent activation of KLF4 is required for PDGF-BB-induced phenotypic modulation of smooth muscle. *Am. J. Physiol. Heart Circ. Physiol.* **2009**, *296*, 1027–1037.
77. Torella, D.; Iaconetti, C.; Catalucci, D.; Ellison, G.M.; Leone. A.; Waring, C.D.; Bochicchio, A.; Vicinanza, C.; Aquila, I.; Curcio, A.; *et al*. MicroRNA-133 controls vascular smooth muscle cell phenotypic switch *in vitro* and vascular remodeling *in vivo*. *Circ. Res.* **2011**, *109*, 880–893.
78. Ji, R.; Cheng, Y.; Yue, J.; Yang, J.; Liu, X.; Chen, H.; Dean, D.B.; Zhang, C. MicroRNA expression signa-ture and antisense-mediated depletion reveal an essential role of microRNA in vascular neointimal lesion formation. *Circ. Res*. **2007**, *100*, 1579–1588.
79. Liu, X.; Cheng, Y.; Zhang, S.; Lin, Y.; Yang, J.; Zhang, C. A necessary role of miR-221and miR-222 in vascular smooth muscle cell proliferation and neointimal hyperplasia. *Circ. Res.* **2009**, *104*, 476–487.
80. Sun, S.G.; Zheng, B.; Han, M.; Fang, X.M.; Li, H.X.; Miao, S.B.; Su, M.; Han, Y.; Shi, H.J.; Wen, J.K. miR-146a and Kruppel-like factor 4 form a feedback loop to participate in vascular smooth musclecell proliferation. *EMBO Reports* **2011**, *12*, 56–62.
81. Raitoharju, E.; Lyytikäinen, L.P.; Levula, M.; Oksala, N.; Mennander, A.; Tarkka, M.; Klopp, N.; Illig, T.; Kähönen, M.; Karhunen, P.J.; *et al*. miR-21, miR-210, miR-34a, and miR-146a/b are up-regulated in human atherosclerotic plaques in the Tampere Vascular Study. *Atherosclerosis* **2011**, *219*, 211–217.
82. Davis, B.N.; Hilyard, A.C.; Nguyen, P.H.; Lagna, G.; Hata, A. Induction of microRNA-221 by platelet-derived growth factor signaling is critical for modulation of vascular smooth muscle phenotype. *J. Biol. Chem*. **2009**, *284*, 3728–3738.
83. Zheng, B.; Han, M.; Wen, J.K. Role of Krüppel-like factor 4 in phenotypic switching and proliferation of vascular smooth muscle cells. *IUBMB Life* **2010**, *62*, 132–139.
84. Horita, H.N.; Simpson, P.A.; Ostriker, A.; Furgeson, S.; van Putten, V.; Weiser-Evans, M.C.; Nemenoff, R.A. Serum response factor regulates expression of phosphatase and tensin homolog through a microRNA network in vascular smooth muscle cells. *Arterioscler. Thromb. Vasc. Biol.* **2011**, *31*, 2909–2919.
85. Maegdefessel, L.; Azuma, J.; Toh, R.; Deng, A.; Merk, D.R.; Raiesdana, A.; Leeper, N.J.; Raaz, U.; Schoelmerich, A.M.; McConnell, M.V.; *et al*. MicroRNA-21 blocks abdominal aortic aneurysm development and nicotine-augmented expansion. *Sci. Transl. Med.* **2012**, *22*, doi:10.1126/scitranslmed.3003441.

86. Cheng, Y.; Liu, X.; Zhang, S.; Lin, Y.; Yang, J.; Zhang, C. MicroRNA-21 protects against the H(2)O(2)-induced injury on cardiac myocytes via its target gene PDCD4. *J. Mol. Cell. Cardiol.* **2009**, *47*, 5–14.
87. Wang, M.; Li, W.; Chang, G.Q.; Ye, C.S.; Ou, J.S.; Li, X.X.; Liu, Y.; Cheang, T.Y.; Huang, X.L.; Wang, S.M. MicroRNA-21 regulates vascular smooth muscle cell function via targeting tropomyosin 1 in arteriosclerosis obliterans of lower extremities. *Arterioscler. Thromb. Vasc. Biol*. **2011**, *31*, 2044–2053.
88. Merlet, E.; Atassi, F.; Motiani, R.K.; Mougenot, N.; Jacquet, A.; Nadaud, S.; Capiod, T.; Trebak, M.; Lompré, A.M.; Marchand, A. miR-424/322 regulates vascular smooth muscle cell phenotype and neointimal formation in the rat. *Cardiovasc. Res.* **2013**, *98*, 458–468.
89. Suárez, Y.; Fernández-Hernando, C.; Yu, J.; Gerber, S.A.; Harrison, K.D.; Pober, J.S.; Iruela-Arispe, M.L.; Merkenschlager, M.; Sessa, W.C. Dicer-dependent endothelial microRNAs are necessary for postnatal angiogenesis. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 14082–14087.
90. Fish, J.E.; Srivastava, D. MicroRNAs: Opening a new vein in angiogenesis research. *Sci. Signal*. **2009**, doi:10.1126/scisignal.252pe1.
91. Wang, S.; Aurora, A.B.; Johnson, B.A.; Qi, X.; McAnally, J.; Hill, J.A.; Richardson, J.A.; Bassel-Duby, R.; Olson, E.N. The endothelial-specific microRNA miR-126 governs vascular integrity and angiogenesis. *Dev. Cell* **2008**, *15*, 261–271.
92. Kane, N.M.; Meloni, M.; Spencer, H.L.; Craig, M.A.; Strehl, R.; Milligan, G.; Houslay, M.D.; Mountford, J.C.; Emanueli, C.; Baker, A.H. Derivation of endothelial cells from human embryonic stem cells by directed differentiation: Analysis of microRNA and angiogenesis *in vitro* and *in vivo*. *Arterioscler. Thromb. Vasc. Biol*. **2010**, *30*, 1389–1397.
93. Nicoli, S.; Knyphausen, C.P.; Zhu, L.J.; Lakshmanan, A.; Lawson, N.D. miR-221 is required for endothelial tip cell behaviors during vascular development. *Dev. Cell* **2012**, *22*, 418–429.
94. Albinsson, S.; Suarez, Y.; Skoura, A.; Offermanns, S.; Miano, J.M.; Sessa, W.C. MicroRNAs are necessary for vascular smooth muscle growth, differentiation, and function. *Arterioscler. Thromb. Vasc. Biol*. **2010**, *30*, 1118–1126.
95. Huang, H.; Xie, C.; Sun, X.; Ritchie, R.P.; Zhang, J.; Chen, Y.E. miR-10a contributes to retinoid acid-induced smooth muscle cell differentiation. *J. Biol. Chem*. **2010**, *285*, 9383–9389.
96. Xie, C.; Ritchie, R.P.; Huang, H.; Zhang, J.; Chen, Y.E. Smooth muscle cell differentiation: Models and underlying molecular mechanisms. *Arterioscler. Thromb. Vasc. Biol*. **2011**, *31*, 1485–1494.
97. Qiu, P.; Li, L. Histone acetylation and recruitment of serum responsive factor and CREB-binding protein onto SM22 promoter during SM22 gene expression. *Circ. Res.* **2002**, *90*, 858–865.
98. Xie, C.; Huang, H.; Sun, X.; Guo, Y.; Hamblin, M.; Ritchie, R.P.; Garcia-Barrio, M.T.; Zhang, J.; Chen, Y.E. MicroRNA-1 regulates smooth muscle cell differentiation by repressing Kruppel-like factor 4. *Stem Cells Dev.* **2011**, *20*, 205–210.
99. Davis, B.N.; Hilyard, A.C.; Lagna, G.; Hata, A. SMAD proteins control DROSHA-mediated microRNA maturation. *Nature* **2008**, *454*, 56–61.
100. Samani, N.J.; Erdmann, J.; Hall, A.S.; Hengstenberg, C.; Mangino, M.; Mayer, B.; Dixon, R.J.; Meitinger, T.; Braund, P.; Wichmann, H.E.; *et al*. WTCCC and the Cardiogenics Consortium. Genome wide association analysis of coronary artery disease. *N. Engl. J. Med.* **2007**, *357*, 443–453.

101. McPherson, R.; Pertsemlidis, A.; Kavaslar, N.; Stewart, A.; Roberts, R.; Cox, D.R.; Hinds, D.A.; Pennacchio, L.A.; Tybjaerg-Hansen, A.; Folsom, A.R.; *et al*. A common allele on chromosome 9 associated with coronary heart disease. *Science* **2007**, *316*, 1488–1491.
102. Broadbent, H.M.; Peden, J.F.; Lorkowski, S.; Goel, A.; Ongen, H.; Green, F.; Clarke, R.; Collins, R.; Franzosi, M.G.; Tognoni, G.; *et al*. PROCARDIS consortium. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. *Hum. Mol. Genet*. **2008**, *17*, 806–814.
103. Holdt, L.M.; Beutner, F.; Scholz, M.; Gielen, S.; Gäbel, G.; Bergert, H.; Schuler, G.; Thiery, J.; Teupser, D. ANRIL expression is associated with atherosclerosis risk at chromosome 9p21. *Arterioscler. Thromb. Vasc. Biol.* **2010**, *30*, 620–627.
104. Yap, K.L.; Li, S.; Muñoz-Cabello, A.M.; Raguz, S.; Zeng, L.; Mujtaba, S.; Gil, J.; Walsh, M.J.; Zhou, M.M. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* **2010**, *38*, 662–674.
105. Leung, A.; Trac, C.; Jin, W.; Lanting, L.; Akbany, A.; Sætrom, P.; Schones, D.E.; Natarajan, R. Novel long non-coding RNAs are regulated by angiotensin II in vascular smooth muscle cells. *Circ. Res.* **2013**, *113*, 266–278.
106. Fish, J.E.; Matouk, C.C.; Yeboah, E.; Bevan, S.C.; Khan, M.; Patil, K.; Ohh, M.; Marsden, P.A. Hypoxia-inducible expression of a natural *cis*-antisense transcript inhibits endothelial nitric-oxide synthase. *J. Biol. Chem.* **2007**, *282*, 15652–15666.
107. Robb, G.B.; Carson, A.R.; Tai, S.C.; Fish, J.E.; Singh, S.; Yamada, T.; Scherer, S.W.; Nakabayashi, K.; Marsden, P.A. Post-transcriptional regulation of endothelial nitric-oxide synthase by an overlapping antisense mRNA transcript. *J. Biol. Chem*. **2004**, *279*, 37982–37996.
108. Li, K.; Blum, Y.; Verma, A.; Liu, Z.; Pramanik, K.; Leigh, N.R.; Chun, C.Z.; Samant, G.V.; Zhao, B.; Garnaas, M.K.; *et al*. A noncoding antisense RNA in tie-1 locus regulates tie-1 function *in vivo*. *Blood* **2010**, *115*, 133–139.
109. Zhao, Y.; Samal, E.; Srivastava, D. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* **2005**, *436*, 214–220.
110. Carè, A.; Catalucci, D.; Felicetti, F.; Bonci, D.; Addario, A.; Gallo, P.; Bang, M.L.; Segnalini, P.; Gu, Y.; Dalton, N.D.; *et al*. MicroRNA-133 controls cardiac hypertrophy. *Nat. Med.* **2007**, *13*, 613–618.
111. Chen, J.F.; Mandel E.M.; Thomson J.M.; Wu. Q.; Callis, T.E.; Hammond, S.M.; Conlon, F.L.; Wang, D.Z. The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat. Genet.* **2006**, *38*, 228–233.
112. Liu, N.; Bezprozvannaya, S.; Williams, A.H.; Qi, X.; Richardson, J.A.; Bassel-Duby, R.; Olson, E.N. microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes Dev*. **2008**, *22*, 3242–3254.
113. Ikeda, S.; He, A.; Kong, S.W.; Lu, J.; Bejar, R.; Bodyak, N.; Lee, K.H.; Ma, Q.; Kang, P.M.; Golub, T.R.; *et al*. MicroRNA-1 negatively regulates expression of the hypertrophy-associated calmodulin and Mef2a genes. *Mol. Cell Biol*. **2009**, *29*, 2193–2204.
114. Van Rooij, E.; Quiat, D.; Johnson, B.A.; Sutherland, L.B.; Qi, X.; Richardson, J.A.; Kelm, R.J., Jr; Olson, E.N. A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. *Dev. Cell* **2009**, *17*, 662–673.

115. Gupta, M.P. Factors controlling cardiac myosin-isoform shift during hypertrophy and heart failure. *J. Mol. Cell. Cardiol*. **2007**, *43*, 388–403.
116. Morton, S.U.; Scherz, P.J.; Cordes, K.R.; Ivey, K.N.; Stainier, D.Y.; Srivastava, D. microRNA-138 modulates cardiac patterning during embryonic development. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17830–17835.
117. Taber, L.A. Biophysical mechanisms of cardiac looping. *Int. J. Dev. Biol.* **2006**, *50*, 323–332.
118. Deacon, D.C.; Nevis, K.R.; Cashman, T.J.; Zhou, Y.; Zhao, L.; Washko, D.; Guner-Ataman, B.; Burns, C.G.; Burns, C.E. The miR-143-adducin3 pathway is essential for cardiac chamber morphogenesis. *Development* **2010**, *137*, 1887–1896.
119. Lagendijk, A.K.; Goumans, M.J.; Burkhard, S.B.; Bakkers J. MicroRNA-23 restricts cardiac valve formation by inhibiting Has2 and extracellular hyaluronic acid production. *Circ. Res.* **2011**, *109*, 649–657.
120. Camenisch, T.D.; Spicer, A.P.; Brehm-Gibson, T.; Biesterfeldt, J.; Augustine, M.L.; Calabro, A., Jr.; Kubalak, S.; Klewer, S.E.; McDonald, J.A. Disruption of hyaluronan synthase-2 abrogates normal cardiac morphogenesis and hyaluronan-mediated transformation of epithelium to mesenchyme. *J. Clin. Investig.* **2000**, *106*, 349–360.
121. Stankunas, K.; Ma, G.K.; Kuhnert, F.J.; Kuo, C.J.; Chang, C.P. VEGF signaling has distinct spatiotemporal roles during heart valve development. *Dev. Biol.* **2010**, *347*, 325–336.
122. Kuhnert, F.; Mancuso, M.R.; Hampton, J.; Stankunas, K.; Asano, T.; Chen, C.Z.; Kuo, C.J. Attribution of vascular phenotypes of the murine Egfl7 locus to the microRNA miR-126. *Development* **2008**, *135*, 3989–3993.
123. Li, Q.; Lin, X.; Yang, X.; Chang, J. NFATc4 is negatively regulated in miR-133a-mediated cardiomyocyte hypertrophic repression. *Am. J. Physiol. Heart Circ. Physiol.* **2010**, *298*, H1340–H1347.
124. Rao, P.K.; Toyama, Y.; Chiang, H.R.; Gupta, S.; Bauer, M.; Medvid, R.; Reinhardt, F.; Liao, R.; Krieger, M.; Jaenisch, R.; *et al*. Loss of cardiac microRNA-mediated regulation leads to dilated cardiomyopathy and heart failure. *Circ. Res*. **2009**, *105*, 585–594.
125. Chen, H.; Untiveros, G.M.; McKee, L.A.; Perez, J.; Li, J.; Antin, P.B.; Konhilas, J.P. Micro-RNA-195 and -451 regulate the LKB1/AMPK signaling axis by targeting MO25. *PLoS One* **2012**, *7*, e41574.
126. Yang, Y.; Ago, T.; Zhai, P.; Abdellatif, M.; Sadoshima, J. Thioredoxin 1 negatively regulates angiotensin II-induced cardiac hypertrophy through upregulation of miR-98/let-7. *Circ. Res.* **2011**, *108*, 305–313.
127. Elia, L.; Contu, R.; Quintavalle, M.; Varrone, F.; Chimenti, C.; Russo, M.A.; Cimino, V.; de Marinis, L.; Frustaci, A.; Catalucci, D.; *et al*. Reciprocal regulation of microRNA-1 and insulin-like growth factor-1 signal transduction cascade in cardiac and skeletal muscle in physiological and pathological conditions. *Circulation* **2009**, *120*, 2377–2385.
128. Obata, K.; Nagata, K.; Iwase, M.; Odashima, M.; Nagasaka, T.; Izawa, H.; Murohara, T.; Yamada, Y.; Yokota, M. Overexpression of calmodulin induces cardiac hypertrophy by a calcineurin-dependent pathway. *Biochem. Biophys. Res. Commun*. **2005**, *338*, 1299–1305.
129. Li, Q.; Song, X.W.; Zou, J.; Wang, G.K.; Kremneva, E.; Li, X.Q.; Zhu, N.; Sun, T.; Lappalainen, P.; Yuan, W.J.; *et al*. Attenuation of microRNA-1 derepresses the cytoskeleton regulatory protein twinfilin-1 to provoke cardiac hypertrophy. *J. Cell Sci*. **2010**, *123*, 2444–2452.

130. Tang, Y.; Zheng, J.; Sun, Y.; Wu, Z.; Liu, Z.; Huang, G. MicroRNA-1 regulates cardiomyocyte apoptosis by targeting bcl-2. *Int. Heart J.* **2009**, *50*, 377–387.
131. Sayed, D.; Hong, C.; Chen, I.Y.; Lypowy, J.; Abdellatif, M. MicroRNA play an essential role in development of cardiac hypertrophy. *Circ. Res.* **2007**, *100*, 416–424.
132. Curcio, A.; Torella, D.; Iaconetti, C.; Pasceri, E.; Sabatino, J.; Sorrentino, S.; Giampà, S.; Micieli, M.; Polimeni, A.; Henning, B.J.; *et al*. MicroRNA-1 downregulation increases connexin 43 displacement and induces ventricular tachyarrhythmias in rodent hypertrophic hearts. *PLoS One* **2013**, *8*, e70158.
133. Dong, D.L.; Chen, C.; Huo, R.; Wang, N.; Li, Z.; Tu, Y.J.; Hu, J.T.; Chu, X.; Huang, W.; Yang, B.F. Reciprocal repression between microRNA-133 and calcineurin regulates cardiac hypertrophy: A novel mechanism for progressive cardiac hypertrophy. *Hypertension* **2010**, *55*, 946–952.
134. Van Rooij, E.; Sutherland, L.B.; Liu, N.; Williams, A.H.; McAnally, J.; Gerard, R.D.; Richardson, J.A.; Olson, E.N. A signature pattern of stress-responsive microRNAs that can evoke cardiac hypertrophy and heart failure. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 18255–18260.
135. Matkovich, S.J.; Wang, W.; Tu, Y.; Eschenbacher, W.H.; Dorn, L.E.; Condorelli, G.; Diwan, A.; Nerbonne, J.M.; Dorn, G.W., 2nd. MicroRNA-133a protects against myocardial fibrosis and modulates electrical repolarization without affecting hypertrophy in pressure-overloaded adult hearts. *Circ. Res.* **2010**, *106*, 166–175.
136. Krützfeldt, J.; Rajewsky, N.; Braich, R.; Rajeev, K.G.; Tuschl, T.; Manoharan, M.; Stoffel, M. Silencing of microRNAs *in vivo* with "antagomirs". *Nature***2005**, *438*, 685–689.
137. Wang, K.; Long, B.; Zhou, J.; Li, P.F. miR-9 and NFATc3 regulate myocardin in cardiac hypertrophy. *J. Biol. Chem.* **2010**, *285*, 11903–11912.
138. Lin, Z.; Murtaza, I.; Wang, K.; Jiao, J.; Gao, J.; Li, P.F. miR-23a functions downstream of NFATc3 to regulate cardiac hypertrophy. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12103–12108.
139. Da Costa Martins, P.A.; Salic, K.; Gladka, M.M.; Armand, A.S.; Leptidis, S.; El Azzouzi, H.; Hansen, A.; Coenen-de Roo, C.J.; Bierhuizen, M.F.; van der Nagel, R.; *et al*. MicroRNA-199b targets the nuclear kinase Dyrk1a in an auto-amplification loop promoting calcineurin/NFAT signalling. *Nat. Cell Biol*. **2010**, *12*, 1220–1227.
140. Van Rooij, E.; Sutherland, L.B.; Qi, X.; Richardson, J.A.; Hill, J.; Olson, E.N. Control of stress-dependent cardiac growth and gene expression by a microRNA. *Science* **2007**, *316*, 575–579.
141. Callis, T.E.; Pandya, K.; Seok, H.Y.; Tang, R.H.; Tatsuguchi, M.; Huang, Z.P.; Chen, J.F.; Deng, Z.; Gunn, B.; Shumate, J.; *et al*. MicroRNA-208a is a regulator of cardiac hypertrophy and conduction in mice. *J. Clin. Investig*. **2009**, *119*, 2772–2786.
142. Shieh, J.T.; Huang, Y.; Gilmore, J.; Srivastava, D. Elevated miR-499 levels blunt the cardiac stress response. *PLoS One* **2011**, *6*, e19481.
143. Van Rooij, E.; Sutherland, L.B.; Thatcher, J.E.; DiMaio, J.M.; Naseem, R.H.; Marshall, W.S.; Hill, J.A.; Olson, E.N. Dysregulation of microRNAs after myocardial infarction reveals a role of miR-29 in cardiac fibrosis. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 13027–13032.
144. Thum, T.; Gross, C.; Fiedler, J.; Fischer, T.; Kissler, S.; Bussen, M.; Galuppo, P.; Just, S.; Rottbauer, W.; Frantz, S.; *et al*. MicroRNA-21 contributes to myocardial disease by stimulating MAP kinase signalling in fibroblasts. *Nature* **2008**, *456*, 980–984.

145. Zavadil, J.; Narasimhan, M.; Blumenberg, M.; Schneider, R.J. Transforming growth factor-beta and microRNA:mRNA regulatory networks in epithelial plasticity. *Cells Tissues Organs* **2007**, *185*, 157–161.
146. Roy, S.; Khanna, S.; Hussain, S.R.; Biswas, S.; Azad, A.; Rink, C.; Gnyawali, S.; Shilo, S.; Nuovo, G.J.; Sen, C.K. MicroRNA expression in response to murine myocardial infarction: miR-21 regulates fibroblast metalloprotease-2 via phosphatase and tensin homologue. *Cardiovasc. Res.* **2009**, *82*, 21–29.
147. Patrick, D.M.; Montgomery, R.L.; Qi, X.; Obad, S.; Kauppinen, S.; Hill, J.A.; van Rooij, E.; Olson, E.N. Stress-dependent cardiac remodeling occurs in the absence of microRNA-21 in mice. *J. Clin. Investig*. **2010**, *120*, 3912–3916.
148. Zhu, S.M.; Si, M.L.; Wu, H.L.; Mo, Y.Y. MicroRNA-21 targets the tumor suppressor gene tropomyosin I (TPM1). *J. Biol. Chem.* **2007**, *282*, 14328–14336.
149. Ganesan, J.; Ramanujam, D.; Sassi, Y.; Ahles, A.; Jentzsch, C.; Werfel, S.; Leierseder, S.; Loyer, X.; Giacca, M.; Zentilin, L.; *et al*. MiR-378 controls cardiac hypertrophy by combined repression of mitogen-activated protein kinase pathway factors. *Circulation* **2013**, *127*, 2097–2106.
150. Van Rooij, E.; Marshall, W.S.; Olson, E.N. Toward microRNA-based therapeutics for heart disease: The sense in antisense. *Circ. Res.* **2008**, *103*, 919–928.
151. Schonrock, N.; Harvey, R.P.; Mattick, J.S. Long noncoding RNAs in cardiac development and pathophysiology. *Circ. Res.* **2012**, *111*, 1349–1362.
152. Congrains, A.; Kamide, K.; Oguro, R.; Yasuda, O.; Miyata, K.; Yamamoto, E.; Kawai, T.; Kusunoki, H.; Yamamoto, H.; Takeya, Y.; *et al*. Genetic variants at the 9p21 locus contribute to atherosclerosis through modulation of ANRIL and CDKN2A/B. *Atherosclerosis* **2012**, *220*, 449–455.
153. He, A.; Ma, Q.; Cao, J.; von Gise, A.; Zhou, P.; Xie, H.; Zhang, B.; Hsing, M.; Christodoulou, D.C.; Cahan, P.; *et al*. Polycomb repressive complex 2 regulates normal development of the mouse heart. *Circ. Res.* **2012**, *110*, 406–415.
154. Grote, P.; Wittler, L.; Hendrix, D.; Koch, F.; Währisch, S.; Beisaw, A.; Macura, K.; Bläss, G.; Kellis, M.; Werber, M.; *et al*. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev. Cell.* **2013**, *24*, 206–214.
155. Ishii, N.; Ozaki, K.; Sato, H.; Mizuno, H.; Saito, S.; Takahashi, A.; Miyamoto, Y.; Ikegawa, S.; Kamatani, N.; Hori, M.; *et al*. Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction. *J. Hum. Genet*. **2006**, *51*, 1087–1099.
156. Guttman, M.; Garber, M.; Levin, J.Z.; Donaghey, J.; Robinson, J.; Adiconis, X.; Fan, L.; Koziol, M.J.; Gnirke, A.; Nusbaum, C.; *et al*. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat. Biotechnol.* **2010**, *28*, 503–510.
157. Luther, H.P. Role of endogenous antisense RNA in cardiac gene regulation. *J. Mol. Med*. **2005**, *83*, 26–32.

Reprinted from *IJMS*. Cite as: Eskildsen, T.V.; Jeppesen, P.L.; Schneider, M.; Nossent, A.Y.; Sandberg, M.B.; Hansen, P.B.L.; Jensen, C.H.; Hansen, M.L.; Marcussen, N.; Rasmussen, L.M.; Bie, P.; Andersen, D.C.; Sheikh, S.P. Angiotensin II Regulates microRNA-132/-212 in Hypertensive Rats and Humans. *Int. J. Mol. Sci.* **2013**, *14*, 11190-11207.

Article

Angiotensin II Regulates microRNA-132/-212 in Hypertensive Rats and Humans

Tilde V. Eskildsen [1,2,†], Pia L. Jeppesen [1,2,†], Mikael Schneider [2], Anne Y. Nossent [1,2,3], Maria B. Sandberg [2], Pernille B. L. Hansen [1], Charlotte H. Jensen [1,2], Maria L. Hansen [2,4], Niels Marcussen [5], Lars M. Rasmussen [2], Peter Bie [1], Ditte C. Andersen [1,2] and Søren P. Sheikh [1,2,*]

[1] Department of Cardiovascular and Renal Research, Institute of Molecular Medicine, University of Southern Denmark, DK-5000 Odense, Denmark; E-Mails: teskildsen@health.sdu.dk (T.V.E.); piajeppesens@hotmail.com (P.L.J.); a.y.nossent@lumc.nl (A.Y.N.); plaerkegaard@health.sdu.dk (P.B.L.H.); charken@health.sdu.dk (C.H.J.); pbie@health.sdu.dk (P.B.); dandersen@health.sdu.dk (D.C.A.)

[2] Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Sdr. Boulevard 29, DK-5000 Odense, Denmark; E-Mails: mikaelschneider@gmail.com (M.S.); maria.sandberg@ouh.regionsyddanmark.dk (M.B.S.); mlhansen@health.sdu.dk (M.L.H.); lars.melholt.rasmussen@ouh.regionsyddanmark.dk (L.M.R.)

[3] Department of Vascular Surgery, Einthoven Laboratory for Experimental Vascular Medicine, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands

[4] Department of Cardiac and Thoracic Surgery, Odense University Hospital, Sdr. Boulevard 29, DK-5000 Odense, Denmark

[5] Department of Pathology, University of Southern Denmark, DK-5000 Odense, Denmark; E-Mail: nmarcussen@health.sdu.dk

[†] These authors contributed equally to this work.

[*] Author to whom correspondence should be addressed; E-Mail: soeren.sheikh@ouh.regionsyddanmark.dk; Tel.: +45-654-14-468; Fax: +45-654-11-911.

Received: 21 March 2013; in revised form: 25 April 2013 / Accepted: 15 May 2013 / Published: 27 May 2013

Abstract: MicroRNAs (miRNAs), a group of small non-coding RNAs that fine tune translation of multiple target mRNAs, are emerging as key regulators in cardiovascular development and disease. MiRNAs are involved in cardiac hypertrophy, heart failure and

remodeling following cardiac infarction; however, miRNAs involved in hypertension have not been thoroughly investigated. We have recently reported that specific miRNAs play an integral role in Angiotensin II receptor (AT_1R) signaling, especially after activation of the Gαq signaling pathway. Since AT_1R blockers are widely used to treat hypertension, we undertook a detailed analysis of potential miRNAs involved in Angiotensin II (AngII) mediated hypertension in rats and hypertensive patients, using miRNA microarray and qPCR analysis. The miR-132 and miR-212 are highly increased in the heart, aortic wall and kidney of rats with hypertension (159 ± 12 mm Hg) and cardiac hypertrophy following chronic AngII infusion. In addition, activation of the endothelin receptor, another Gαq coupled receptor, also increased miR-132 and miR-212. We sought to extend these observations using human samples by reasoning that AT_1R blockers may decrease miR-132 and miR-212. We analyzed tissue samples of mammary artery obtained from surplus arterial tissue after coronary bypass operations. Indeed, we found a decrease in expression levels of miR-132 and miR-212 in human arteries from bypass-operated patients treated with AT_1R blockers, whereas treatment with β-blockers had no effect. Taken together, these data suggest that miR-132 and miR-212 are involved in AngII induced hypertension, providing a new perspective in hypertensive disease mechanisms.

Keywords: hypertension; Angiotensin II; AT_1R; AT_1 receptor blocker; microRNA

1. Introduction

Persistent elevation of systemic blood pressure (hypertension) is one of the most prevalent medical conditions involving the cardiovascular system and affects as many as one billion people worldwide [1]. Hypertension is an undisputed risk factor for cardiovascular diseases, including stroke, cardiac failure and renal diseases [1]. Several mechanisms have been implicated in the pathogenesis of hypertension, including increased activity of the sympathetic nervous system, dysfunction of the vascular endothelium, vascular smooth muscle and cardiac hypertrophy, as well as overactivation of the renin-angiotensin-aldosterone system (RAAS) [2]. Angiotensin II (AngII) controls blood pressure and fluid homeostasis through its receptors, AT_1R and AT_2R, and through stimulation of aldosterone [1]. AngII receptors are expressed in tissues that have an impact on blood pressure control, including heart, kidney and vasculature [3,4]. The classical AngII responses in the cardiovascular and renal systems are mediated mainly by AT_1R signaling [3–5], including heterotrimeric G-protein activation and downstream signaling through the canonical MAP kinases ERK1/2, which, in turn, regulate gene transcription [4]. Accordingly, specific inhibitors of AngII pathways, including AT_1R blockers, dramatically lower blood pressure in hypertensive patients and slow the progression of cardiovascular disease [1,3].

We speculated that altered expression of microRNAs (miRNA) may be part of the pathogenesis behind AngII-related hypertension. MicroRNAs are small non-coding RNAs that regulate gene expression by pairing to and destabilizing the mRNAs of protein coding genes, resulting in decreased mRNA levels [6]. The mammalian miRNAs are highly conserved, and each miRNA is predicted to

target mRNAs of hundreds of distinct genes, fine-tuning and optimizing the expression patterns of most protein-coding genes [7]. Theoretically, these miRNAs are ideally suited to co-regulate gene expression events in cellular responses to vasopressors, such as AngII. Most miRNAs are solitary and expressed under the control of their own promoters and regulatory sequences, while others are arranged as clusters and may be co-regulated with additional members of the cluster [7]. For example, miR-132 and miR-212 are clustered closely in the genome and are transcribed together under the regulation of cAMP response element binding protein [8], which is a known AngII regulated gene [9,10]. Several miRNAs are aberrantly expressed in cardiovascular diseases [2,11–13]. miR-21, miR-155 and miR-221/222 have recently been shown to regulate AngII signaling in cardiac fibroblasts [14–16] and in endothelial cells [17], while miR-29 regulates fibrotic pathways [18]. We have recently shown that AngII, via the Gαq pathway, regulates five miRNAs during *in vitro* stimulation of primary cardiac fibroblasts and of HEK293N cells overexpressing the AT_1-receptor [19]. Most of the miRNA studies are based on *in vitro* experiments, and very few studies have examined the relation between AngII mediated hypertension and miRNA regulation *in vivo*.

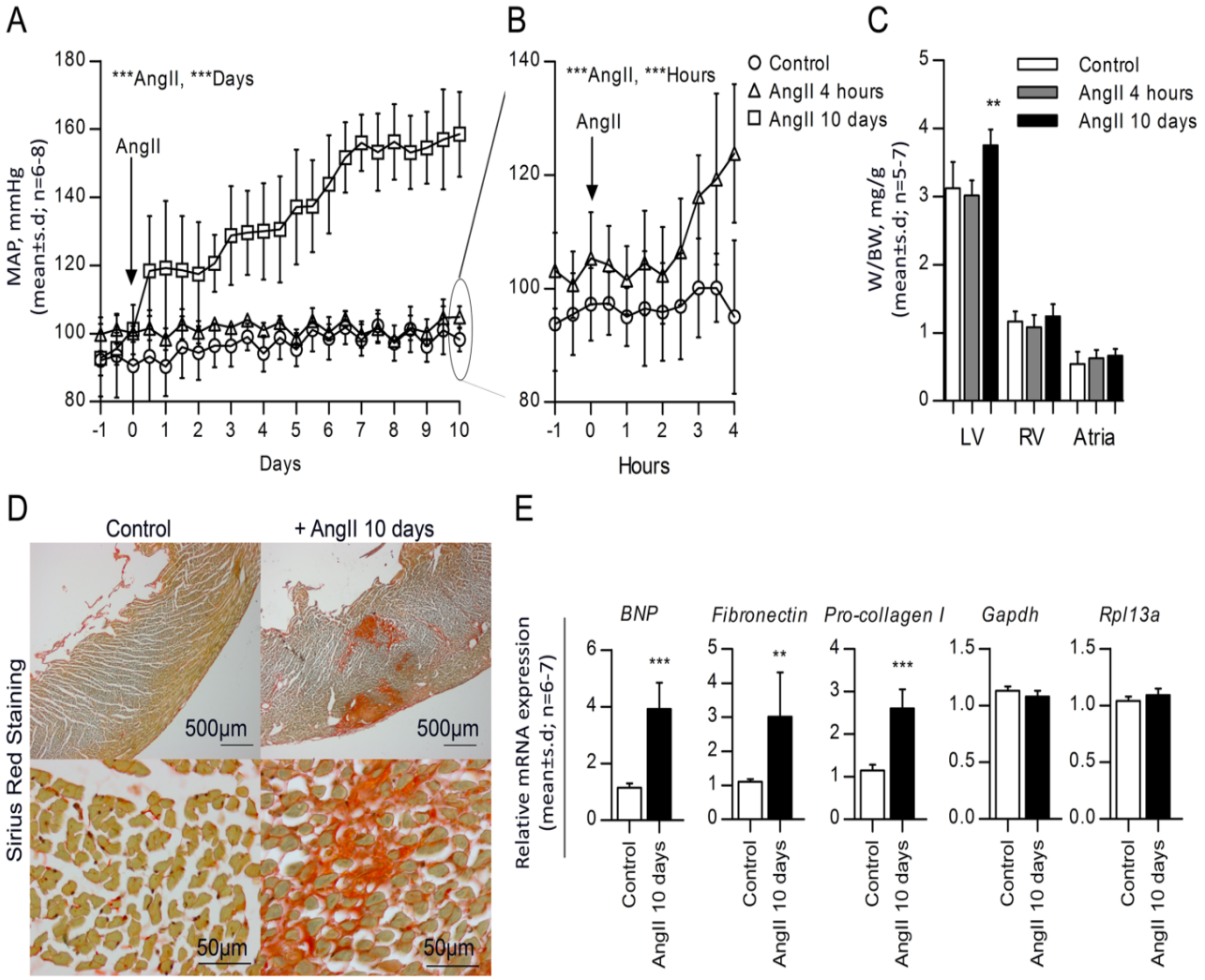
In this study, we hypothesized that *in vivo* AngII mimics the "five miRNA" expression signature obtained by AT_1R overexpression [19]. We examined the miRNA expression in heart, aorta and kidney from a rat model with a constant intravenous infusion of AngII and expanded these results to data concerning miRNA expression in the mammary artery of patients treated with AT_1R blockers. Our results suggest that miR-132 and miR-212 are involved in AngII-induced Gαq-signaling pathway leading to hypertension. Further understanding of the importance of these miRNAs will come from future miRNA knockdown experiments or knockout in whole animals.

2. Results

2.1. High Blood Pressure, Cardiac Hypertrophy and Fibrosis Are Sustained in the Rat Model

Infusion of AngII for 10 days resulted in a stable and significant elevation in blood pressure to 159 ± 12 mm Hg ($p < 0.001$, $n = 7$) at day 10, as compared to control rats that remained constant at 98 ± 4 mm Hg ($n = 8$) (Figure 1A). Likewise, we found that short time (4 h) AngII infusion resulted in an acute and significant 29 mm Hg increase in blood pressure ($p < 0.001$, $n = 6$) (Figure 1B). AngII hypertensive rats exhibited cardiac hypertrophy, as evidenced by a significant 17% increased left ventricle to body weight ratio ($p < 0.01$, $n = 7$) *versus* control rats ($n = 6$) (Figure 1C). The mass of the left ventricle increased from 614 ± 82 mg ($n = 6$) in control rats to 780 ± 75 mg ($n = 7$) in the hypertensive rats ($p < 0.01$), whereas no increase was observed for the right ventricle or atria weight (Figure 1C). Left ventricular hypertrophy was further validated by a significantly higher expression level of *B-type natriuretic peptide* (*BNP*) ($p < 0.001$) (Figure 1E). Likewise, cardiac fibrosis after infusion of AngII for 10 days was confirmed by increased collagen deposition (Figure 1D) and an increased expression of genes generally associated with fibrosis, including *Fibronectin* ($p < 0.01$) and *Procollagen I* ($p < 0.001$) (Figure 1E). These results thus showed that continuous AngII infusion for 10 days resulted in clear and sustained hypertension, leading to hypertrophic and fibrotic changes of the heart.

Figure 1. Characterization of AngII-induced hypertensive rat. (**A**) Mean daily averages of mean arterial blood pressure from seven rats treated with chronic infusion of 30 ng/kg/min AngII for 10 days (□) and six rats treated with acute infusion of 30 ng/kg/min AngII for 4 h (Δ), compared to eight control rats (O); (**B**) Mean hourly averages of mean arterial blood pressure from acute infusion of rats (Δ) for four h compared to control rats (O). Data are shown as the mean ± SD. Arrows shows the start of AngII infusion (day 0). Statistical significance was tested by two-way ANOVA for either control *versus* AngII for 10 days or control *versus* AngII for 4 h. *** $p < 0.001$. A and B, duplicate figure [20]; (**C**) Weight to body weight ratio (mg/g) of chronic ($n = 7$), acute ($n = 6$) and control ($n = 8$) rat hearts divided into left ventricle, right ventricle and atria. Data is presented as the mean ± SD, and statistical significance was tested by one-way ANOVA using Tukey's multiple comparison test. ** $p < 0.01$; (**D**) Representative sections of left ventricles of AngII affected hearts compared to control hearts, stained with Sirius Red for collagen deposition; (**E**) qRT-PCR for the early marker of hypertrophy, *BNP,* the fibrotic markers, *Fibronectin* and *Procollagen-I*, and the two stably expressed reference genes, *Gapdh and Rpl13a* (M: 0.140 and CV: 0.049). Data is presented as the mean ± SD, and statistical significance was tested by un-paired *t*-test. * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$.



2.2. Chronic AngII-Mediated Hypertension in Rats Increases miR-132/-212 Cluster Expression in Blood Pressure Regulating Organs: Heart, Aorta and Kidney

Table 1. miRNA microarray analysis showing significantly altered miRNA expression in the left ventricles from rats infused with AngII for 10 days compared to controls. Data is presented as $\log_2$ fold expression (n = 6–7) and sorted by p-value.

Microarray: miRNAs regulated by AngII		
Name of miRNA	**Log fold change**	***p*-value**
		*** $p < 0.001$
21	0.7016	1.89×10^{-5}
132	0.1261	5.90×10^{-5}
105	0.1883	9.28×10^{-5}
155	0.1425	0.00012
221	0.3414	0.00052
223	0.4459	0.00085
208b	0.5560	0.00089
		** $p < 0.01$
222	0.1681	0.0022
147b	0.1151	0.0032
26b	−0.1153	0.0034
15b	0.2580	0.0057
613	−0.1047	0.0065
31 *	0.1054	0.0075
520b	0.0938	0.0082
30c-1 *	−0.1409	0.0084
18b	0.1334	0.0092
		* $p < 0.05$
301a	0.1537	0.010
143	0.1242	0.011
434-5p	0.1692	0.012
484	0.0855	0.014
155	0.1135	0.014
379	0.0813	0.015
29c	−0.2193	0.017
936	0.1682	0.018
199a-5p	0.1523	0.021
201	−0.1204	0.021
101	−0.1860	0.021
363 *	0.2352	0.021
760	−0.0944	0.022
944	0.1056	0.023
200b *	0.1079	0.024
30b	−0.0888	0.025
322	−0.1890	0.026

Table 1. *Cont.*

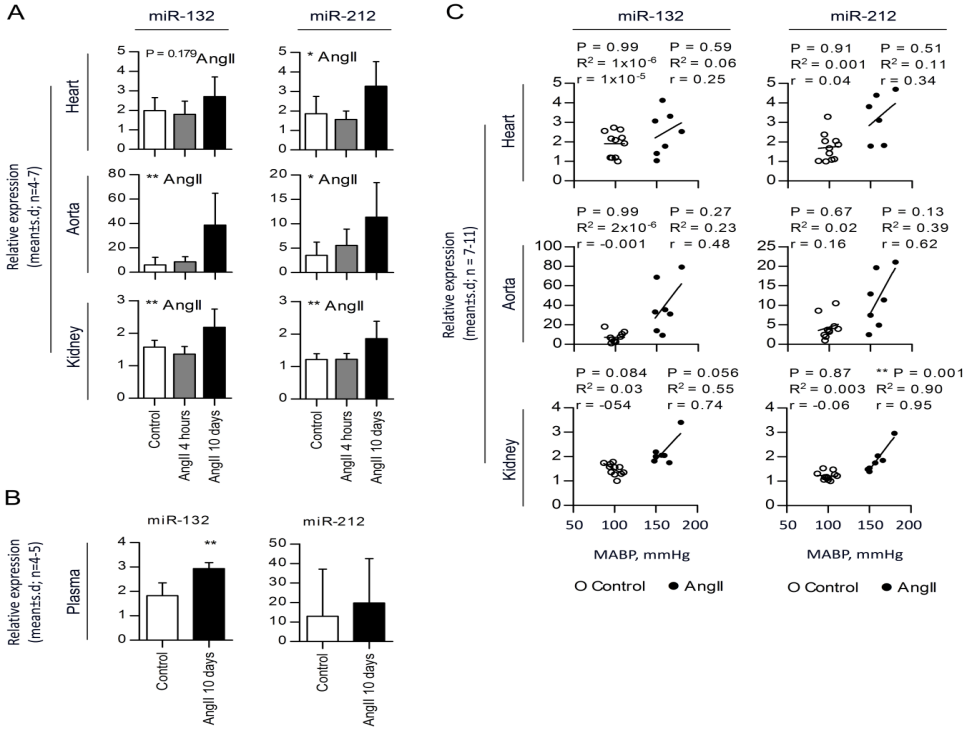
Microarray: miRNAs regulated by AngII		
Name of miRNA	**Log fold change**	***p*-value**
337-3p	0.1065	0.026
29c *	−0.1247	0.028
302c *	0.1453	0.030
193a-3p	0.1665	0.030
888	−0.0852	0.033
299-3p	0.2128	0.035
142-3p	0.3121	0.035
517 *	0.0746	0.041
31	0.3286	0.042
194	−0.0860	0.042
701	0.0671	0.042
545	−0.0834	0.043
609	0.0617	0.044
141 *	−0.0928	0.045
211	0.0588	0.047
373	0.0712	0.048
218-1 *	−0.0678	0.049

* means the passenger strand of the miRNA-miRNA* duplex; Passenger miRNA, may also have constitute a bioactive miRNA itself.

Fifty miRNAs were identified as differentially expressed ($p < 0.05$) in hearts of sustained hypertension (10 days of AngII), as compared to controls (Table 1), whereas no miRNAs were differentially expressed in acute hypertensive rats (4 h of AngII). AngII affects the blood pressure by two separate mechanisms: firstly, an acute contractile effect on the arterial walls arising within a few hours, followed by a chronic compensative response arising after a few days. Gene expression is primarily effected by the long-term AngII signaling. However, we sought to investigate if miRNAs could rapidly change, due to the acute effect, and found no regulation of miRNAs in the short period of AngII infusion. Of the 50 miRNAs differentially regulated in chronic hypertension, miR-21 was the most upregulated and was used as a positive control in our analysis. Interestingly, among the many dysregulated miRNAs, the second most significantly regulated miRNA was miR-132. Since the miR-132 gene is clustered with the miR-212 gene and they are likely expressed together [21], we included miR-212 in further analyses. In summary, we found that the expression of miR-21, -132 and -212 were up-regulated by 3.4-, 1.4- and 1.8-fold, respectively, in the left ventricle of AngII-induced hypertensive animals, as compared to controls (Figures 2A and S1). Besides the heart, the arterial wall and kidneys are involved in systemic blood pressure homeostasis, and we, therefore, examined whether the miR-132/-212 levels were affected also in these tissues. miR-132 and -212 were increased 6.4- and 3.2-fold, respectively, in aortas from hypertensive animals, as compared to controls (Figure 2A). Likewise, we observed a significant regulation of miR-132 (1.4-fold) and -212 (1.5-fold) in the kidneys of AngII-infused rats *versus* controls (Figure 2A). Furthermore, miR-132 was found to be significantly upregulated in the plasma of AngII-induced hypertensive animals, whereas no regulation was observed for plasma miR-212 levels compared to the control rats (Figure 2B). Lack of

miR-212 upregulation in the plasma of AngII-induced hypertensive rats compared to control rats might be caused by the low concentration and high variability of the miR-212 levels found in the plasma. Of note, the miRNA expression levels were found to significantly correlate with AngII-induced hypertension in the kidneys, a strong tendency of correlation in aorta and a lower tendency for correlation in the heart as tested by linear regression and correlation analysis (Figure 2C). By contrast, no changes in miR-132/-212 levels were found in any of the three tissues in the acutely hypertensive rats (Figure 2A). These data confirmed the microRNA array data obtained for chronic and acute hypertensive rats. Altogether, these results strongly suggest that the miR-132/-212 cluster may be a general and novel mediator of AngII-induced hypertension.

Figure 2. Validation of miRNA regulation in AngII-induced hypertensive rat heart, aorta and kidney. (**A**) qRT-PCR identification of miRNAs in left ventricle of hearts, aortas and kidneys from AngII affected rat hearts. Statistical significance was tested by one way ANOVA. * $p < 0.05$. Values are shown as relative expression with the mean ± SD; $n = 4–7$; (**B**) qRT-PCR identification of miRNAs in plasma of AngII affected rat hearts. Statistical significance was tested by un-paired *t*-test. ** $p < 0.01$. Values are shown as relative expression with the mean ± SD; $n = 4–5$. miRNA expression in the three organs and in plasma was individually normalized to two reference genes stably expressed among the samples. Reference genes used for normalization: Heart, miR-17 and miR-191 (M: 0.356, CV: 0.123); aorta, miR-103 and miR-191 (M: 0.832, CV: 0.290); kidney, miR-17 and miR-191 (M: 0.145, CV: 0.050); and plasma, miR-17 and miR-103 (M: 1.143, CV: 0.390); (**C**) Correlation analysis for miRNA expression levels at day 10 of AngII or isotone glucose infusion. Statistical significance was tested by linear regression (R^2) and correlation analysis (Pearson's *r*). Data are shown as two individual correlations per miRNA is each organ for control (○) and AngII for 10 days (●). ** $p < 0.01$; $n = 7–11$.



2.3. miR-132 and -212 Regulation in Response to AngII in Mice

Studying mice in this particular field of hypertension research is of interest, because their genome can easily be modified. Infusion of AngII at 15, 30 and 60 ng/kg/min for seven days resulted in a stable and significant elevation in blood pressure to 119 ± 10, 127 ± 5 and 128 ± 10 mm Hg ($p < 0.001$, $n = 3$), respectively, at day seven, as compared to control mice that remained constant at 95 ± 2 mm Hg ($n = 4$) (Figure 3A). AngII-infused mice did not exhibit cardiac hypertrophy when examining the left ventricle to body weight ratio ($n = 3$) *versus* control mice ($n = 4$) (Figure 3B). Likewise, miR-132 and miR-212 were not significantly increased in AngII-infused mice at any of the AngII concentrations when compared to control mice. These results thus showed that continuous AngII infusion for seven days in mice resulted in a modest increase in blood pressure without hypertrophic changes of the heart and no regulation of the miR-132/-212 cluster.

Figure 3. Validation of miRNA regulation in AngII-induced hypertensive mice hearts. (**A**) Mean daily averages of mean arterial blood pressure from three mice treated with chronic infusion of 60, 30 or 15 ng/kg/min AngII for seven days (□, Δ or ◊, respectively), compared to three control mice (O). Data are shown as the mean ± SD. AngII infusion is started at day 0. Statistical significance was tested by two-way ANOVA. *** $p < 0.001$, *$p < 0.05$; (**B**) Left ventricle (LV) weight to body weight ratio (mg/g) of mice infused with 60, 30 or 15 ng/kg/min of AngII for seven days ($n = 3$) and control ($n = 3$) mice. Data is presented as the mean ± SD, and statistical significance was tested by one-way ANOVA using Tukey's multiple comparison test; (**C**) qRT-PCR identification of miRNAs in left ventricle of hearts from AngII affected mice hearts. Statistical significance was tested by one-way ANOVA using Tukey's multiple comparison test. Values are shown as relative expression with the mean ± SD; $n = 3$. Reference genes used for normalization: miR-103 and miR-191 (M: 0.826, CV: 0.288).

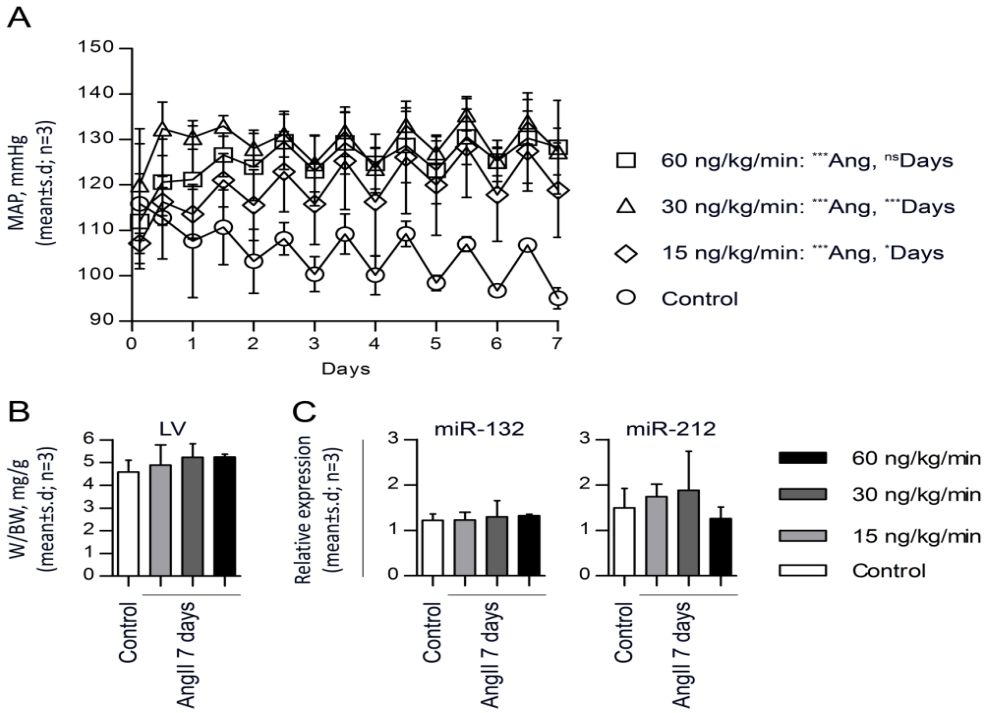
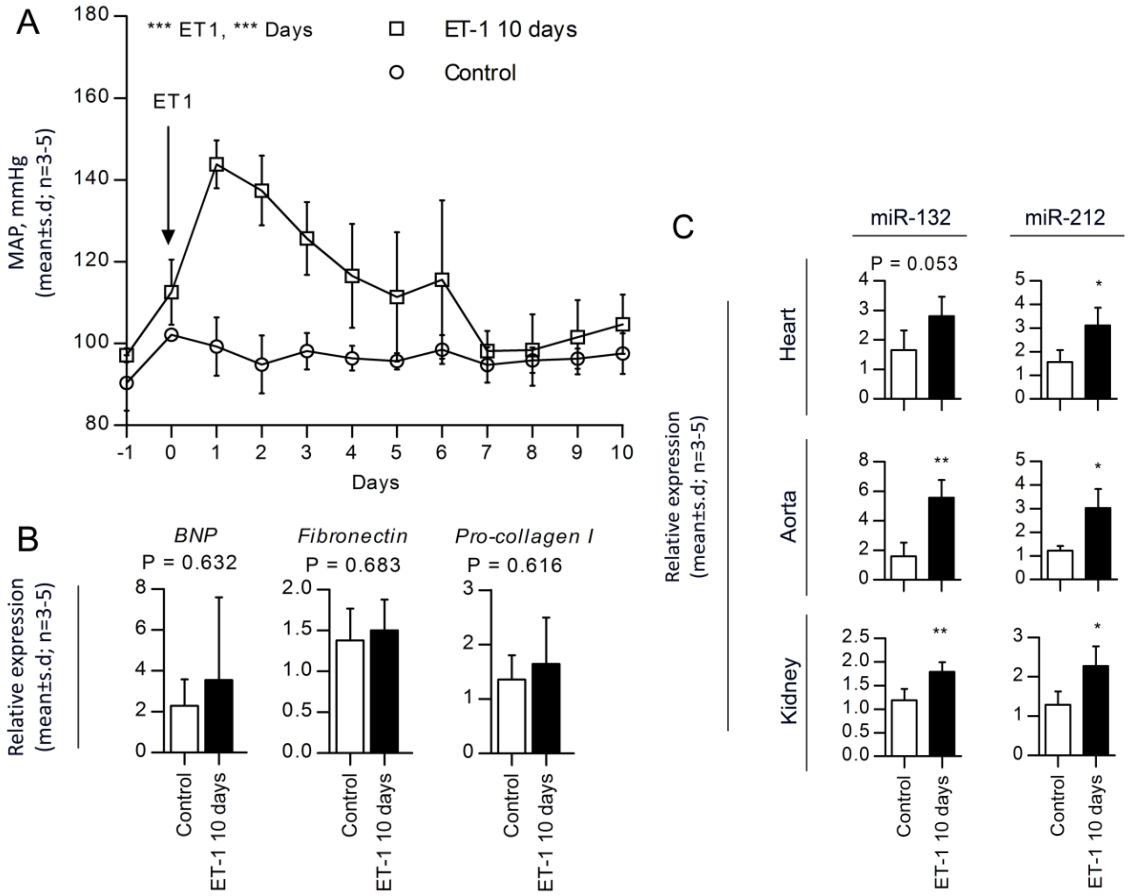


Figure 4. Endothelin 1-induced hypertensive rats. (**A**) Mean daily averages of mean arterial blood pressure from five rats treated with chronic infusion of 5 pmol/kg/min ET-1 for 10 days (□), compared to three control rats (O). Statistical significance was tested by two-way ANOVA. *** $p < 0.001$; (**B**) qRT-PCR for the early hypertrophy marker, *BNP* and the fibrosis markers, *Fibronectin* and *Procollagen-I*. Statistical significance was tested by un-paired *t*-test. Data is shown as the mean ± SD, $n = 3–5$. Reference genes used for normalization: *GAPDH* and *Rpl13a* (M: 0.153, CV: 0.053); (**C**) qRT-PCR identification of miRNAs in the left ventricle of hearts, aortas and kidneys from ET-1 affected rat hearts. Statistical significance was tested by un-paired *t*-test. * $p < 0.05$. Values are shown as relative expression with the mean ± SD, $n = 3–5$. miRNA expression was individually normalized to two reference genes stably expressed among the samples. Reference genes used for normalization: heart; miR-17 and miR-191 (M: 0.309, CV: 0.107); aorta, miR-103 and miR-191 (M: 0.667, CV: 0.232); kidney, miR-17 and miR-191 (M: 0.180, CV: 0.062).



2.4. miR-132 and -212 Regulations in Response to ET-1, Vasopressor-Induced Hypertension

In order to test whether the miR-132/-212 cluster response is specific for AT_1R induction, we examined the effect of continuous infusion of a second vasopressor (ET-1) [22]. In contrast to the sustained high blood pressure in AngII-infused rats, ET-1-mediated hypertension peaked at day one (144 ± 6 mm Hg ($n = 5$) and, hereafter, gradually declined, reaching 105 ± 7 mm Hg at day 10 (Figure 4A). In line with this, we did not observe any cardiac hypertrophy or fibrosis in ET-1 treated hearts. Thus, we found no change in *BNP*, *Fibronectin* or *Procollagen-I* expression levels (Figure 4B).

However, despite the partially compensated blood pressure and the absence of cardiac hypertrophy and fibrosis, we found significantly higher levels of miR-132 and miR-212 in heart (1.7- and 2.0-fold), aorta (3.5- and 2.5-fold) and kidney (1.5- and 1.8-fold) of the ET-1 treated animals as compared to controls (Figure 4C). In line with the previous findings by Jeppesen *et al.* [19], this suggests that the upregulation of miR-132/-212 is a direct consequence of Gαq vasopressor stimulation pathways and not a result of secondary heart hypertrophy and fibrosis. In summary, these data indicate that upregulation of the miR-132/-212 cluster likely is part of a general response to Gαq-vasopressor stimulation of the ERK1/2 pathway and may be involved in a common AngII- and ET-1-mediated signaling pathway leading to hypertension.

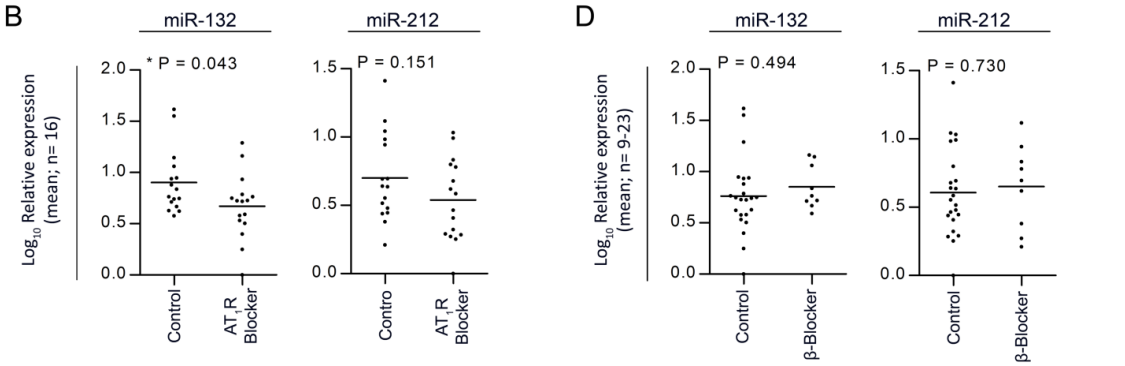
Figure 5. AngII receptor blockers and β-blockers in internal mammary artery (IMA) patients. (**A**,**C**) Study group information on the AngII receptor blocker (ARB) and β-blocker patient groups, respectively. None of the characteristics within the patient groups of ARBs and β-blockers were significantly different. The groups were matched for age, sex, diabetes and treatment using statins. None of the patients were simultaneously treated with angiotensin-converting enzyme (ACE)-inhibitors; (**B**) Scatter plot represents Log_{10} relative miRNA expression with mean bar, in patients treated with ARBs compared to patients treated with non-ARB (n = 16). (**D**) Scatter plot representing Log_{10} relative miRNA expression with mean bar, in patients treated with β-blocker (n = 9) compared to patients treated with non-β-blockers (n = 23). Expression is normalized to two stably expressed reference genes. Reference genes used for normalization: miR-103 and miR-191 (M: 0.804, CV: 0.279).

A

Internal Mammary Artery	Control	AT_1R blocker
n	16	16
Sex, females/males, n	3/13	3/13
Age, years	71.2 (6.2)	71.5 (7.7)
Systolic blood pressure, mmHg	139.9 (9.7)	141.3 (24.5)
Diastolic blood pressure, mmHg	75.5 (10.6)	77.3 (14.4)
β-blockers	11	12
Diuretics	6	9

C

Internal Mammary Artery	Control	β-blocker
n	23	9
Sex, females/males, n	4/19	2/7
Age, years	70.1 (7.3)	73.7 (4.9)
Systolic blood pressure, mmHg	142.9 (18.6)	133.6 (12.1)
Diastolic blood pressure, mmHg	77.8 (12.3)	72.0 (11.8)
ARB	12	4
Diuretics	11	4



2.5. Treatment with Angiotensin II Receptor Blocker Attenuates the Expression of the miR-132/-212 Cluster in Human Hypertension

We measured miR-132/-212 levels in the internal mammary artery (IMA) tissue obtained from patients treated with AngII receptor blocker (ARB) or β-blocker in age and sex matched patients undergoing by-pass surgery (Figure 5A,C). The four patient groups showed similar pre-operative

blood pressure (Figure 5A,C). ARB treated patients (n = 16) revealed a significant attenuation of miR-132 expression (0.55-fold), as well as a tendency for miR-212 downregulation (0.64-fold), as compared to non-ARB-treated patients (miR-132; 0.93 and miR-212; 1.01) (n = 16) (Figure 5B). To further investigate whether this miR-132 and -212 regulation is specific to AngII or related to putative direct influences from blood pressure, we examined patients treated with β-blockers (Figure 5C). Interestingly, we did not find any attenuation or downregulation of miR-132/-212 expression in IMA patients receiving β-blockers, as compared to non-β-blocker-treated patients (Figure 5D), indicating that not all blood pressure-reducing agents can downregulate miR-132/-212 expression, which further supports the notion that AngII mediates a global upregulation of miR-132/-212 in humans. These results suggest that the miR-132/-212 cluster in humans may also be part of the response to Gαq-vasopressors, such as AngII.

3. Discussion

We found increased expression of miR-132 and -212 in the left ventricle, aorta and kidney, as well as in the plasma (Figure 2) after 10 days of sustained AngII-induced hypertension in rats, which is compatible with our pervious published *in vitro* study [19]. Even though miR-132 and miR-212 are expressed from the same precursor, we observed independent regulation in the different tissues in response to the same AngII infusion. This could be due to differences in stability and degradation [21,23]. Moreover, the degree of miR-132/-212 increase shows a tendency to correlate with blood pressure suggesting that these miRNAs could play a novel role in AT1 receptor pharmacology, both *in vitro* and *in vivo*.

Blood pressure has previously been associated with miRNA. Recently, Nossent *et al.* investigated the association between single-nucleotide polymorphisms (SNPs) located in the miRNA binding sites of genes associated with Renin-AngiotensinII-system (RAS), blood pressure and myocardial infarction in a large population study [24]. Several SNPs located in RAS genes were associated with changes in blood pressure and were shown to interfere with miRNA regulation [24].

miR-132/-212 has been described in both the central nervous and cardiovascular systems. In one study, miR-132 was reported to be constitutively expressed and released by pericyte progenitor cells, and transplantation of these cells into mice with myocardial infarction improves cardiac function through proangiogenic activities [25]. Several studies identify miR-132/-212 involvement in the central nervous system, *i.e.*, in neuronal function and plasticity [8,21,26]. In addition, miR-132/-212 has also been found to be involved in neovascularization, inflammation and adipocyte differentiation in the peripheral tissues [21,23,27]. Interestingly, miR-132/-212 was upregulated in the aortas of mice stimulated with AngII from osmotic pumps for 14 days, but this study did not report blood pressure values [28]. We performed similar studies with chronic AngII infusion in mice for seven days, simultaneously measuring blood pressure. In contrast to our rat infusion model, only a modest blood pressure increase was observed in this mice strain, and miR-132/-212 was not significantly upregulated in mouse hearts (Figure 3). Differences in AngII effects between rats and mice have previously been described by Cassis *et al.*, stating that AngII infusion in mice has minimal effects compared to same doses given to rats [29]. Based on these data, we decided that the rat model was more suitable for studying AngII-induced miRNA changes in hypertensive animals. We previously demonstrated that

AT_1R signaling regulates miR-132 and -212 in HEK293N cells and in primary cultures of cardiac fibroblasts through the Gαq dependent pathway [19]. These results were recently confirmed in primary cultures of rat vascular smooth muscle cells [28]. Additionally, we investigated the regulation of the miR-132/-212 cluster in endothelial cell lines and in primary cultures of vascular smooth muscle cells (VSMCs) and leukocytes and found no regulation in either of the cell types (data not shown). Furthermore, by inhibition of the Gαq subunit in cardiac fibroblasts, we demonstrated significant decreases in miR-132 and -212 expressions, pointing to Gαq protein activation as the responsible pathway for AngII-induced miRNA regulation *in vitro* [19].

Using a different vasopressor (ET-1) that binds to the endothelin receptor and also activates Gαq-ERK1/2 signaling [30,31], we examined whether this receptor could regulate the miRNAs. Interestingly, even though the ET-1-induced hypertension had a much shorter duration than the sustained hypertension induced by AngII, both miR-132 and miR-212 were upregulated at a point in time when blood pressure was not (Figure 4). From a mechanistic point of view, these findings indicate that miR-132/-212, also *in vivo*, may be regulated through Gαq-ERK1/2 activation, which is one of the mutual steps in the AngII and ET-1 signaling pathways leading to hypertension [30–32]. Following these observations, we examined whether the increased miRNA expression levels could be attenuated by blocking the AngII signaling in humans (Figure 5). We examined the arterial walls from two different groups of patients treated with either one of several ARBs, including Losartan, Candesartan, Irbesartan and Telmisartan, or no ARB. These ARBs are widely used in hypertension treatment and have similar chemical structures, but different pharmacological properties and efficacy [33]. None of the patients were treated with ACE-inhibitors. Our patient groups were selected based on treatment with ARBs and individual ARB-types grouped and analyzed together in order to allow for statistical testing. Most patients in the ARB group probably suffered from hypertension; however, it has not been possible from the patient files to deduce who strictly fulfills the definitions of hypertension. Likewise, in the non-ARB group, the precise prevalence of hypertension is not known. Another limitation of our human observations is the lack of knowledge about the exact time point of medication, which could not be controlled. Despite these limitations, we observed a significant downregulation of miR-132, as well as a robust attenuation of miR-212 in the ARB-treated patients. We next asked whether treatment with β-blockers, often a first choice antihypertensive drug, would also decrease miR-132 and -212, following the notion that it could be reduction in blood pressure *per se*, which may alter the miRNA levels. However, the levels of miR-132/-212 were not downregulated in patients treated with β-blockers. This observation is compatible with the notion that alterations in miRNA is not only a spurious factor found in rodents or related to experimental systems *in vitro*, but that these molecules also could be suggested to play a role in the arterial wall among patients receiving medication that alters the activity in the AngII system. Our data indicates that AT_1R control of these miRNAs is evolutionary conserved between rat and man. Thus, AT_1R activation in rats increases miR-132 and miR-212, while blocking the AT_1R decreases miRNA levels in humans. Since no anti-miR experiments have been conducted, it has not been possible to deduce the specific cause and effect relationship; however, the regulation of miR-132 and miR-212 is likely biological important, because although rats and humans share biological features in blood pressure control, they have multiple differences in the molecular subtypes of ion channels, receptors and signaling pathways in blood vessel cells. Further studies are necessary to assess the relative biological and pharmacological

impact of individual miR-132 and miR-212 levels on systemic blood pressure, in the heart, arterial wall and kidney. We speculate that the AT_1R could perform these effects through Gαq and downstream activation of the ERK1/2 pathway and not through blood pressure or aldosterone. Three lines of evidence support this notion. Firstly, the effect was inhibited by the pharmacological blockade of AngII *in vivo* and Gαq and ERK1/2 *in vitro* [19]. Secondly, another Gαq activator (ET-1) reproduced the effects, while, finally, inhibiting Gs signaling with the β-adrenergic blocker had no effects on the miRNA levels.

4. Experimental Section

4.1. Animal Care

Female Sprague-Dawley rats (8 weeks old, Taconic, Ry, Denmark) were housed in air-conditioned rooms with a 12 h dark-light cycle and fed standard diet (Altromin® Standard 1320, Lage, Germany) with free access to tap water. All rat experiments were approved by the Danish National Animal Experiment Inspectorate (Permission #2009/561/1753).

4.2. Angiotensin II (AngII) Model

Under Hypnorm/midazolam anesthesia, chronic catheters were placed in the left femoral vein and artery and connected to a swivel via a skin button between the scapulae, allowing the rat full mobility [34]. Following a 5–7 day recovery period, the arterial line was connected to a pressure transducer (Föhr Medical Instruments, Hessen, Germany), and data were collected continuously using Lab View software (National Instruments, Austin, TX, USA). The rats were infused with 5% glucose solution at a rate of 10 μL/kg/min with or without AngII (Sigma-Aldrich, St. Louis, MO, USA) for 10 days. Rats were randomly divided into three groups: (1) 10 days' continuous infusion of AngII at 30 ng/kg/min (sustained hypertension); (2) 9 days' and 20 h infusion with 5% glucose, followed by a 4 h period of AngII infusion at 30 ng/kg/min (acute hypertension); and (3) 10 days' infusion of 5% glucose (controls). Rats were sacrificed with an overdose of pentobarbital, then perfused with ice-cold sterile isotonic saline via the heart. The heart, kidney and aorta were dissected under sterile conditions. Hearts were carefully sectioned into atria, right ventricle and left ventricle, whereas aortas were cleaned from fat, as well as connective tissue, and kidneys were decapsulated.

4.3. Endothelin 1 (ET-1) Model

The setup was similar to that described above for the AngII model, but rats were randomly divided into two groups: (1) 10 days' infusion of endothelin-1 (ET-1) (Bachem) at a concentration of 5.0 pmol/kg/min in isotonic saline; and (2) 10 days' infusion with isotonic saline, as previously described by Mortensen L.H. & Fink G.D. 1992 [35].

4.4. Mice AngII Model

The mice experiments were approved by the Danish National Animal Experiment Inspectorate (Permission #2009/561/1749). The setup was similar to that described above for the AngII rat model

with minor modifications as described below. Mixed gender C57/BL6 mice were anaesthetized using ketamine/xylazine and catheters placed in the femoral artery and vein for measurements of arterial blood pressure and AngII infusions, respectively. The mice recovered for five days before beginning the continuous measurements of mean arterial blood pressure and heart rate for seven days. Mice were randomly divided into four groups: (1–3) 7 days' continuous infusion of AngII at 15, 30 and 60 ng/kg/min respectively and (4) 7 days' of continuous infusion with 5% glucose (controls). Mice were sacrificed and processed, as described above.

4.5. Patients, Internal Mammary Artery Study

Surplus arterial tissue was obtained from the repair vessel, *i.e.*, the internal mammary artery (IMA), from patients undergoing coronary artery by-pass graft surgery at the Department of Cardiac and Thoracic Surgery, Odense University Hospital, Denmark. Four groups of patients were analyzed, based on intake or no intake of either AngII receptor blocker or β-blocker. Study population characteristics are described in (Figure 5A,C). As previously described [36], the intima/media of the artery was carefully dissected immediately after removal, frozen in liquid nitrogen and stored in the Odense Artery Biobank. The study was approved by the Ethical Review Board of Southern Denmark (No. S-20100044), and informed consent was obtained from each subject.

4.6. Sirius Red Staining

Fibrosis was identified by Sirius Red Staining for collagen deposits. Cryosections were equilibrated at room temperature for 30 min, rinsed in a mixture of NBF (37%) and ethanol (93%) for 45 s, rehydrated (5 min) and subsequently counterstained (15 min) with Weigert's Iron Hematoxylin (Sigma-Aldrich, HT-107/109). Following a wash, sections were further incubated (60 min) with 0.1% Sirius red (Sigma-Aldrich, P6744) in saturated picric acid (Sigma-Aldrich, 365548) and finally washed twice in 99% ethanol. Sections were mounted with Pertex and analyzed by bright-field microscopy using a Leica DML332 equipped with a Leica DFC300F camera.

4.7. Microarray and Data Processing

The miRNA expression profiling was performed as two-color common reference hybridizations on LNA-based arrays (miRCURY LNA™ microRNA Array ready to spot probe set, Exiqon, Denmark) spotted in-house on CodeLink™ HD Activated slides (DHD1-0023, SurModics, Eden Prairie, MN, USA), according to the manufacturer's recommendation. Samples were labeled with Hy3, and the common reference (pool of all samples) was labeled with Hy5, by use of miRCURY LNA microRNA Array Power labeling kit (208032-A, Exiqon) and hybridized for 16 h. Slides were washed (208021, Exiqon), scanned on an Agilent (G2565CA) Microarray scanner and analyzed by the Genepix 6.0 software. Normalization and background correction was performed with "R" software using the "vsn" package (Bioconductor, open source software), and quadruplicate spots were averaged. Differential expression was assayed using the "limma" package (Bioconductor) by fitting the eBayes linear model and contrasting individual treatments with untreated controls. Log_2 fold changes were calculated using the toptable function of the limma package.

4.8. mRNA and miRNA Analysis

Relative qRT-PCR of mRNA and miRNA were performed, as previously described [37,38]. Briefly, total RNA was extracted using the TriReagent protocol (Molecular Research Center, Inc., Cincinnati, Ohio), and RNA purity, integrity and quantity were examined by nanodrop (Nanodrop® Technologies, Wilmington, DE, USA) and Bioanalyzer (Agilent 2100, Santa Clara, California) measurements. Relative quantitative mRNA PCR was performed on reverse transcribed cDNA (High Capacity cDNA RT kit; Applied Biosystems, Foster City, CA, USA). Primer sequences, amplification efficiencies (AE) and standard errors (SE): *BNP* forward: 5'-GATGCAGAAGCTGGAGA-3', reverse: 5'-TCTGCTGGACCCGGAGGGTG-3', AE: 2.046 and SE: 0.011; *Fibronectin* forward: 5'-CAGGGGTCACGTACCTCTTCAAAG-3', reverse: 5'-CGAGGTGGAGTCCAAGTTACCAGA-3', AE: 1.995 and SE: 0.010; *Procollagen-I* forward: 5'-ATCGTGGCTTCTCTGGTCTCCAG-3', reverse: 5'-CAGGGAGACCGTTGAGTCCATCT-3', AE: 2.039, SE: 0.004; *GAPDH* forward: 5'-GTCGGTGTGAACGGATTTGGC-3', reverse: 5'-TGAAGGGGTCGTTGATGGCA-3', AE: 2.079, SE: 0.008; *RPL13A* forward: 5'-GAAAGCGGATGAACACCAACCC-3', reverse: 5'-GGGATCCCATCCAACACCTTGA-3', AE: 1.958, SE: 0.003.

For miRNA qRT-PCR primers specific for rat and human miR-21 (#000397), miR-132 (#000457), miR-212 (#002551), miR-17 (#002308), miR-103 (#000439), miR-191 (#002299) and let-7f (#000382) were purchased from Applied Biosystems. Amplification and detection were performed using 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). As recommended by others [39,40] and previously described [37,38], we used the qBase$^+$ software to normalize all qRT-PCR data against multiple stably expressed control genes (information applied in respective figure legends).

4.9. Statistical Analysis

Results are represented as the mean ± SD. All analyses comprised independent experiments and two-way ANOVA, and two-tailed student's *t*-tests were performed, as indicated [GraphPad Prism (5.0 version) software, La Jolla, CA, USA] to test significant levels. In the animal experiments, data on the effect of AngII and ET-1 were obtained from 6–8 and 3–5 animals, respectively. Data not following Gaussian distribution, as tested by D'Agostino's normality test, were subjected to a Log_{10} transformation before statistical analysis and normality evidenced to follow Gaussian distribution prior to significance tests. Differences were considered to be significant at $p < 0.05$.

5. Conclusions

In conclusion, we found that miR-132 and -212 are increased in AngII-induced hypertension *in vivo*, in organs associated with blood pressure control, which partly mimics the "five miRNA" expression signature obtained by AT_1R overexpression [19]. Cleary, the *in vivo* model adds to our understanding, as we can narrow down the miRNA changes after AngII, to miR-132 and -212. Importantly, our results show that functional inhibition of the AT_1R reversed the miR-132 and -212 expression levels, demonstrating that AngII is responsible for the regulation, possibly via the Gαq-dependent pathway.

Acknowledgments

We thank Lene Bundgaard Andersen, Marlene Louise Christensen, Tonja Lyngse Jørgensen and Vivi Monrad for excellent technical assistance. SPS is funded by the John and Birthe Meyer Foundation, LMR is funded by the Novo Nordisk foundation and TVE and PLJ are funded by The Danish Cardiovascular Academy and the Danish heart foundation.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Coffman, T.M. Under pressure: The search for the essential mechanisms of hypertension. *Nat. Med*. **2011**, *17*, 1402–1409.
2. Batkai, S.; Thum, T. MicroRNAs in hypertension: Mechanisms and therapeutic targets. *Curr. Hypertens. Rep*. **2012**, *14*, 79–87.
3. Crowley, S.D.; Gurley, S.B.; Coffman, T.M. AT(1) receptors and control of blood pressure: The kidney and more. *Trends Cardiovasc. Med*. **2007**, *17*, 30–34.
4. Bader, M.; Ganten, D. Update on tissue renin-angiotensin systems. *J. Mol. Med. (Berl.)* **2008**, *86*, 615–621.
5. Hansen, J.L.; Aplin, M.; Hansen, J.T.; Christensen, G.L.; Bonde, M.M.; Schneider, M.; Haunso, S.; Schiffer, H.H.; Burstein, E.S.; Weiner, D.M.; *et al*. The human angiotensin AT(1) receptor supports G protein-independent extracellular signal-regulated kinase 1/2 activation and cellular proliferation. *Eur. J. Pharmacol*. **2008**, *590*, 255–263.
6. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297.
7. Ambros, V. The functions of animal microRNAs. *Nature* **2004**, *431*, 350–355.
8. Vo, N.; Klein, M.E.; Varlamova, O.; Keller, D.M.; Yamamoto, T.; Goodman, R.H.; Impey, S. A cAMP-response element binding protein-induced microRNA regulates neuronal morphogenesis. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16426–16431.
9. Neyses, L.; Nouskas, J.; Luyken, J.; Fronhoffs, S.; Oberdorf, S.; Pfeifer, U.; Williams, R.S.; Sukhatme, V.P.; Vetter, H. Induction of immediate-early genes by angiotensin II and endothelin-1 in adult rat cardiomyocytes. *J. Hypertens*. **1993**, *11*, 927–934.
10. Christensen, G.L.; Kelstrup, C.D.; Lyngso, C.; Sarwar, U.; Bogebo, R.; Sheikh, S.P.; Gammeltoft, S.; Olsen, J.V.; Hansen, J.L. Quantitative phosphoproteomics dissection of seven-transmembrane receptor signaling using full and biased agonists. *Mol. Cell Proteomics* **2010**, *9*, 1540–1553.
11. Olson, E.N. Gene regulatory networks in the evolution and development of the heart. *Science* **2006**, *313*, 1922–1927.
12. Van Rooij, E.; Sutherland, L.B.; Qi, X.; Richardson, J.A.; Hill, J.; Olson, E.N. Control of stress-dependent cardiac growth and gene expression by a microRNA. *Science* **2007**, *316*, 575–579.
13. Care, A.; Catalucci, D.; Felicetti, F.; Bonci, D.; Addario, A.; Gallo, P.; Bang, M.L.; Segnalini, P.; Gu, Y.; Dalton, N.D.; *et al*. MicroRNA-133 controls cardiac hypertrophy. *Nat. Med*. **2007**, *13*, 613–618.

14. Thum, T.; Gross, C.; Fiedler, J.; Fischer, T.; Kissler, S.; Bussen, M.; Galuppo, P.; Just, S.; Rottbauer, W.; Frantz, S.; *et al*. MicroRNA-21 contributes to myocardial disease by stimulating MAP kinase signalling in fibroblasts. *Nature* **2008**, *456*, 980–984.
15. Van Rooij, E.; Marshall, W.S.; Olson, E.N. Toward microRNA-based therapeutics for heart disease: the sense in antisense. *Circ. Res*. **2008**, *103*, 919–928.
16. Martin, M.M.; Lee, E.J.; Buckenberger, J.A.; Schmittgen, T.D.; Elton, T.S. MicroRNA-155 regulates human angiotensin II type 1 receptor expression in fibroblasts. *J. Biol. Chem*. **2006**, *281*, 18277–18284.
17. Zhu, N.; Zhang, D.; Chen, S.; Liu, X.; Lin, L.; Huang, X.; Guo, Z.; Liu, J.; Wang, Y.; Yuan, W.; *et al*. Endothelial enriched microRNAs regulate angiotensin II-induced endothelial inflammation and migration. *Atherosclerosis* **2011**, *215*, 286–293.
18. Van Rooij, E.; Sutherland, L.B.; Thatcher, J.E.; DiMaio, J.M.; Naseem, R.H.; Marshall, W.S.; Hill, J.A.; Olson, E.N. Dysregulation of microRNAs after myocardial infarction reveals a role of miR-29 in cardiac fibrosis. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 13027–13032.
19. Jeppesen, P.L.; Christensen, G.L.; Schneider, M.; Nossent, A.Y.; Jensen, H.B.; Andersen, D.C.; Eskildsen, T.; Gammeltoft, S.; Hansen, J.L.; Sheikh, S.P. Angiotensin II type 1 receptor signalling regulates microRNA differentially in cardiac fibroblasts and myocytes. *Br. J. Pharmacol*. **2011**, *164*, 394–404.
20. Nossent, A.Y.; Eskildsen, T.V.; Andersen, L.B.; Bie, P.; Brønnum, H.; Schneider, M.; Andersen, D.C.; Welten, S.M.J.; Jeppesen, P.L.; Hamming, J.F.; *et al*. The 14q32 MicroRNA-487b targets the anti-apoptotic insulin receptor substrate 1 in hypertension induced remodeling of the aorta. *Ann. Surg.* **2013**, in press.
21. Tognini, P.; Pizzorusso, T. MicroRNA212/132 family: Molecular transducer of neuronal function and plasticity. *Int. J. Biochem. Cell Biol.* **2012**, *44*, 6–10.
22. Mortensen, L.H.; Pawloski, C.M.; Kanagy, N.L.; Fink, G.D. Chronic hypertension produced by infusion of endothelin in rats. *Hypertension* **1990**, *15*, 729–733.
23. Anand, S.; Majeti, B.K.; Acevedo, L.M.; Murphy, E.A.; Mukthavaram, R.; Scheppke, L.; Huang, M.; Shields, D.J.; Lindquist, J.N.; Lapinski, P.E.; *et al*. MicroRNA-132-mediated loss of p120RasGAP activates the endothelium to facilitate pathological angiogenesis. *Nat. Med.* **2010**, *16*, 909–914.
24. Nossent, A.Y.; Hansen, J.L.; Doggen, C.; Quax, P.H.; Sheikh, S.P.; Rosendaal, F.R. SNPs in microRNA binding sites in 3'-UTRs of RAAS genes influence arterial blood pressure and risk of myocardial infarction. *Am. J. Hypertens*. **2011**, *24*, 999–1006.
25. Katare, R.; Riu, F.; Mitchell, K.; Gubernator, M.; Campagnolo, P.; Cui, Y.; Fortunato, O.; Avolio, E.; Cesselli, D.; Beltrami, A.P.; *et al*. Transplantation of human pericyte progenitor cells improves the repair of infarcted heart through activation of an angiogenic program involving micro-RNA-132. *Circ. Res.* **2011**, *109*, 894–906.
26. Nudelman, A.S.; DiRocco, D.P.; Lambert, T.J.; Garelick, M.G.; Le, J.; Nathanson, N.M.; Storm, D.R. Neuronal activity rapidly induces transcription of the CREB-regulated microRNA-132, *in vivo*. *Hippocampus* **2010**, *20*, 492–498.
27. Strum, J.C.; Johnson, J.H.; Ward, J.; Xie, H.; Feild, J.; Hester, A.; Alford, A.; Waters, K.M. MicroRNA 132 regulates nutritional stress-induced chemokine production through repression of SirT1. *Mol. Endocrinol*. **2009**, *23*, 1876–1884.

28. Jin, W.; Reddy, M.A.; Chen, Z.; Putta, S.; Lanting, L.; Kato, M.; Park, J.T.; Chandra, M.; Wang, C.; Tangirala, R.; *et al*. Small RNA sequencing reveals microRNAs that modulate angiotensin II effects in vascular smooth muscle cells. *J. Biol. Chem.* **2012**, *287*, 15672–15683.
29. Cassis, L.A.; Huang, J.; Gong, M.C.; Daugherty, A. Role of metabolism and receptor responsiveness in the attenuated responses to Angiotensin II in mice compared to rats. *Regul. Pept*. **2004**, *117*, 107–116.
30. Rautureau, Y.; Schiffrin, E.L. Endothelin in hypertension: An update. *Curr. Opin. Nephrol. Hypertens*. **2012**, *21*, 128–136.
31. Gray, G.A.; Webb, D.J. *Molecular Biology and Pharmacology of the Endothelins*; Springer & R.G. Landes, Co.: New York, NY, USA, 1995; pp. 116–145.
32. Roberts, R.E. The extracellular signal-regulated kinase (ERK) pathway: A potential therapeutic target in hypertension. *J. Exp. Pharmacol.* **2012**, *4*, 77–83.
33. Zaman, M.A.; Oparil, S.; Calhoun, D.A. Drugs targeting the renin-angiotensin-aldosterone system. *Nat. Rev. Drug Discov*. **2002**, *1*, 621–636.
34. Bruner, C.A.; Mangiapane, M.L.; Fink, G.D. Subfornical organ. Does it protect against angiotensin II-induced hypertension in the rat? *Circ. Res*. **1985**, *56*, 462–466.
35. Mortensen, L.H.; Fink, G.D. Captopril prevents chronic hypertension produced by infusion of endothelin-1 in rats. *Hypertension* **1992**, *19*, 676–680.
36. Cangemi, C.; Skov, V.; Poulsen, M.K.; Funder, J.; Twal, W.O.; Gall, M.A.; Hjortdal, V.; Jespersen, M.L.; Kruse, T.A.; Aagard, J.; *et al*. Fibulin-1 is a marker for arterial extracellular matrix alterations in type 2 diabetes. *Clin. Chem.* **2011**, *57*, 1556–1565.
37. Andersen, D.C.; Jensen, C.H.; Schneider, M.; Nossent, A.Y.; Eskildsen, T.; Hansen, J.L.; Teisner, B.; Sheikh, S.P. MicroRNA-15a fine-tunes the level of Delta-like 1 homolog (DLK1) in proliferating 3T3-L1 preadipocytes. *Exp. Cell Res*. **2010**, *316*, 1681–1691.
38. Andersen, D.C.; Andersen, P.; Schneider, M.; Jensen, H.B.; Sheikh, S.P. Murine "cardiospheres" are not a source of stem cells with cardiomyogenic potential. *Stem Cells* **2009**, *27*, 1571–1581.
39. Vandesompele, J.; De Preter, K.; Pattyn, F.; Poppe, B.; Van Roy, N.; de Paepe, A.; Speleman, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. **2002**, *3*, RESEARCH0034.
40. Hellemans, J.; Mortier, G.; de Paepe, A.; Speleman, F.; Vandesompele, J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol*. **2007**, *8*, R19.

Reprinted from *IJMS*. Cite as: Fenoglio, C.; Ridolfi, E.; Galimberti, D.; Scarpini, E. An Emerging Role for Long Non-Coding RNA Dysregulation in Neurological Disorders. *Int. J. Mol. Sci.* **2013**, *14*, 20427-20442.

Review

An Emerging Role for Long Non-Coding RNA Dysregulation in Neurological Disorders

Chiara Fenoglio *, Elisa Ridolfi, Daniela Galimberti and Elio Scarpini

Department of Pathophysiology and Transplantation, University of Milan, "Dino Ferrari" Center, IRCCS Cà Granda Foundation Ospedale Maggiore Policlinico, Via F.Sforza 35, Milan 20122, Italy; E-Mails: elisa.ridolfi@unimi.it (E.R.); daniela.galimberti@unimi.it (D.G.); elio.scarpini@unimi.it (E.S.)

* Author to whom correspondence should be addressed; E-Mail: chiara.fenoglio@unimi.it; Tel.: +39-02-55033858; Fax: +39-02-50320430.

Received: 14 June 2013; in revised form: 18 September 2013 / Accepted: 25 September 2013 / Published: 14 October 2013

Abstract: A novel class of transcripts, long non coding RNAs (lncRNAs), has recently emerged as key players in several biological processes, including dosage compensation, genomic imprinting, chromatin regulation, embryonic development and segmentation, stem cell pluripotency, cell fate determination and potentially many other biological processes, which still are to be elucidated. LncRNAs are pervasively transcribed in the genome and several lines of evidence correlate dysregulation of different lncRNAs to human diseases including neurological disorders. Although their mechanisms of action are yet to be fully elucidated, evidence suggests lncRNA contributions to the pathogenesis of a number of diseases. In this review, the current state of knowledge linking lncRNAs to different neurological disorders is discussed and potential future directions are considered.

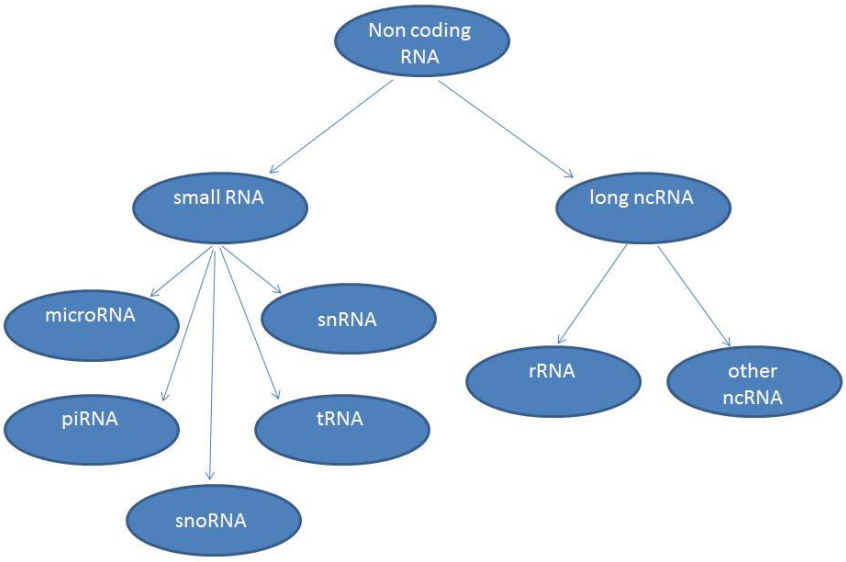
Keywords: lncRNA; RNA; neurodegenerative diseases; epigenetics

1. Introduction

For the last few decades of the 20th century, the underlying dogma of molecular biology has been that the purpose of RNA is to direct the assembly of proteins from amino acids. A few exceptions to this paradigm were known (such as ribosomal RNA and transfer RNA, which are functional RNA macromolecules that do not code for protein). Non-coding RNAs (ncRNAs) include the familial

"housekeeping" RNAs and the thousands of regulatory RNAs that are the subject of recent intense investigation. NcRNAs come in many different sizes and for this reason have been divided into small and long classes: small ncRNAs (sncRNA) being less than 200 nucleotides (nt), and long ncRNAs (lncRNA) greater than 200 nt to over 100 kb in length [1]. The current cut-off is arbitrary and corresponds to certain biochemical fractionation protocols and excludes most categories of small infrastructural or regulatory RNAs (tRNAs, snRNAs, miRNAs, siRNAs, piRNAs, tiRNAs, spliRNAs, sdRNAs and others, Figure 1).

Figure 1. Non-coding RNAs (ncRNAs) are arbitrarily grouped basing on size. Small ncRNAs being less than 200 nucleotides and long ncRNAs greater than 200 nucleotides.



An estimation of the number of lncRNA loci from transcriptional surveys in mammals suggests that they are at least as numerous as protein-coding genes [2] with many of lncRNAs identified in intergenic regions alone [3]. With a few exceptions, it is only within the past few years that the functions and mechanisms of lncRNAs have begun to emerge.

The lncRNAs that have been studied in detail were found to be involved in different biological processes including X chromosome inactivation, nuclear structure, genomic imprinting and development.

In the following paragraphs some of the known functions of lncRNAs will be discussed with particular regard to their role in neurological disorders.

2. Identification of lncRNAs

Data derived from massive cloning and traditional sequencing methods have demonstrated that mammalian genomes produce thousands of RNA transcripts in addition to protein-coding genes [2]. These studies modified our simplistic view of the genome and suggested the presence of thousands of previously unknown transcripts. Recently more than 3000 intergenic ncRNAs have been found in the mammalian genome, by using information from chromatin modifications [3,4]. Evidence gained from ChIP-Seq analyses demonstrated that transcribed protein-coding genes have unusual chromatin modifications; these modifications are trimethylation of histone H3 at lysine 4 (H3K4me3) at the promoter region and trimethylation of histone H3 at lysine 36 (H3K36me3) in the body of the

gene [5]. By eliminating these chromatin domains that correspond to protein-coding genes more than 3000 intergenic domains transcribed into lncRNAs have been found. Bioinformatics analyses showed that the majority of these transcripts have no protein-coding capability. Moreover, it was shown that many of these lncRNAs are able to interact with multiple chromatin-modifying complexes in different human and mouse cell types suggesting that lncRNAs may be involved in epigenetic regulation [3].

Recent advances in RNA sequencing technology (RNA-Seq) addressed further interrogation of the total cellular RNA, or, transcriptome, at a much higher resolution [6]. Thousands of transcripts, in addition to protein-coding mRNAs and microRNAs, have been found to be expressed in a wide range of tissues and cell types [7].

3. Epigenetics

Many of the lncRNAs seem to be involved in epigenetic mechanisms of gene regulation. Although the field of epigenetics earned its name over 50 years ago, just in this past decade the significance of epigenetics has been recognized in human health and disease [8]. The term epigenetics refers to changes in gene expression and/or phenotype that can be heritable without a change in the underlying DNA sequence. Several factors contribute to epigenetic mechanisms of gene regulation including DNA methylation, histone modifications, and ncRNAs. DNA methylation in the promoter region of genes is typically associated with transcriptional repression. Several mammalian enzymes are responsible for establishing and maintaining DNA methylation in the genome [9].

The modification of histone proteins has also been found in epigenetic regulation [10]. Histones are highly conserved proteins that package DNA in the nucleus and modulate the accessibility of transcription factors and RNA polymerases to DNA. Histone modifications typically take place at amino acids located in the *N*-termini of histones such as lysine, arginine and serine residues that can be acetylated or methylated (lysine and arginine) and phosphorylated (serine). Histone modifications are placed on and removed from histone residues by numerous enzymes that usually work as part of multi-protein complexes [10].

A relevant question in biology is how chromatin-modifying complexes are targeted to specific genomic loci since many of these enzymes lack DNA binding capacity. Recent studies suggest a potential role for ncRNA in driving chromatin-modifying complexes to genomic loci [11–13], however, the exact mechanism by which ncRNAs drive complexes is not known and is currently under investigation.

4. Functions of lncRNAs

Basing on our current knowledge of lncRNAs, it appears that such molecules are involved in many different aspects of cellular functions (reviewed in [14,15]). The roles of lncRNAs in the regulation of gene expression and organismal development are different and just beginning to be discovered. Biological processes dependent upon lncRNAs include imprinting and gene dosage regulation, stem cell pluripotency, embryonic development and segmentation, hematopoiesis, and neural cell fate determination (see for review ref [16]). LncRNAs may employ a number of mechanisms to impact gene expression via *cis* and *trans* processes.

4.1. Gene Imprinting

While the function of parental gene imprinting is still unclear, lncRNAs have been found to participate in imprinting processes. These are referred to as the events that influence the monoallelic expression of a gene according to its parents of origin. Imprinting control regions (ICRs) are DNA regions that are differentially methylated depending on their parental origins. Unmethylated ICRs cause specific expression of nearby lncRNAs, which then suppress neighboring genes in *cis*. Several imprinted clusters contain protein-coding genes and lncRNAs that are reciprocally expressed, such as *IGf2r*/*Air* [17], *Dlk1*/*Gtl2* [18,19], and *Nesp*/*Nespas*/*Gnas* [20]. Some of these lncRNAs can exert their function by recruiting epigenetic factors, such as PRC2 and G9a, in order to control the imprinted expression of neighboring coding genes [21,22].

Airn and Kcnq1ot1 are examples of lncRNAs that cause suppression of paternally inherited genes. Kcnq1ot1, in particular, is involved in the repression of several protein-coding genes in *cis* through interaction with repressive chromatin modifying complexes [23,24].

H19 was one of the first mammalian lncRNAs to be identified and is highly expressed in the embryo [25,26]. Though it does not seem to act as an lncRNA [27], H19 likely functions as an miRNA precursor [28,29] and is mutually imprinted with the protein-coding gene *Igf2*.

4.2. Gene Dosage and X Chromosome Inactivation

The X chromosome inactivation (XCI) indicates the mechanisms by which the difference in X-linked gene dosage between XX females and XY males is exerted in therian mammals in which one of the two X chromosomes in females is silenced (the inactive X or Xi) so that only one X remains active and is expressed in each female cell (the active X, or Xa) [30]. It is well-known that XCI in placental mammals is largely controlled by a cluster of lncRNA loci known as the X-inactivation center (*Xic*) [31]. The X (inactive)-specific transcript (*Xist*) is highly expressed from Xi during the onset of *XCI* but not from *Xa*, contributing a defining moment for the realization that lncRNAs can have profound roles in the control of gene expression even though the exact mechanism of action is still not completely understood. Some evidence suggests that Xist mediates the chromosome X silencing effects by interacting with repressive chromatin-modifying complexes such as PRC2 [22].

Xist itself is also regulated by other lncRNAs. Initially, *Xist* and *Tsix*, its antisense transcript, transcribed from a promoter downstream of *Xist*, are expressed on both X chromosomes. However, *Tsix* expression continues on the X that will remain active (*Xa*) and this activity recruits dnmt3a to suppress *Xist* from being transcribed on *Xa* [32]. Conversely on *Xi*, it is Tsix that is suppressed, potentially through another lncRNA that is part of the X inactivation center, Jpx [33]. With *Tsix* suppressed, the protein PRC2 is recruited to induce histone modification marks at the 5' end of *Xist*. This upregulates *Xist* expression on *Xi* and causes further propagation of these silencing marks throughout Xi, which are maintained across the lifetime of the organism [33].

4.3. Embryonic Development and Segmentation

LncRNAs are likely implicated in processes involving animal development. The *Hox* genes encode homeodomain TFs that are crucial for anterior-posterior pattern formation in bilateral metazoans [34].

Hox genes are structured in linear clusters along the chromosome, and mammals have four paralogous clusters, *HoxA*, -*B*, -C, and -D. Several lncRNAs are encoded within these clusters, including HOTAIR (Hox antisense intergenic RNA) from *HoxC*, and HOTTIP and Mistral from *HoxA* [35–37]. The expression of *Hox* genes is also regulated by lncRNAs [36]. Some Hox-related lncRNAs operate in *cis*, having either enhancing or repressive effects. However, some like the human HOTAIR works in *trans*, and it is expressed from the *Hox* locus marking a boundary of active and inactive chromatin [35]. Furthermore, HOTAIR, similar to *Xist*, interacts with chromatin-modifying complexes such as PRC2 and the corepressor complex CoREST [35] and may guide these complexes to genomic loci. Overexpression of HOTAIR caused cells to become metastatic when injected into mice compared to control cells with an empty vector. Moreover, HOTAIR may serve as a scaffold for targeting chromatin-modifying complexes to chromatin [35].

Recently HOTAIR was found to be crucial for cell growth and viability and that its knockdown induced apoptosis in breast cancer cells Moreover it was found that HOTAIR is transcriptionally induced by estradiol [38]. It is possible that other lncRNAs that interact with chromatin-modifying complexes also function in a manner similar to HOTAIR.

4.4. Stem Cell Pluripotency and Cell Fate Determination

The promoters of more than 100 lncRNAs are bound by stem cell factors. Disruption of these lncRNAs can alter cell differentiation. The human lncRNA-RoR (RoR) is a recently identified lncRNA that is capable of reprogramming differentiated cells to induce pluripotent stem cells [39,40]. RoR is highly expressed both in embryonic stem cells and in induced pluripotent stem cells, due to the regulation of RoR by pluripotency transcription factors such as Oct4, Sox2 and Nanog. Interestingly it was observed that knockdown of RoR leads to a modest increase in apoptosis and activation of p53 pathways [40]. Although the underlying mechanisms still remain to be fully clarified, Zhang *et al.* recently [41] demonstrated that human RoR is a strong negative regulator of p53 influencing the inhibition of p53-mediated cell cycle arrest and apoptosis.

Recently, a refined analysis from Guttman *et al.* [39] performed loss-of-function studies on 226 lncRNAs expressed in mouse embryonic stem cells characterizing the effects on gene expression. The authors identified 26 lncRNAs able to maintain the pluripotent status. In particular, knockdown of these lncRNAs resulted in a loss of pluripotency markers, and reduction of Nanog promoter activity. Simultaneously, expression patterns similar to differentiation into specific lineages were produced, suggesting that lncRNAs repressed differentiation programs in mouse embryonic stem cells. Altogether these findings support the hypothesis that some lncRNAs are integral members of a regulatory network, together with key pluripotent transcription factors, which modulate pluripotency and lineage-specific differentiation pathways in mouse embryonic stem cells.

lncRNAs are implicated in cell fate determination events in multiple cells lineages, including the nervous system. Taurine upregulated gene 1 (*TUG1*) is an lncRNA expressed in the developing retina and brain, as well as in adult tissues. It has been found that in the newborn retina, loss of TUG1 resulted in malformed or non-existent outer segments of transfected photoreceptors, thus suggesting that TUG1 is required for the proper formation of photoreceptors in the developing rodent retina [42]. Evf2 is instead a mouse lncRNA that appears to recruit Dlx and Mecp2 transcription factors to

important DNA regulatory elements in the *Dlx5-Dlx6* intergenic regions and controlled Dlx5, *Dlx6* and Gad1 expression through *cis* and *trans* acting mechanisms. Evf2 mouse mutants appeared to have reduced numbers of GABAergic interneurons in early postnatal hippocampus and the dentate gyrus. This situation is restored to normality in Evf2 mutant adult hippocampus although reduced synaptic inhibition still occurred [43,44].

Although many lncRNAs have been shown to regulate gene expression only a few have been shown to have other cellular functions. For example, NEAT1 has been shown to play an important role in paraspeckle formation [45]. Also NRON has a role in nuclear import/export [46]. All together, these studies suggest that the lncRNAs have different cellular functions, many of which are yet to be identified and characterized for the mechanism of their function.

5. LncRNAs in Human Diseases

As the functions and mechanisms of lncRNAs are beginning to emerge, there is an intense interest in identifying any potential role of these molecules in human diseases. Several studies have shown that lncRNAs are dysregulated in human pathologies, however it has yet to be shown that these molecules are enough to drive the disease status.

lncRNAs have been strongly associated with cancer [47]. Recently, the lncRNA PCAT-1, was found to promote cell proliferation and is a target of PRC2 regulation, [48]. Moreover, ANRIL, which is upregulated in prostate cancer, is required for the expression of the tumor suppressors INK4a/p16 and INK4b/p15 [49]. HOTAIR upregulation is associated with poor prognosis in breast cancer [11], liver [50], colorectal [51], gastrointestinal [52] and pancreatic [53] cancers and probably contributes to increase also tumor invasiveness and metastasis [11].

MALAT-1, which is another lncRNA associated with various cancers and metastasis [54] was found to affect the transcriptional and post-transcriptional regulation of cytoskeletal and extracellular matrix genes [55]. Although lncRNAs have been extensively investigated in cancers several lines of evidence suggest a possible role also in different disease conditions such as cardiovascular diseases. For example, two lncRNAs have been found to be dysregulated in heart disease; the expression of the lncRNA MIAT is associated with increased risk of myocardial infarction, whereas the lncRNA ANRIL is associated with increased risk to coronary heart disease [56,57]. Recently, a novel lncRNA has been discovered, named DBE-T, that functions in *cis* and localizes to the Facioscapulohumeral muscular dystrophy (FSHD) locus. FSHD is the third most common myopathy and is predominantly caused by a contraction of specific repeats mapping on chromosome 4q35 [58]. It is suggested that DBE-T likely acts as a locus control element by promoting active chromatin domains and thus FSHD would be caused from lncRNA “promoter mutations” able to perturb DBE-T regulation [58]. Furthermore, a novel lncRNA seems to be involved in the pathogenic mechanisms underneath the HELLP syndrome (hemolysis, elevated liver enzymes, low platelets) that is a recessively inherited life-threatening pregnancy complication [59].

6. Role of lncRNAs in the Central Nervous System

Recent evidence demonstrate that lncRNAs contribute to the complex biological system organization and gene regulatory networks of the central nervous system (CNS), affecting brain

patterning, neural stem cell maintenance, neurogenesis and gliogenesis, stress responses, and synaptic and neural plasticity.

Mercer and colleagues identified 849 lncRNAs (among the 1328 examined), that are expressed in the adult mouse brain and found that the majority were associated with specific neuroanatomical regions, cell types, or subcellular compartments [60]. A complementary study showed that over 200 of these lncRNAs are expressed in developing and adult mouse brain and are largely derived from genomic loci located proximal to protein-coding genes with similar expression profiles in the brain [61].

Guttman *et al.* discovered more than 1000 evolutionarily conserved intergenic lncRNAs in mouse by analyzing chromatin signatures from four mouse cell types [5]. A functional analysis of the expression of these lncRNAs revealed the presence of a "brain cluster" of lncRNAs that is associated with biological processes including hippocampal development, oligodendrocyte (OL) myelination, brain aging, CREB and PGC1-alpha transcriptional regulation, and GABAergic neuronal (GABAN), G protein coupled receptor and calcineurin signaling pathways. An additional study demonstrated that 169 lncRNAs are differentially expressed during the sequential processes of mouse ventral forebrain-derived neural precursor cells mediated lineage restriction, GABAN and OL lineage specification, progressive OL lineage maturation, and terminal differentiation including myelination [62].

Detailed analyses of specific lncRNAs, dynamically expressed in the CNS, reveal potential roles in mediating neural cell fate decisions. The Sox2OT lncRNA, which contains the *Sox2* gene within one of its introns and is subsequently transcribed in the same direction [63], is expressed in regions of constitutive adult neurogenesis [60]. Moreover, Sox2OT is dynamically regulated in CNS structures during development, where it may be responsible for modulating Sox2 expression [64]. The lncRNA Nkx2.2AS regulates Nkx2.2, a transcription factor that is critical for OL lineage specification. A recent study reported that forced expression of Nkx2.2AS in NSCs *in vitro* enhances their differentiation along the OL lineage, in part, by inducing an increase in Nkx2.2 mRNA levels [65].

LncRNAs also modulate synaptic plasticity and promote long-term changes in synaptic strength. The rodent-specific *BC1* and primate-specific BC200 lncRNAs, are selectively targeted to postsynaptic dendritic compartments, where they regulate local protein synthesis by repressing the initiation of translation through an eIF4A-dependent mechanism [66–69]. Similarly, NTAB is a lncRNA that is expressed in developing and adult rat brain, where it is also found in neuronal processes [70]. Another lncRNA, MALAT-1, is enriched in hippocampal neurons, where it regulates several serine/arginine splicing factors important for synapse formation, density and maturation [71].

6.1. Dysregulation of lncRNAs in Neurological Disorders

6.1.1. LncRNAs Play a Role in the Pathophysiology of Several Neurological Disorders

Angelman syndrome (AS) is a neurodevelopmental disorder associated with genomic imprinting and characterized by severe neurologic abnormalities [72]. Ube3a-as is a lncRNA transcribed antisense to the maternally expressed *Ube3a* gene, mutated or deleted in AS, suggesting that Ube3a-as may repress paternal Ube3a expression. Other studies have shown that repression of Ube3a is dependent on Ube3a-as [73,74]. However, other data has demonstrated that silencing of paternal *Ube3a* can occur in

the absence of Ube3a-as and implies a more complex regulatory relationship underlying the imprinting of *Ube3a* [75].

LncRNAs may influence the pathogenesis of fragile X syndrome (FXS), which is characterized by a triplet nucleotide repeat expansion in the 5'UTR of *FMR1*, the gene encoding the neuronal development protein, FMRP. The lncRNAs ASFMR1 and FMR4 are generated from the *FMR1* gene locus. ASFMR1 has multiple alternative splicing patterns and overlaps the 5' untranslated region (UTR) CGG repeat region of *FMR1* [76]; FMR4 is initiated upstream of the *FMR1* start site. Alternative splicing of ASFMR1 seems to exhibit pre-mutation-specific profiles and is also silenced in FXS patients and upregulated in pre-mutation carriers, suggesting that a common process is responsible for regulating the expression of these transcripts. *FMR4* is also silenced in FXS patients because of a CGG expansion repeat in the 5'UTR of the *FMR1* gene and upregulated in pre-mutation carriers [77]; thus, their absence in the neurons of affected patients might contribute to the pathogenesis of this neurological disorder.

Another microsatellite expansion disease in which lncRNAs are involved is the spinocerebellar ataxia type 8 (SCA8), an autosomal dominant disorder, characterized by bidirectional transcription of this expansion repeat from opposite strands, forming both a protein-coding transcript encoding a polyglutamine expansion, ATXN8, and a lncRNA transcript containing a CUG expansion, ATXN8OS [78]. This suggests that SCA8 pathogenesis involves a toxic gain of function at both the protein and RNA levels. A recent study found that the expanded ATXN8OS transcript accumulates in ribonuclear inclusions in the GABAergic neurons of SCA8 patients [79]. These inclusions co-localized with splicing factor, MBNL1, altering the activity of MBNL alternative splicing proteins [79].

Huntington's disease (HD) is caused by an expansion repeat mutation in the *Htt* gene, which encodes a ubiquitously expressed 3144 amino acid protein of unknown function, leading to a toxic gain of function in the mutant protein [80], which promotes aberrant nuclear-cytoplasmic trafficking of the master neuronal regulator REST. The result is the deregulation of REST target gene expression in tissues from animal models of HD and human HD, which include both protein-coding genes as well as ncRNAs, such as lncRNAs. It is therefore likely that HD tissues are also characterized by dysregulation of lncRNA expression. ChiP-seq data showed that the *HAR1* locus is under control of REST [81]. The *HAR1* region contains lncRNAs, *HAR1F* and *HAR1R*. HAR1F is specifically expressed in the neurons of the marginal zone during development of the cortex and in the frontal cortex, hippocampus, cerebellum, thalamus and hypothalamus in the adult brain [82]. The levels of HAR1F and HAR1R are decreased in HD brains compared with normal brains [83].

6.1.2. LncRNAs in Neurodegenerative Diseases, such as Alzheimer's Disease (AD)

β-site amyloid precursor protein-cleaving enzyme 1 (BACE1) is a crucial enzyme in AD pathophysiology, involved in the cleavage of the amyloid precursor protein (*APP*) and the generation of amyloid peptides which can aggregate and form plaques. Faghihi and colleagues characterized a conserved non-coding antisense transcript for BACE1, called BACE1-AS, which functions as a regulator of BACE1 gene expression. BACE1-AS upregulates BACE1 levels in response to a variety of stresses, including Aβ 1–42 exposure, and is elevated in several brain regions of patients with AD. These findings imply that BACE1-AS is deregulated in AD, which induces feed-forward regulation of BACE1, increases Aβ levels, and thus may promote the pathogenesis of the disease [84].

Mus *et al.* found a link between altered levels of a lncRNA, BC200, and AD [85]. Increased levels of BC200 were found in brain regions that are preferentially affected in AD, such as the hippocampus, which correlated with disease severity. Further, in advanced stages of AD, BC200 was mis-localized and clustered in the perikaryon. These observations suggest that deregulation of this synaptic lncRNA is involved in the synaptic and neural network dysfunction that is found in both early and later stages of AD [85].

Together with neurological diseases, a number of psychiatric disorders have also been associated with lncRNAs (Table 1). The risk of developing schizophrenia (SZ), schizoaffective disorder, bipolar disorder, major depression, and autistic spectrum disorders has been linked to the disruption of the *DISC* genomic locus, which encodes both the *DISC1* protein-coding gene and the *DISC2* lncRNA [86–88]. *DISC2* controls the expression of its partner, *DISC1*, which modulates multiple aspects of CNS structure and function such as embryonic and adult neurogenesis [89]. Another lncRNA, Gomafu, implicated in brain and retinal development [62,90] binds directly to the splicing factors QKI and SRSF1 and the dysregulation of Gomafu induces alternative splicing that resemble those observed in SZ for the archetypal SZ-associated genes *DISC1* and *ERB4*. Moreover, Gomafu is downregulated in post-mortem cortex of SZ subjects, suggesting a role in SZ pathogenesis for this lncRNA [91].

Table 1. Examples of lncRNAs that are dysregulated in neurological disorders.

lncRNA	Disease association	Biological function	Reference
Ube3a-as	Angelman Syndrome	Repress paternal *Ube3a* expression	[72]
ASFMR1 FMR4	Fragile X Syndrome	Regulate the expression of *ASFMR1* and *FMR1* genes	[76,77]
ATXN8OS	SCA8	Alteration of the activity of the splicing factor *MBNL1*	[78,79]
HAR1F HAR1R	Hungtington's Disease	Influence genes promoting aberrant nuclear-cytoplasmatic trafficking of *REST* gene	[83]
BC200 BACE1-AS	Alzheimer's	Involved in the synaptic and neural network dysfunction Regulates *BACE1* gene expression	[84,85]
DISC2 Gomafu	Psychiatric disorders	Controls the expression of *DISC1* Implicated in brain and retinal development	[86–89] [91]

7. Conclusions and Future Perspectives

Several recent studies suggest that lncRNAs play a pivotal role in many key biological processes, although their mechanisms of action are yet to be fully elucidated. Currently there is great interest in identifying the functions of this novel class of transcripts. There is strong evidence that many lncRNAs are biologically relevant, with a large percentage of these molecules functioning through their interactions with chromatin-modifying complexes to alter gene expression [3]. These findings are beginning to shed light on how chromatin-modifying complexes are targeted to specific genomic loci and suggest the interesting idea that lncRNAs are in some way driving chromatin-modifying complexes to genomic loci.

LncRNAs have, in a relatively short period of time, become recognized as a major new class of genes that may potentially comprise a major component of the genome's information content, complementary and comparable in abundance and complexity to the proteome. Furthermore lncRNAs have already been reported in a wide range of human diseases suggesting that their activity is crucial for human health.

In addition, therapeutic strategies that target endogenous mRNA molecules, such as those employing RNA interference (RNAi) and other customized oligonucleotide approaches with the capacity to reprogram disease-associated mRNAs, are now being developed [92]. These approaches could be adapted to target lncRNAs whose expression is dysregulated in CNS disorders. These observations suggest that lncRNAs represent a versatile class of factors that are centrally important to the modulation of different CNS processes and may represent a major layer underlying the genetic programming of brain development that could potentially be utilized for developing novel diagnostic and therapeutic tools for the cure of CNS disorders.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Mattick, J.S. Non-coding RNAs: The architects of eukaryotic complexity. *EMBO* **2001**, *2*, 986–991.
2. Harrow, J.; Frankish, A.; Gonzalez, J.M.; Tapanari, E.; Diekhans, M.; Kokocinski, F.; Aken, B.L.; Barrell, D.; Zadissa, A.; Searle, S.; *et al.* GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res*. **2012**, *22*, 1760–1774.
3. Khalil, A.M.; Guttman, M.; Huarte, M.; Garber, M.; Raj, A.; Rivea Morales, D.; Thomas, K.; Presser, A.; Bernstein, B.E.; van Oudenaarden, A.; *et al.* Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11667–11672.
4. Marques, A.C.; Ponting, C.P. Catalogues of mammalian long noncoding RNAs: Modest conservation and incompleteness. *Genome Biol*. **2009**, *10*, R124.
5. Guttman, M.; Amit, I.; Garber, M.; French, C.; Lin, M.F.; Feldser, D.; Huarte, M.; Zuk, O.; Carey, B.W.; Cassady, J.P.; *et al.* Lander Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **2009**, *458*, 223–227.
6. Nagalakshmi, U.; Waern, K.; Snyder, M. RNA-Seq: A method for comprehensive transcriptome analysis. *Curr. Protoc. Mol. Biol.* **2010**, doi:10.1002/0471142727.
7. Wang, E.T.; Sandberg, R.; Luo, S.; Khrebtukova, I.; Zhang, L.; Mayr, C.; Kingsmore, S.F.; Schroth, G.P.; Burge, C.B. Alternative isoform regulation in human tissue transcriptomes. *Nature* **2008**, *456*, 470–476.
8. Egger, G.; Liang, G.; Aparicio, A.; Jones, P.A. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **2004**, *429*, 457–463.
9. Kaneda, M.; Okano, M.; Hata, K.; Sado, T.; Tsujimoto, N.; Li, E.; Sasaki, H. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* **2004**, *429*, 900–903.

10. Bernstein, B.E.; Meissner, A.; Lander, E.S. The mammalian epigenome. *Cell* **2007**, *128*, 669–681.
11. Gupta, R.A.; Shah, N.; Wang, K.C.; Kim, J.; Horlings, H.M.; Wong, D.J.; Tsai, M.C.; Hung, T.; Argani, P.; Rinn, J.L.; *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **2010**, *464*, 1071–1076.
12. Reis, E.M.; Louro, R.; Nakaya, H.I.; Verjovski-Almeida, S. As antisense RNA gets intronic. *OMICS* **2005**, *9*, 2–12.
13. Tsai, M.C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J.K.; Lan, F.; Shi, Y.; Segal, E.; Chang, H.Y. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **2010**, *329*, 689–693.
14. Mattick, J.S. The genetic signatures of noncoding RNAs. *PLoS Genet.* **2009**, *5*, e1000459.
15. Mattick, J.S.; Amaral, P.P.; Dinger, M.E.; Mercer, T.R.; Mehler, M.F. RNA regulation of epigenetic processes. *Bioessays* **2009**, *31*, 51–59.
16. Clark, M.B.; Mattick, J.S. Long noncoding RNAs in cell biology. *Semin. Cell Dev. Biol.* **2011**, *22*, 366–376.
17. Lyle, R.; Watanabe, D.; Vruchte, W.; Lerchner, O.; Smrzka, W. The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. *Nat. Genet.* **2000**, *25*, 19–21.
18. Schmidt, J.V.; Matteson, P.G.; Jones, B.K.; Guan, X.J.; Tilghman, S.M. The *Dlk1* and *Gtl2* genes are linked and reciprocally imprinted. *Genes Dev.* **2000**, *14*, 1997–2002.
19. Da Rocha, S.T.; Edwards, C.A.; Ito, M.; Ogata, T.; Ferguson-Smith, A.C. Genomic imprinting at the mammalian Dlk1-Dio3 domain. *Trends Genet.* **2008**, *24*, 306–316.
20. Williamson, C.M.; Turner, M.D.; Ball, S.T.; Nottingham, W.T.; Glenister, P. Identification of an imprinting control region affecting the expression of all transcripts in the Gnas cluster. *Nat. Genet.* **2006**, *38*, 350–355.
21. Nagano, T.; Mitchell, J.A.; Sanz, L.A.; Pauler, F.M.; Ferguson-Smith, A.C.; Feil, R.; Fraser, P. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **2008**, *322*, 1717–1720.
22. Zhao, J.; Sun, B.K.; Erwin, J.A.; Song, J.J.; Lee, J.T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **2008**, *322*, 750–756.
23. Pandey, R.R.; Mondal, T.; Mohammad, F.; Enroth, S.; Redrup, L.; Komorowski, J.; Nagano, T.; Mancini-Dinardo, D.; Kanduri, C. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell* **2008**, *32*, 232–234.
24. Kanduri, C.; Thakur, N.; Pandey, R. The length of the transcript encoded from the Kcnq1ot1 antisense promoter determines the degree of silencing. *EMBO J.* **2006**, *25*, 2096–2106.
25. Bartolomei, M.S.; Zemel, S.; Tilghman, S.M. Parental imprinting of the mouse *H19* gene. *Nature* **1991**, *351*, 153–155.
26. Brannan, C.I.; Dees, E.C.; Ingram, R.S.; Tilghman, S.M. The product of the *H19* gene may function as an RNA. *Mol. Cell Biol.* **1990**, *20*, 28–36.
27. Gabory, A.; Jammes, H.; Dandolo, L. The H19 locus: Role of an imprinted non-coding RNA in growth and development. *Bioessays* **2010**, *32*, 473–480.
28. Cai, X.; Cullen, B.R. The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* **2007**, *13*, 313–316.

29. Keniry, A.; Oxley, D.; Monnier, P.; Kyba, M.; Dandolo, L. The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nat. Cell Biol.* **2012**, *14*, 859–865.
30. Lee, J.T. Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 815–826.
31. Brown, C.J.; Lafrenière, R.G.; Powers, V.E.; Sebastio, G.; Ballabio, A. Localization of the X inactivation centre on the human X chromosome in Xq13. *Nature* **1991**, *349*, 82–84.
32. Do, J.T.; Han, D.W.; Gentile, L.; Sobek-Klocke, I.; Stehling, M.; Schöler, H.R. Enhanced reprogramming of Xist by induced upregulation of Tsix and Dnmt3a. *Stem Cells* **2008**, *26*, 2821–2831.
33. Tian, D.; Sun, S.; Lee, J.T. The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell* **2010**, *143*, 390–403.
34. Pearson, J.C.; Lemons, D.; McGinnis, W. Modulating *Hox* gene functions during animal body patterning. *Nat. Rev. Genet.* **2005**, *6*, 893–904.
35. Rinn, J.L.; Kertesz, M.J.; Wang, K.; Squazzo, S.L.; Xu, X. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **2007**, *129*, 1311–1323.
36. Bertani, S.; Sauer, S.; Bolotin, E.; Sauer, F. The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. *Mol. Cell* **2011**, *43*, 1040–1046.
37. Wang, K.C.; Yang, Y.W.; Liu, B.; Sanyal, A.; Corces-Zimmerman, R. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **2011b**, *472*, 120–124.
38. Bhan, A.; Hussain, I.; Ansari, K.I.; Kasiri, S.; Bashyal, A.; Mandal, S.S. Antisense Transcript Long Noncoding RNA (lncRNA) HOTAIR is transcriptionally induced by Estradiol. *J. Mol. Biol.* **2013**, doi:10.1016/j.jmb.2013.01.022.
39. Guttman, M.; Donaghey, J.; Carey, B.W.; Garber, M.; Grenier, J.K.; Munson, G.; Young, G.; Lucas, A.B.; Ach, R.; Bruhn, L.; *et al.* lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **2011**, *477*, 295–300.
40. Loewer, S.; Cabili, M.N.; Guttman, M.; Loh, Y.H.; Thomas, K.; Park, I.H.; Garber, M.; Curran, M.; Onder, T.; Agarwal, S.; *et al.* Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat. Genet.* **2010**, *42*, 1113–1117.
41. Zhang, A.; Zhou, N.; Huang, J.; Liu, Q.; Fukuda, K.; Ma, D.; Lu, Z.; Bai, C.; Watabe, K.; Mo, Y.Y. The human long non-coding RNA-RoR is a p53 repressor in response to DNA damage. *Cell Res.* **2013**, *23*, 340–350.
42. Young, T.L.; Matsuda, T.; Cepko, C.L. The noncoding RNA taurine upregulated gene 1 is required for differentiation of the murine retina. *Curr. Biol.* **2005**, *15*, 501–512.
43. Feng, J.; Bi, C.; Clark, B.S.; Mady, R.; Shah, P.; Kohtz, J.D. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev.* **2006**, *20*, 1470–1484.
44. Bond, A.M.; Vangompel, M.J.; Sametsky, E.A.; Clark, M.F.; Savage, J.C.; Disterhoft, J.F. Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry. *Nat. Neurosci.* **2009**, *12*, 1020–1027.

45. Clemson, C.M.; Hutchinson, J.N.; Sara, S.A.; Ensminger, A.W.; Fox, A.H.; Chess, A.; Lawrence J.B. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell* **2009**, *33*, 717–726.
46. Willingham, A.T.; Orth, A.P.; Batalov, S.; Peters, E.C.; Wen, B.G.; Aza-Blanc, P.; Hogenesch, J.B.; Schultz, P.G. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **2005**, *309*, 1570–1573.
47. Gutschner, T.; Diederichs, S. The hallmarks of cancer: A long non-coding RNA point of view. *RNA Biol.* **2012**, *9*, 703–719.
48. Prensner, J.R.; Iyer, M.K.; Balbin, O.A.; Dhanasekaran, S.M.; Cao, Q. Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. *Nat. Biotechnol.* **2011**, *29*, 742–749.
49. Kotake, Y.; Nakagawa, T.; Kitagawa, K.; Suzuki, S.; Liu, N. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. *Oncogene* **2011**, *30*, 1956–1962.
50. Yang, Z.; Zhou, L.; Wu, L.M.; Lai, M.C.; Xie, H.Y.; Zhang, F.; Zheng, S.S. Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. *Ann. Surg. Oncol.* **2011**, *18*, 1243–1250.
51. Kogo, R.; Shimamura, T.; Mimori, K.; Kawahara, K.; Imoto, S.; Sudo, T.; Tanaka, F.; Shibata, K.; Suzuki, A.; Komune, S.; *et al.* Long noncoding RNA HOTAIR regulates polycombdependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res.* **2011**, *71*, 6320–6326.
52. Niinuma, T.; Suzuki, H.; Nojima, M.; Nosho, K.; Yamamoto, H. Upregulation of miR-196a and HOTAIR drive malignant character in gastrointestinal stromal tumors. *Cancer Res.* **2012**, *72*, 1126–1136.
53. Kim, K.; Jutooru, I.; Chadalapaka, G.; Johnson, G.; Frank J.; Burghardt, R.; Kim, S.; Safe, S. HOTAIR is a negative prognostic factor and exhibits pro-oncogenic activity in pancreatic cancer. *Oncogene* **2012**, doi:10.1038/onc.2012.193.
54. Lin, R.; Roychowdhury-Saha, M.; Black, C.; Watt, A.T.; Marcusson, E.G. Control of RNA processing by a large non-coding RNA over-expressed in carcinomas. *FEBS Lett.* **2011**, *585*, 671–676.
55. Tano, K.; Mizuno, R.; Okada, T.; Rakwal, R.; Shibato, J. MALAT-1 enhances cell motility of lung adenocarcinoma cells by influencing the expression of motility-related genes. *FEBS Lett.* **2010**, *584*, 4575–4580.
56. Ishii, N.; Ozaki, K.; Sato, H.; Mizuno, H.; Saito, S.; Takahashi, A.; Miyamoto, Y.; Ikegawa, S.; Kamatani, N.; Hori, M.; *et al.* Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction. *J. Hum. Genet.* **2006**, *51*, 1087–1099.
57. Broadbent, H.M.; Peden, J.F.; Lorkowski, S.; Goel, A.; Ongen, H.; Green, F.; Clarke, R.; Collins, R.; Franzosi, M.G.; Tognoni, G.; *et al.* PROCARDIS consortium. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. *Hum. Mol. Genet.* **2008**, *17*, 806–814.

58. Cabianca, D.S.; Casa, V.; Bodega, B.; Xynos, A.; Ginelli, E.; Tanaka, Y.; Gabellini, D. Long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* **2012**, *149*, 819–831.
59. Van Dijk, M.; Thulluru, H.K.; Mulders, J.; Michel, O.J.; Poutsma, A. Windhorst, HELLP babies link a novel lincRNA to the trophoblast cell cycle. *J. Clin. Invest.* **2012**, *122*, 4003–4011.
60. Mercer, T.R.; Dinger, M.E.; Sunkin, S.M.; Mehler, M.F.; Mattick, J.S. Specific expression of long noncoding RNAs in the mouse brain. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 716–721.
61. Ponjavic, J.; Oliver, P.L.; Lunter, G.; Ponting, C.P. Genomic and transcriptional co-localization of protein-coding and long non-coding RNA pairs in the developing brain. *PLoS Genet.* **2009**, *5*, e1000617.
62. Mercer, T.R.; Qureshi, I.A.; Gokhan, S.; Dinger, M.E.; Li, G.; Mattick, J.S.; Mehler, M.F. Long noncoding RNAs in neuronal-glial fate specification and oligodendrocyte lineage maturation *BMC Neurosci.* **2010**, *11*, 14.
63. Fantes, J.; Ragge, N.K.; Lynch, S.A.; McGill, N.I.; Collin, J.R.; Howard-Peebles, P.N.; Hayward, C.; Vivian, A.J.; Williamson, K.; van Heyningen V.; *et al.* Mutations in *SOX2* cause anophthalmia. *Nat. Genet.* **2003**, *33*, 461–463.
64. Amaral, P.P.; Neyt, C.; Wilkins, S.J.; Askarian-Amiri, M.E.; Sunkin, S.M.; Perkins, A.C.; Mattick, J.S. Complex architecture and regulated expression of the Sox2ot locus during vertebrate development. *RNA* **2009**, *15*, 2013–2027.
65. Tochitani, S.; Hayashizaki, Y. Nkx2.2 antisense RNA overexpression enhanced oligodendrocytic differentiation. *Biochem. Biophys. Res. Commun.* **2008**, *372*, 691–696.
66. Brosius, J. RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. *Gene* **1999**, *238*, 115–134.
67. Kondrashov, A.V.; Kiefmann, M.; Ebnet, K.; Khanam, T.; Muddashetty, R.S.; Brosius, J. Inhibitory effect of naked neural BC1 RNA or BC200 RNA on eukaryotic *in vitro* translation systems is reversed by poly(A)-binding protein (PABP). *J. Mol. Biol.* **2005**, *353*, 88–103.
68. Lin, D.; Pestova, T.V.; Hellen, C.U.; Tiedge, H. Translational control by a small RNA: Dendritic BC1 RNA targets the eukaryotic initiation factor 4A helicase mechanism. *Mol. Cell. Biol.* **2008**, *28*, 3008–3019.
69. Martignetti, J.A.; Brosius, J. BC200 RNA: A neural RNA polymerase III product encoded by a monomeric Alu element. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11563–11567.
70. French, P.J.; Bliss, T.V.; O'Connor, V. Ntab, a novel non-coding RNA abundantly expressed in rat brain. *Neuroscience* **2001**, *108*, 207–215.
71. Bernard, D.; Prasanth, K.V.; Tripathi, V.; Colasse, S.; Nakamura, T.; Xuan, Z.; Zhang, M.Q.; Sedel, F.; Jourdren, L.; Coulpier. F.; *et al.* A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. *EMBO J.* **2010**, *29*, 3082–3093.
72. Williams, C.J. Angelman syndrome scientific symposium on the structure and function of UBE3A/E6AP. *Child Neurol.* **2009**, *24*, 904–908.
73. Chamberlain, S.J.; Brannan, C.I. The Prader–Willi syndrome imprinting center activates the paternally expressed murine Ube3a antisense transcript but represses paternal Ube3a. *Genomics* **2001**, *73*, 316–322.

74. Johnstone, K.A.; DuBose, A.J.; Futtner, C.R.; Elmore, M.D.; Brannan, C.I.; Resnick, J.L. A human imprinting centre demonstrates conserved acquisition but diverged maintenance of imprinting in a mouse model for Angelman syndrome imprinting defects. *Hum. Mol. Genet.* **2006**, *15*, 393–404.
75. Yamasaki, K.; Joh, K.; Ohta, T.; Masuzaki, H.; Ishimaru, T.; Mukai, T.; Niikawa, N.; Ogawa, M.; Wagstaff, J.; Kishino, T. Neurons but not glial cells show reciprocal imprinting of sense and antisense transcripts of Ube3a. *Hum. Mol. Genet.* **2003**, *15*, 837–847.
76. Khalil, A.M.; Faghihi, M.A.; Modarresi, F.; Brothers, S.P.; Wahlestedt, C. A novel RNA transcript with antiapoptotic function is silenced in fragile X syndrome. *PLoS One* **2008**, *3*, e1486.
77. Ladd, P.D.; Smith, L.E.; Rabaia, N.A.; Moore, J.M.; Georges, S.A.; Hansen, R.S.; Hagerman, R.J.; Tassone, F.; Tapscott, S.J.; Filippova, G.N. An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals. *Hum. Mol. Genet.* **2007**, *16*, 3174–3187.
78. Moseley, M.L.; Zu, T.; Ikeda, Y.; Gao, W.; Mosemiller, A.K.; Daughters, R.S.; Chen, G.; Weatherspoon, M.R.; Clark, H.B.; *et al.* Ranum Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nat. Genet.* **2006**, *38*, 758–769.
79. Daughters, R.S.; Tuttle, D.L.; Gao, W.; Ikeda, Y.; Moseley, M.L.; Ebner, T.J.; Swanson, M.S.; Ranum, L.P. RNA gain-of-function in spinocerebellar ataxia type 8. *PLoS Genet.* **2009**, *5*, e1000600.
80. Scahill, R.I.; Wild, E.J.; Tabrizi, S.J. Biomarkers for Huntington's disease: An update. *Expert Opin. Med. Diagn.* **2012**, *6*, 371–375.
81. Johnson, D.S.; Mortazavi, A.; Myers, R.M.; Wold, B. Genome-Wide mapping of *in vivo* protein-DNA interactions. *Science* **2007**, *316*, 1497–1502.
82. Pollard, K.S.; Salama, S.R.; Lambert, N.; Lambot, M.A.; Coppens, S.; Pedersen, J.S.; Katzman, S.; King, B.; Onodera, C.; *et al.* An RNA gene expressed during cortical development evolved rapidly in humans. *Nature* **2006**, *443*, 167–172.
83. Johnson, R.; Richter, N.; Jauch, R.; Gaughwin, P.M.; Zuccato, C.; Cattaneo, E.; Stanton, L.W. The human accelerated region 1 noncoding RNA is repressed by REST in Huntington's disease. *Physiol. Genomics* **2010**, *41*, 269.
84. Faghihi, M.A.; Modarresi, F.; Khalil, A.M.; Wood, D.E.; Sahagan, B.G.; Morgan, T.E.; Finch, C.E.; St Laurent, G.; Kenny, P.J.; Wahlestedt, C. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat. Med.* **2008**, *14*, 723–730.
85. Mus, E.; Hof, P.R.; Tiedge, H. Dendritic BC200 RNA in aging and in Alzheimer's disease. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10679–10684.
86. Chubb, J.E.; Bradshaw, N.J.; Soares, D.C.; Porteous, D.J.; Millar, J.K. The DISC locus in psychiatric illness. *Mol. Psychiatr.* **2008**, *13*, 36–64.
87. Millar, J.K.; Wilson-Annan, J.C.; Anderson, S.; Christie, S.; Taylor, M.S.; Semple, C.A.; Devon, R.S.; St Clair, D.M.; Muir, W.J.; Blackwood, D.H.; *et al.* Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum. Mol. Genet.* **2000**, *9*, 1415–1423.

88. Williams, J.M.; Beck, T.F.; Pearson, D.M.; Proud, M.B.; Cheung, S.W.; Scott, D.A. A 1q42 deletion involving DISC1, DISC2, and TSNAX in an autism spectrum disorder. *Am. J. Med. Genet.* **2009**, *149A*, 1758–1762.
89. Brandon, N.J.; Millar, J.K.; Korth, C.; Sive, H.; Singh, K.K.; Sawa, A. Understanding the role of DISC1 in psychiatric disease and during normal development. *J. Neurosci.* **2009**, *29*, 12768–12775.
90. Rapicavoli, N.A.; Blackshaw, S. New meaning in the message: Noncoding RNAs and their role in retinal development. *Dev. Dyn.* **2009**, *238*, 2103–2114.
91. Barry, G.; Briggs, J.A.; Vanichkina, D.P.; Poth, E.M.; Beveridge, N.J.; Ratnu, V.S.; Nayler, S.P.; Nones, K.; Hu, J.; Bredy, T.W.; *et al.* The long non-coding RNA Gomafu is acutely regulated in response to neuronal activation and involved in schizophrenia-associated alternative splicing. *Mol. Psychiatr.* **2013**, doi:10.1038/mp.2013.45.
92. Wood, M.; Yin, H.; McClorey, G. Modulating the expression of disease genes with RNA-based therapy. *PLoS Genet.* **2007**, *3*, e109.

Reprinted from *IJMS*. Cite as: Erriquez, D.; Perini, G.; Ferlini, A. Non-Coding RNAs in Muscle Dystrophies. *Int. J. Mol. Sci.* **2013**, *14*, 19681-19704.

Review

Non-Coding RNAs in Muscle Dystrophies

Daniela Erriquez [1], Giovanni Perini [1,2,*] and Alessandra Ferlini [3,*]

[1] Department of Pharmacy and Biotechnology, University of Bologna, Bologna 40126, Italy; E-Mail: daniela.erriquez@hotmail.it

[2] Health Sciences and Technologies–Interdepartmental Center for Industrial Research, University of Bologna, Bologna 40064, Italy

[3] Section of Microbiology and Medical Genetics, Department of Medical Sciences, University of Ferrara, Ferrara 44100, Italy

***** Authors to whom correspondence should be addressed; E-Mails: giovanni.perini@unibo.it (G.P.); fla@unife.it (A.F.); Tel.: +39-051-2094167 (G.P.); +39-0532-974404 (A.F.); Fax: +39-051-2094286 (G.P.); +39-0532-974406 (A.F.).

Received: 5 August 2013; in revised form: 5 September 2013 / Accepted: 9 September 2013 / Published: 30 September 2013

Abstract: ncRNAs are the most recently identified class of regulatory RNAs with vital functions in gene expression regulation and cell development. Among the variety of roles they play, their involvement in human diseases has opened new avenues of research towards the discovery and development of novel therapeutic approaches. Important data come from the field of hereditary muscle dystrophies, like Duchenne muscle dystrophy and Myotonic dystrophies, rare diseases affecting 1 in 7000–15,000 newborns and is characterized by severe to mild muscle weakness associated with cardiac involvement. Novel therapeutic approaches are now ongoing for these diseases, also based on splicing modulation. In this review we provide an overview about ncRNAs and their behavior in muscular dystrophy and explore their links with diagnosis, prognosis and treatments, highlighting the role of regulatory RNAs in these pathologies.

Keywords: microRNAs (miRNAs); long non-coding RNAs (lncRNAs); Duchenne muscular dystrophy (DMD); Becker muscular dystrophy (BMD); Myotonic dystrophies (DM1 and DM2); Facioscapulohumeral dystrophy (FSHD)

---

1. Introduction

Transcription of the eukaryotic genome yields only 1%–2% of protein coding transcripts and the remainder is classified as non-coding RNAs (ncRNAs). In other words, non-coding RNAs are the main output of the global transcription process, highlighting the idea that such an intense cellular effort cannot be just simple noise. Rather, it is reasonable to speculate that this underscored transcriptome possesses specific vital functions [1–3].

In general, non-coding RNAs are divided into structural and regulatory RNAs. The first ones include ribosomal, transfer, small nuclear and small nucleolar RNAs (rRNAs, tRNAs, snRNAs and snoRNAs respectively), which have been deeply characterized at the functional level. The second ones are a very broad class of RNAs whose main categorization essentially relies on their length.

Small ncRNAs are defined as transcripts shorter than 200 nucleotides. The most functionally characterized are microRNAs (miRNAs), piwi-interacting RNAs (piRNAs) and small interfering RNAs (siRNAs), which are critical for the assembly and the activity of the RNA interference machinery.

RNAs longer than 200 nucleotides are named long non-coding RNAs (lncRNAs) and are a very heterogeneous group of molecules. Because there is not an official way to classify them, they can be placed in one or more categories depending on their genome localization and/or on their orientation (sense, antisense, bidirectional, intronic or intergenic lncRNAs) [4,5].

In the past years, several reports have increased our knowledge about additional levels of regulation of many physiological processes that are mediated by ncRNAs. Even more interesting, these flexible molecules have been found to be dysregulated in many pathological human disorders.

In this review, we will focus on RNAs involved in human skeletal muscle dystrophies. There is a continuous flow of new scientific reports that underpin functional links between ncRNAs and skeletal muscle biology, suggesting that these molecules can play a crucial function both in physiological muscle development and in pathological muscle disorders.

Muscular dystrophies (MDs) are strictly inherited conditions recognized as a common pathogenic mechanism of disruption/impairment of the muscle cell membrane (sarcolemma) which causes a cascade of pathogenic events, including: inflammation, cell necrosis and cell death with progressive fibrosis replacing the muscle mass. MDs represent diseases of extraordinary interest both in medical genetics and biology. The high translational value of research about MDs has recently driven scientific findings toward precise genetic diagnoses as well as novel therapies [6–9].

There are more than 30 different types of inherited dystrophies that are characterized by muscle wasting and weakness of variable distribution and severity, manifesting at any age from birth to middle years, resulting in mild to severe disability and even short life expectancy in the worse cases. Clinical and pathological features are generally the parameters to classify the most common type of MDs. The broad spectrum of MDs arises from many different genetic mutations that reflect defects not only in structural proteins, but also in signaling molecules and enzymes. Dystrophin was the first mutant structural protein shown to cause MD. Mutations in the dystrophin gene lead to two more common type of dystrophy: the severe Duchenne muscular dystrophy (DMD OMIM 300677) due to out-of-frame mutations, and the milder Becker muscular dystrophy (BMD OMIM 300376) associated with in-frame mutations. Some “exceptions to the reading frame rule” are associated with intermediate

phenotypes. The genetic causes of the highly heterogeneous Limb Girdle Muscular Dystrophies (LGMDs) reside in many genes (such as α, β, γ, δ, and ε sarcoglycans) encoding for structural proteins that are part of the complex sarcolemma network and deeply involved, together with dystrophin, in force transduction [10,11].

There are other muscular dystrophies, such as the Facioscapulohumeral muscular dystrophy (FSHD OMIM 158900) and Myotonic dystrophies (DM1 and DM2, see below), that are due to mutations in genes with a main regulatory function. FSHD is due to deletions in non-coding RNA which cause modification of the chromatin assembly in the 4q34 chromosomal region; Myotonic dystrophies (DMs) are related to trinucleotide (DM1) and tetranucleotide (DM2) repeat expansions that produce toxic mutant mRNA with subsequent interference of RNA-splicing mechanisms [12,13].

Many lines of evidence reveal that aberrant expression levels of non-coding RNAs can result in novel types of defects that cause remarkable changes in processes such as mRNA maturation, translation, signaling pathways or gene regulation. To date, it is clear that there is involvement of several miRNAs in the muscular dystrophies, on the contrary, very little is known about the role of long ncRNAs [14].

In this review we try to recapitulate the emerging studies about this intriguing category of molecules, summarizing what is known in muscle, both in physiological and in pathological contexts; new insights are revealing that they are important players in processes such as cellular lineage commitment, growth and differentiation of skeletal muscle. Since muscle differentiation and regeneration are key features that require to be considered when designing novel therapies, addressing the role of ncRNAs in MDs is of high clinical relevance.

2. Muscle-Specific and Ubiquitously Expressed miRNAs in Skeletal Muscle

microRNAs control the stability and/or the translational efficiency of target messenger RNAs, thus causing post-transcriptional gene silencing. Mammalian miRNAs are transcribed as long primary transcripts (pri-miRNAs) and encode one or more miRNAs. Pri-miRNAs are processed by RNase III Drosha in the nucleus to generate stem-loop structures of ~70 nucleotides (pre-miRNAs) and then exported to the cytoplasm where they are further processed by RNase Dicer to yield ~22 bp mature miRNAs. A mature miRNA, incorporated into the RNA-induced silencing complex (RISC), anneals to the 3' UTRs of its target mRNAs by its complementary strand, thus causing post-transcriptional gene silencing via translational repression or mRNA degradation. In the last years new paradigms of miRNA biogenesis are also emerging in which the processing of miRNA does not require all steps mentioned above [15].

Vertebrate skeletal muscle is derived from the somites, the first metameric structures in mammalian embryos, that progressively subdivide into embryonic compartments, thus giving rise to dermomyotome and subsequently to myotome to produce differentiated muscular tissue. The process of generating muscle—myogenesis—is highly complex and requires a broad spectrum of signaling molecules, either during embryonic development and in postnatal life, that converges on specific transcription and chromatin-remodeling factors, as well as on regulatory RNAs, to activate gene and microRNA expression program [16,17].

The fate of myogenic precursor cells first determined by paired -homeodomain transcription factors, Pax3/Pax7, followed by regulation of highly conserved MyoD (also named MyoD1, myogenic differentiation 1), Myf5 (myogenic factor 5), MyoG (myogenin), and MRF4 factors, expressed in the skeletal muscle lineage and therefore referred as myogenic regulatory factors (MRFs). The MRFs differ in the timing and the stages of myogenesis, reflecting their different roles during muscle cell commitment and differentiation. MyoD and Myf5 are both considered markers of terminal commitment to muscle fate. Myf5 is the first MRF expressed during the formation of the myotome, followed by expression of MyoD. Specifically, in the majority of muscle progenitors, MyoD functions downstream from Pax3 and Pax7 in the genetic hierarchy of myogenic regulators, whereas Myf5, depending on the context, can also act in parallel with the Pax transcription factors [18–20]. Instead, MyoG and MRF4 act subsequently to specify the immature muscle cells (myoblasts) for terminal differentiation. Myoblasts exit from cell cycle after a defined proliferation time, to become terminally differentiated myocytes [21,22]. Muscle-specific genes such as myosin heavy chain genes (*MyHC* genes) and muscle creatine kinase (*M-CK*) are expressed in the last phase of this multi-regulated program, where mononucleated myocytes specifically fuse to each other to form multinucleated myotubes [22–28].

Dicer loss-of function studies clarified the importance of miRNAs in normal skeletal muscle development [29]. miRNAs actively take part in the proliferation and differentiation of skeletal muscle cells as an integral component of genetic regulatory circuitries.

miR-1, miR-133a/b and miR-206 are largely studied and defined muscle-specific miRNAs (myomiRs). They are regulated in muscular transcriptional networks via MRFs and via others key-regulators of the myogenic program, MEF2 (myocyte enhancer factor 2) and SRFs (serum response factors). Recently, a new regulatory pathway, the mechanistic target of rapamycin (mTOR) signaling was seen to regulate miR-1 expression and was also found responsible for MyoD stability [30–36]. It is possible to functionally define miR-133 as enhancer of myoblast proliferation while miR-1 and miR-206 as enhancers of skeletal muscle differentiation [37–40]. An up-to-date list of the identified targets of miR-1, miR-133 and miR-206, together with a plethora of specific muscular pathways they are involved in, is reported in a recent review [40] and some of these will be also discussed in the next paragraph to highlight how these important families of miRNAs contribute to determine typical deficiencies occurring in a pathological muscular context. Intriguingly, these myomiRs have been shown to behave as serum biomarkers in DMD patients. They are released into the bloodstream as a consequence of fiber damage and their power as diagnostic tools is promising since increased miRNA levels correlate with severity of the disease, significantly better than other commonly utilized markers, such as creatine kinase (CK). Moreover, their major serum stability is another aspect that may make them useful not only for diagnosis but also for monitoring the condition of affected individuals after a therapeutic treatment [41,42].

miR-208b/miR-499, also named myomiRs because of their muscle-restricted expression, are produced from the introns of two myosin genes, *β-MHC* and *Myh7b*. They are functionally redundant and play a dominant role in the specification of muscle fiber identity by activating slow and repressing fast myofiber gene programs [43].

Interestingly, many miRNAs are defined as “non-muscle specific” (or also ubiquitously expressed), because essentially they are not exclusively expressed in muscular tissue. It has been, however,

demonstrated that they play key-roles in modulating important pathways involved in the regulation of muscular metabolism and cellular commitment. Many miRNAs fall into this category and we report here a few relevant examples, providing for each miRNA the context in which they were studied and highlighting their global effects on muscular metabolism (Table 1).

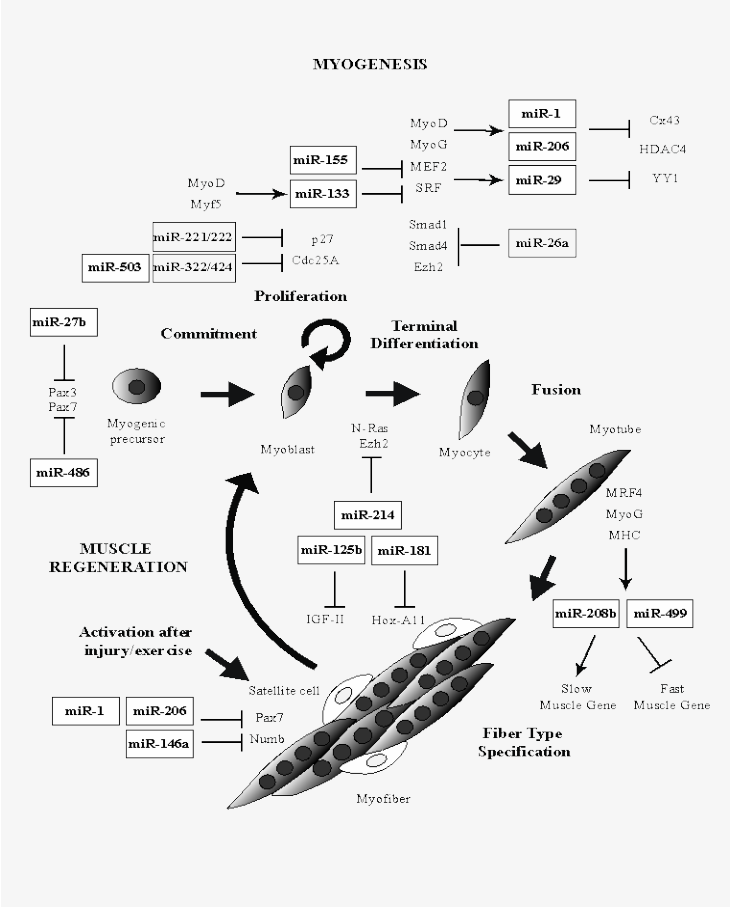
Table 1. miRNAs expressed in muscular tissue (in an exclusive manner or not) and their global effect on muscle metabolism.

miRNA	Role in Muscle Metabolism [Refs.]	Tissue Expression
miR-1	enhancer of skeletal muscle differentiation [37–40]	muscle-specific
miR-133a/b	enhancer of myoblast proliferation [37–40]	muscle-specific
miR-206	enhancer of skeletal muscle differentiation [37–40]	muscle-specific
miR-208b	involved in specification of muscle fiber identity [43]	muscle-specific
miR-499	involved in specification of muscle fiber identity [43]	muscle-specific
miR-24	promotes myoblast differentiation [44]	ubiquitous
miR-26a	promotes myoblast differentiation [45,46]	ubiquitous
miR-27b	promotes entry into differentiation program [47]	ubiquitous
miR-29	enhancer of differentiation [48,49]	ubiquitous
miR-125b	negatively contributes to the myoblast differentiation and muscle regeneration [50–52]	ubiquitous
miR-155	represses myoblast differentiation [53]	ubiquitous
miR-181	regulates skeletal muscle differentiation and regeneration after injury [54]	ubiquitous
miR-146a	promotes satellite cell differentiation [55,56]	ubiquitous
miR-214	promotes cell cycle exit and differentiation [57]	ubiquitous
miR-221/222	promote cell cycle progression [58]	ubiquitous
miR-322/424; miR-503	promote myogenesis interfering with the progression through the cell cycle [59]	ubiquitous
miR-486	positively regulates myoblast differentiation [60,61]	muscle-enriched

Some of these miRNAs counteract the differentiation process since their activity is aimed to positively regulate the proliferation phase during muscular development.

miR-125b, one of the few down-regulated miRNAs during myogenesis, together with miR-221/222, negatively contributes to myoblast differentiation and muscle regeneration, taking part in the regulatory axis that includes mTOR and IGF-II [50–52]. Similarly, miR-155 mediates the repression of differentiation targeting MEF2A, a member of MEF2 family of transcription factors. By this negative regulation, miR-155 functions as an important regulator of muscle gene expression and myogenesis [53]. miR-221/222 instead are involved in maintenance of the proliferative state promoting cell cycle progression. They are under control of the Ras-MAPK axis and inhibit the cell-cycle regulator p27 (Cdkn1b/Kip1). Their ectopic expression, indeed, lead to defects in the transition from myoblasts to myocytes and in the assembly of sarcomeres in myotubes [58].

Figure 1. Overview of muscle-specific and ubiquitously expressed miRNAs that contribute to myogenesis and muscle regeneration processes and their regulatory activity on the muscular specific targets/chromatin modifying enzymes/cell cycle regulators (for details see the text). The main regulatory factors that exert a fundamental role during each step of normal muscle development are also reported as well as their eventual regulatory activity on the described miRNAs.



In contrast to this set of miRNAs, many other "non-muscle specific" miRNAs exert an active role in muscle differentiation through different mechanisms: miR-24, for example, has been shown to be essential for the modulation of transforming growth factor β/bone morphogenetic protein (TGF-β/BMP) pathway, a well-known inhibitor of differentiation, although its specific muscular targets are yet unknown [44]; miR-26a is involved in TGF-β/BMP pathway, where it negatively regulates the transcription factors Smad1 and Smad4, critical components of that signaling; miR26a targets the polycomb complex member Ezh2, involved in chromatin silencing of skeletal muscle genes [45,46]; miR-27b promotes entry into differentiation program both *in vitro* and *in vivo* regenerating muscles by down-regulating Pax3 [47]; miR-29 in general is defined as an enhancer of differentiation. During myogenesis it is up-regulated by SRFs and MEF2, and in a self-regulatory manner, it suppresses YY1 and HDAC4 translation by targeting their 3'-UTRs [48,49]; miR-146a is another positive regulator of myogenesis, since it modulates the activity of NUMB protein, which promotes satellite cell differentiation towards muscle cells by inhibiting Notch signaling [55,56]; miR-181 is involved in skeletal muscle differentiation and regeneration after injury and one of its targets is Hox-A11, which in turn represses transcription of MyoD [54]; miR-214 was identified in

zebrafish as regulating the muscle development. Here it is expressed in skeletal muscle cell progenitors and was shown to specify muscle cell type during somitogenesis by modulating the response of muscle progenitors to Hedgehog proteins signaling [57]. Its involvement in muscle is also confirmed in C2C12 myoblasts and in skeletal myofibers of mouse where it promotes cell cycle exit and thus differentiation, targeting proto-oncogene N-*Ras* and the repressor of myogenesis Ezh2 respectively [62,63]; miR-322/424 and -503 promote myogenesis interfering with the progression through the cell cycle [59]; while miR-486 was reported to positively regulate myoblast differentiation targeting phosphatase and tensin homolog (PTEN) and Foxo1a, which negatively affect phosphoinositide-3-kinase (PI3K)/Akt signaling and down-regulate the transcription factor Pax7, required only for muscle satellite cell biogenesis and for specification of the myogenic precursor lineage [60,61]. All these data clearly show the vast scenario of functions in which miRNAs are involved and their specific activities they play in the skeletal muscle physiology (Figure 1).

3. miRNAs in Muscular Dystrophies

Muscle is a dynamic tissue that goes through many recurrent phases of degeneration and regeneration throughout an individual's lifetime. During normal muscle development, specific molecular circuitries and signaling pathways control several events in different cell types such as activation of satellite cell proliferation, progenitor cell maintenance, myoblast differentiation, muscle cell homeostasis and immune cell recruitment. It is therefore not surprising that their deregulation heavily contributes to the degeneration of dystrophic muscles and is the object of intense research [64] (Table 2).

Table 2. miRNAs found deregulated in MDs and their specific activity on muscular targets or involvement in muscular processes.

miRNA/miRNAs	Deregulated in MDs [References]	Type of Deregulation	Muscular Targets/Process [References]
miR-1 (myomiR)	DMD [65,66]; DM1 [67,68]	down-regulated	HDAC4; Cx43; Pax7; c-Met; G6PD [40]
miR-133 (myomiR)	DMD [66]	down-regulated	SRF; nPTB; UCP2 [40]
miR-206 (myomiR)	DMD [65]; DM1 [69]	up-regulated	DNApolα; Fstl1; Utrn; Pax7; Cx43; HDAC4; c-Met [40]
miR-29b/c	DMD [65,66]; DM1 [67]	down-regulated	YY1; Col1a1; Eln; HDAC4 [40,62,63]
miR-135a	DMD [65]	down-regulated	muscle degeneration [65]
miR-30c	DMD [66]	down-regulated	-
miR-31	DMD [65,70]	up-regulated	DMD [70]
miR-34c; miR-449; miR-494	DMD [65]	up-regulated	muscle regeneration [65]
miR-146b; miR-155	DMD; BMD; LGMD; FSHD [71]	up-regulated	-; MEF2A [53]
miR-214	DMD; BMD; LGMD; FSHD [71]	up-regulated	Ezh2; N-Ras [40,62,63]

Table 2. *Cont.*

miRNA/miRNAs	Deregulated in MDs [References]	Type of Deregulation	Muscular Targets/Process [References]
miR-221; miR-222	DMD; BDM; LGMD; FSHD [71]	up-regulated	p27(Cdkn1b/Kip1) [58]; Sntb1 [72]
miR-223	DMD [65]	up-regulated	muscle inflammation [65]
miR-335	DMD [65]; DM1 [67]	up-regulated	muscle regeneration [65]
miR-33	DM1 [67]	down-regulated	-
miR-34a-5p; miR-34b-3p; miR-34c-5p; miR-146b-5p; miR-208a; miR-221-3p; miR-381	DM2 [73]	up-regulated	-
miR-125b-5p; miR-193a-3p; miR-193b-3p; miR-378a-3p	DM2 [73]	down-regulated	-

Eisenberg *et al.* analyzing 10 primary muscular disorders (including DMD, BDM, LGMD and FSHD samples) have identified five miRNAs (miR-146b, miR-221, miR-155, miR-214, and miR-222) consistently deregulated in almost all samples taken into consideration, suggesting their involvement in common regulatory mechanisms. Other miRNAs however showed a disease-specific profile. Functional correlation between miRNAs and mRNA targets in DMD biopsies draw a tight posttranscriptional regulation network in secondary response functions and in muscle regeneration [71].

Greco and coworkers have divided a DMD-signature of miRNAs into three main classes relative to their functional link to specific muscular pathway. Regeneration-miRNAs were up-regulated (miR-31, miR-34c, miR-206, miR-335, miR-449, and miR-494), while degenerative-miRNAs (miR-1, miR-29c, and miR-135a) were down-regulated in *mdx* mice and in DMD patients' muscles. The third class are named inflammatory-miRNAs, (miR-222 and miR-223), being expressed in damaged muscle areas only [65].

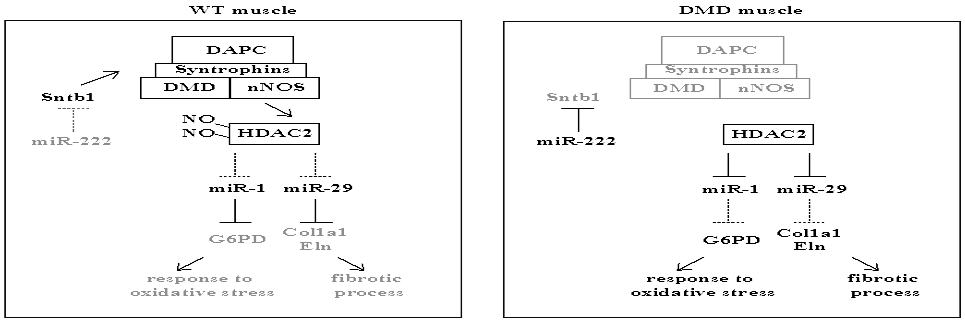
Muscle specific myomiR miR-1 and miR-133 and the ubiquitous miR-29c and miR-30c are down-regulated in *mdx* mice. It is possible to restore WT levels of these miRNAs by treating animals with an exon-skipping approach to restore a partially functional dystrophin protein, an experimental strategy that overcomes an out-frame mutation in the *DMD* locus. The same results are confirmed also in human DMD samples. These results corroborate the direct correlation between miRNAs levels and dystrophin protein levels. In contrast with the other myomiRs, miR-206 shows an increased expression in distrophic *mdx* muscle because it activates satellite cell differentiation program through Pax7 and HDAC4 repression. Another interesting target of miR-206 is Utrophin (Utrn), a dystrophin protein homolog, involved in a compensatory mechanism in DMD pathology [31,66,74].

miR-31-repressing activity seems to regulate muscle terminal differentiation directly targeting the 3'-UTR of dystrophin. Also miR-31, as miR-206, has a preferential localization in regenerating myoblasts, and is highly expressed in Duchenne muscles, probably due to an intensified activation of satellite cells. In both human and murine wild-type conditions its expression is detected in early phases of myoblast differentiation, supporting the idea that it contributes to avoid early expression of late differentiation markers. For this reason it is linked to a delay in the maturation program occurring in the pathological context [70].

Dystrophin is a structural protein that links the cytoskeleton to a large membrane-associated multiprotein complex (dystrophin-associated protein complex, DAPC) to stabilize the sarcolemma. Via

Syntrophins (SNTA1, SNTB1, SNTB2, SNTG2), members of DAPC, the enzyme neuronal Nitric Oxide Synthase (nNOS) is localized to the membrane of muscle fibers and regulates intramuscular generation of nitric oxide (NO) [75–77]. nNOS signaling determines the status of nitrosilation of Histone Deacetilases (HDACs) and thus their chromatin association to muscular specific gene-targets. Upon myoblast differentiation, HDACs are displaced from chromatin to promote muscle-specific gene transcriptional activation [78,79]. Some miRNAs involved in DMD pathology have been recently discovered to undergo this type of transcriptional regulation [66]. The absence of dystrophin in DMD patients and *mdx* mice leads to a dramatic decrease of DAPC and a consequential impairment of NO production [80,81]. The expression of a specific subset of miRNAs is modulated by HDAC2 via Dystrophin/nNOS pathway. In particular the activation of both human and murine miR-1 and miR-29 is tightly linked to HDAC2 release from their respective promoters. The functional role of these two miRNAs in muscular metabolism is also been highlighted. miR-1 controls Glucose-6-phosphate dehydrogenase (G6PD), a relevant enzyme involved in the response to oxidative stress while miR-29 controls fibrotic process since it targets the structural component of extracellular matrix, collagen (Col1a1) and elastin (Eln). Moreover, miR-222 targeting β1-Syntrophin (Sntb1) may also contribute to deregulation of the Dystophin-Syntrophins-nNOS pathway [72] (Figure 2).

Figure 2. Schematic representation of the functional/physical relationship between Dystrophin-Syntrophins-nNOS pathway and miRNAs involved in such signaling, both in a WT and a DMD context.



Myotonic dystrophy (DM) is the most common adult onset, progressive muscular dystrophy. DM is a multi-systemic disease and it is characterized by a generalized muscle weakness and wasting, associated with peripheral neuropathy, heart rhythm defects, and cataracts. The myotonia phenomenon

is due to the peculiar muscle membrane depolarization activities. Two type of DM exist, type-1 (DM1, OMIM 160900) and type-2 (DM2, OMIM 602668). DM1 is caused by an expansion of the CTG triplet repeats in the 3'-untraslated region (UTR) of the Dystrophic Myotonic Protein Kinase (*DMPK*), while DM2 is caused by the expansion of a tetranucleotide repeat CCTG in the first intron of CCHC-type zinc finger nucleic acid binding protein (*CNPB*). These gene expansions do not disrupt the relative protein coding sequence, the repeats being in non-coding regions. However, both expanded RNAs accumulate in the nucleus and trigger a toxic gain of function that interferes with RNA splicing of other genes [82–86]. Perbellini and colleagues have performed expression analysis in DM1 biopsies obtained from 15 patients. They found specific deregulated miRNAs: miR-1 and -335 are up-regulated, whereas miR-29b, -29c and -33 are down-regulated compared to control muscles [67,68]. Gambardella and co-workers profiled a specific pattern of myomiRs involved in myogenesis of cardiac and skeletal muscle and found lines of evidence of miR-206 overexpression in five DM1 patients [69]. A similar investigation has been made in DM2 patients. Eleven miRNAs have been shown to be deregulated. Nine displayed higher levels compared to controls (miR-34a-5p, miR-34b-3p, miR-34c-5p, miR-146b-5p, miR-208a, miR-221-3p and miR-381), while four were decreased (miR-125b-5p, miR-193a-3p, miR-193b-3p and miR-378a-3p). Moreover the potential involvement of these miRNAs in relevant skeletal muscle pathways and functions has been validated by bioinformatics analyses [73]. Recently a novel therapeutic approach has been proposed to target the CTG repeat expansion on RNA using antisense oligonucleotides [8,87,88]. Therefore improving knowledge concerning the transcription regulation of the *DMPK* gene, also via ncRNAs, will greatly benefit this new therapy.

4. Long Non-Coding RNAs in Skeletal Muscle and Muscular Dystrophies

Increasing lines of evidence support the biological relevance of lncRNAs. They are regulated during development and involved in almost all levels of gene expression and cellular functions including chromosomal dosage compensation, chromatin modification, cell cycle regulation, control of imprinting, alternative splicing, intracellular trafficking, cellular differentiation, and reprogramming of stem cells [89]. Recently, lncRNAs related to muscle are emerging both in physiological and pathological context (Table 3).

Key features of dystrophic muscle include central nuclei, small regenerating fibers and accumulation of connective tissue and fatty tissue. Muscle differentiation *in vitro* is a useful system to investigate the activity of long non-coding RNAs that show muscular specific pattern of expression. Recently, a new regulatory network involving cross-talk of several ncRNAs has been identified by Cesana and colleagues. Relying on ability of myomiRs to orchestrate muscular proliferation and differentiation, the genomic region of miR-206/-133b has been analyzed in detail. Thus a novel muscle specific transcript has been identified. Because of its non-coding potential and its activated expression upon myoblast differentiation it was termed linc-MD1. More specifically linc-MD1 is expressed in newly regenerating fibers and is abundant in dystrophic condition, however no expression is detected in mature differentiated fibers. linc-MD1 is localized in the cytoplasm and is a polyadenylated transcript. Through a series of functional studies it was possible to define its competing endogenous activity (ceRNA). linc-MD1 acts as a natural decoy for miR-133 and -135, thus interfering with miRNA repressing activity on the important targets involved in myogenic differentiation MAML1 (Mastermind-like 1) and MEF2, respectively [90].

Table 3. Recently discovered lncRNAs related to muscle, both in physiological and pathological context.

lncRNA/lncRNAs [References]	Expression in Muscular Districts	Deregulated in MDs	Activity
linc-MD1 [90]	Expressed in newly regenerating fibers	DMD	natural decoy for miR-133 and -135 (ceRNA)
Malat1 [91]	up-regulated during the differentiation of myoblasts into myotubes	?	regulation of cell growth
Men ε/β lncRNAs [92–94]	up-regulated upon differentiation of C2C12 myoblats	?	critical structural/organizational components of paraspeckles
SRA ncRNA [95–97]	increased expression during myogenic differentiation	DM1	co-activator of MYOD transcription factor
NRON [98,99]	enriched also in muscle	?	regulates NFAT's subcellular localization (scaffold)
lncINT44s; lncINT44s2; lncINT55s [14]	transcribed contextually with dystrophin isoforms and upon MYOD-induced myogenic differentiation	?	negative modulation of endogenous dystrophin full-length isoforms
KUCG1 [100]	expressed at low levels in the brain	DMD with mental retardation	possible candidate gene that contribute to develop of mental retardation in the index case
DBT-E [101]	not-physiological lncRNA	FSHD	coordinates de-repression of genes located in the 4q35 region

Metastasis associated lung adenocarcinoma transcript 1 (Malat1) is a highly conserved 8.7 kb non-coding transcript that is abundantly expressed in cancer cells and a strong predictor of metastasis [102]. Malat1 has been proposed to regulate alternative splicing [103], transcriptional activation and the expression of nearby genes [104,105]. Numerous experimental examples support its functional role in the regulation of cell growth, but the exact mechanism of action of Malat1 in different physiological and pathological conditions still needs to be elucidated. By a microarray data analysis obtained using skeletal muscle of mice (gastrocnemius muscle) treated with recombinant myostatin it was observed that the Malat1 expression levels are significantly decreased. Myostatin is a potent negative regulator of myogenesis that inhibits myoblast proliferation and differentiation [106,107]. Further expression analysis confirmed a persistent up-regulation of Malat1 during the differentiation of myoblasts into myotubes in C2C12 cells as well as in primary human skeletal muscle cells. Conversely, targeted knockdown of Malat1 using siRNA suppressed myoblast proliferation by arresting cell growth in the G0/G1 phase. These results reveal Malat1 as a novel downstream target of myostatin with a considerable ability to regulate myogenesis. Although Malat1 appears largely dispensable for normal mouse development [108,109] it is plausible that Malat1 has a role in the transition from the proliferative phase to differentiation in skeletal myogenesis, as well as in the commitment to muscle differentiation [91].

Many lncRNA have been discovered but not yet fully characterized, as for example Men ε/β lncRNAs. To date it is known that two long non-coding isoforms (Men ε/β lncRNAs) which are expressed in several human tissues, including muscle, arise from the Multiple Endocrine Neoplasia I locus (*MEN1*). Experimental lines of evidence show their up-regulation upon differentiation of C2C12

myoblats, although their biological role in muscular development is not yet clear. Men ε (also known as NEAT1) and Men β are transcribed from the same RNA polymerase II promoter and are both retained in the nucleus. Suwoo and colleagues formally demonstrated that Men ε/β transcripts are critical structural/organizational components of paraspeckles, organelles localized in the nucleoplasm close to nuclear speckles, where RNA-binding proteins and *Cat2*-transcribed nuclear RNA (CTN-RNA) are stored [110]. Moreover, large-scale analysis revealed that many other lncRNAs are differentially expressed in C2C12 cells upon myoblast differentiation into myotubes, although their biological functions have not been investigated [92–94].

Between the many functions ascribed to lncRNA there are examples of lncRNAs modulating the activity of transcriptional activators or co-activators, directly or through the regulation of their sub-cellular localization [89]. Two of these have been seen also in a muscular context. The steroid receptor RNA activator (SRA) RNA is a very peculiar transcript that exists as both a non-coding and a coding RNA (yielding SRA ncRNA and protein SRAP respectively). The SRA ncRNA is highly expressed in skeletal muscle and works as a co-activator of MYOD transcription factor, a master regulator of skeletal myogenesis. To address the significance of the enigmatic bifunctional property of this transcript, Hube and colleagues performed an exhaustive analysis clarifying the opposite function of non-protein coding SRA versus ORF-containing transcripts. The balance between coding and non-coding SRA isoforms changes during myogenic differentiation in primary human cells. In particular it is shown that an increased expression of SRA ncRNA and a parallel decrease of protein SRAP occurs during myogenic differentiation in healthy muscle satellite cells. This does not happen in cells isolated from DM1 patients, probably because of a delay in differentiation program. Remarkably, only the ncRNA species enhances MYOD transcriptional activity. The protein SRAP prevents this SRA RNA-dependent co-activation through interaction with its RNA counterpart [95–97]. However how this is achieved is not known.

Non-coding repressor of NFAT (NRON) is another case of lncRNA that shows a regulatory activity on a transcription factor. NRON is not highly expressed but it has a distinct tissue specific expression. It has been found enriched in placenta, muscle, and lymphoid tissues. NFAT is a transcription factor responsive to local changes in calcium signals. It is essential for the T cell receptor–mediated immune response and plays a critical role in the development of heart and vasculature, musculature, and nervous tissue. The first study about the role of NRON showed that it regulates NFAT's subcellular localization rather than its transcriptional activity. Sharma and coworkers confirmed these data demonstrating that NRON takes part in a large cytoplasmic RNA-protein complex that acts as a scaffold for NFAT to modulate its nuclear trafficking and thus its response activity [98,99].

Little is yet known about the dystrophin gene regulation. *DMD* is the largest gene in the human genome that comprises 79 exons spanning >2500 kb on chromosome Xp21.2, which gives rise to 7 isoforms that are finely regulated in terms of tissue specificity [111]. Mutations in the *DMD* gene range from single-nucleotide changes to chromosomal abnormalities (http://www.dmd.nl/). Deletions encompassing one or more exons of the dystrophin gene are the most common cause of the severe Duchenne muscular dystrophy (DMD) resulting in an absence of dystrophin or expression of a non-functional protein. Becker muscular dystrophy (BMD) instead is a milder form of dystrophy because it is associated with reduction of wild-type dystrophin or expression of a partially functional protein. DMD is the most common inherited muscle disease affecting approximately one in

3500 males and is characterized by progressive muscle wasting during childhood. Heterozygous females for dystrophin mutations are named carriers of DMD mutations [112,113]. Many of them are asymptomatic, but a certain number, defined as "manifesting" or "symptomatic", develop symptoms of the disease, which vary from a mild muscle weakness to a DMD-like clinical course. Despite intensively explored, the pathogenic mechanism underlying clinical manifestation in DMD female carriers still remains a controversial issue [114]. For these reasons *DMD* regulation is a field of intense interest to shed light on this complex scenario.

Using a custom-made tiling array the entire *DMD* gene has been explored in the search for non-coding transcripts originating within the dystrophin locus. The major tissues of dystrophin synthesis, namely human brain, heart and skeletal muscle, were used as test tissues in array. The data analysis has highlighted a variety of novel long non-coding RNAs (lncRNAs), both sense and antisense oriented, whose expression profiles mirror that of *DMD* gene. Importantly, these transcripts are intronic in origin, specifically localized to the nucleus and are transcribed contextually with dystrophin isoforms or in fibroblast upon MYOD-induced myogenic differentiation. To characterize their possible functional role on the *DMD* locus three sense-oriented lncRNAs (lncINT44s, lncINT44s2 and lncINT55s) isolated from skeletal muscle were further investigated. Their forced ectopic expression in both human muscle and neuronal cells causes a negative regulation of endogenous full-length dystrophin isoforms, denoted B for brain (Dp427b), M for muscle (Dp427m) and P for Purkinje (Dp427p). Importantly, no variation was observed with regard to the ubiquitous Dp71 transcript, suggesting that the effect of sense lncRNAs on full-length dystrophin isoforms may be specific. In particular, reporter assay confirmed their repressive role on the minimal promoter regions of the muscle dystrophin isoform. A possible mechanism of action involves specific DMD lncRNAs that control muscle dystrophin isoforms by down-modulating dystrophin transcription levels. An inverse correlation between ncRNAs expression and muscle dystrophin has been also found *in vivo*, analyzing muscle samples of DMD female carriers, either healthy or mildly affected, reinforcing the idea that a negative relationship between lncRNAs and dystrophin mRNA levels may exist [14].

In severe DMD one third of patients display also mental retardation, but the pathogenesis is unknown. In a singular case of DMD complicated by mental retardation, an intra-chromosomal inversion (inv(X)p21.2;q28) has been identified. The genetic rearrangement has been molecularly characterized to find a possible disrupted gene because of the inversion, and that might be responsible for the neurological symptoms associated with dystrophy. A novel gene named *KUCG1* was discovered at break point on Xq28. The 658-bp transcript displays an mRNA-like structure but not having coding potential is been classified as long non-coding RNA. KUCG1 lncRNA is expressed at low levels in a tissue-specific manner, as well in the brain. It is possible that the disruption of KUCG1 transcript contributes to the development of mental retardation in the index case [115] since other experimental lines of evidence suggest that a subset of lncRNAs could contribute to neurological disorders when they become deregulated [100].

Polycomb (PcG) and Trithorax (TrxG) group proteins antagonistically act in the epigenetic regulation of gene expression. Typically, TrxG counteracts PcG-mediated epigenetic gene silencing. Among the many lncRNAs interacting with chromatin remodeling enzymes the most famous are Xist and HOTAIR, both acting as a negative regulators of gene expression by recruitment of PRC2 (Polycomb Repressive Complex 2) on PcG target genes [116,117]. Cabianca *et al*. were the first to

discover an lncRNA interacting with the TrxG in the Facioscapulohumeral muscular dystrophy (FSHD). FSHD is an autosomal-dominant disease characterized by progressive wasting of facial, upper arm, and shoulder girdle muscles. In up to 95% of cases, the genetic defect is mapped to the subtelomeric region of chromosome 4q35 containing a macrosatellite tandem array of 3.3 Kb long D4Z4 repeats. FSHD is caused by deletions reducing copy number of D4Z4 below 11 units rather than a classical mutation in a coding-protein gene. D4Z4 deletion is associated to a loss of repressive epigenetic marks and thus to a switch from a heterochromatic/close state to a more euchromatic/open conformation of chromatin structure. A novel long non-coding RNA, named DBT-E is produced selectively in FSHD patients. DBT-E is transcribed from D4Z4 repeats and is a chromatin-associated lncRNA that coordinates de-repression of genes located in the 4q35 region. DBT-E recruits the Trithorax group protein Ash1L to the FSHD locus driving histone H3 lysine dimethylation and thus chromatin remodeling [101].

5. Discussion

It is surprising how ncRNAs are tightly interconnected with the main fundamental aspects of muscular tissue: development, differentiation and regeneration. At the molecular level miRNAs and lncRNAs take part in almost all levels of regulation in these key processes. Chromatin modifying enzymes, positive and negative transcription factors, cell cycle regulators, and enzymatic and structural proteins involved in signaling circuitries are under their fine-modulation. Moreover, ncRNAs are often found to be under the regulation of their own targets, thus determining feedback loops that drive developmental switches ensuring a perfect synergy between stimuli and responses. Both time- and tissue-specific gene regulation are the fulcrum on which the fine-tuning of a healthy organism is based. Disrupting the physiological pattern of expression not only in codifying genes, but also in regulatory RNAs, can heavily modify specific cell processes.

If this is true in physiological conditions, increased lines of evidence show that regulatory RNAs play a crucial role also in the etiology of many human diseases. Among these, muscular dystrophies represent a field of intense research, also because of the recent creation of novel experimental treatments. This has encouraged studies on expression regulation in diseases muscle cells, both *in vivo* (animal models) and *in vitro*. These studies have shown that mutant proteins in MDs result in perturbations of many cellular components. Indeed MDs have been associated with mutations in structural proteins, signaling molecules and enzymes as well as mutations that result in aberrant processing of mRNA or alterations in post-translational modifications of proteins. These findings have not only revealed important insights for cell biologists, but have also provided unexpected and exciting new approaches for therapy. Moreover, in muscular dystrophies as well as in other diseases, such as cancer, regulatory RNAs may serve as biomarkers, providing information on disease course, disease severity and response to therapies. miRNA dosing in serum is a very appealing field of investigation since they are easily accessible, peculiar to defined conditions and can facilitate the early identification of the muscular disease, potentially avoiding invasive techniques such as a biopsy, or in some cases to reduce the time and the costs of diagnosis. Biomarkers are particularly important in the field of personalized treatments. Pharmacogenomics aims at predicting which drug will be most effective and safe in the individuals. This can be established via genome sequence and SNP association (pharmacogenetics) and expression profiling. miRNAs have been shown to play a pivotal role in drug

efficacy and toxicity, having powerful implications in personalized medicine [118]. Indeed, as we have described, miRNAs can negatively regulate gene expression and can profile the disease severity, as in the case of myomiRs and DMD [41]. miRNAs show a linear relationship with genes and drugs, since drug function can be influenced or even hampered by changes in genes expression level or in specific isoforms representation, as supported by several data on cancer [119]. Many pharmacogenomically relevant genes are regulated by miRNAs, as summarized and shown in the Pharmacogenomcis Knowledge Base (PharmGKB, www.pharmgkb.org/), a very useful resource listing genes known to be relevant for drug response.

Conversely, miRNAs can vary in their expression level following drug treatments [120]. Within the muscle field, we do have increased knowledge on the miRNAs network, especially those governing the muscle transcriptional network. It is clearly emerging how miRNAs can regulate differentiation and homeostasis of skeletal muscle progenitor cells, providing robustness to the MYOD-induced myoblast differentiation and myogenesis [121,122]. Disclosing the role miRNAs have in regulating the intermediate steps of the myogenesis cascade will be of outmost importance in identifying drugs that may act as adjuvants/enhancers of gene/protein re-synthesis in clinical trials, as for exon skipping therapies in DMD.

More complex is our current understanding of the role of lncRNAs in muscle biology and pathology. We have just started to explore the peripheral areas of this "*terra incognita*". So far lncRNAs have been involved in numerous molecular processes such as remodelling of chromatin architecture, or regulation of gene transcription. For instance, some pharmacodynamic studies on corticosteroids, which represent the gold standard in the routine therapy of DMD, revealed that the steroid receptor RNA activator (SRA) transcript functions as both a lncRNA and template for synthesis of a protein (SRAP). Interestingly, the SRA ncRNA increases the activity of nuclear receptors (not only for corticosteroids) and acts as a master regulator of MYOD expression [95]. lncRNAs can also exert their function through a more passive role. For instance they are valued for their ability to work as molecular sponges by annealing to small RNAs and thereby preventing them from their normal activity. Furthermore, in some other cases lncRNAs have been shown to provide a kind of structural backbone for the assembly of ribonucleic particles whose functions are still to be disclosed. In this respect, it has been crucial to determine in which intracellular compartments these RNA/protein particles form. Despite the fact that so far most investigated lncRNAs are confined to nuclei, a few recent studies have, indeed, shown that some lncRNAs can also abundantly localize inside the cytoplasm with functions that still remain to be determined.

Given that just a few lncRNAs have been tackled on functional levels and that the annotated ones are in the order of thousands with many more expected to be discovered, it is plausible to speculate that their involvement in novel functions and roles will be rapidly identified with repercussions on many fields of cell biology and pathology and with the possibility to potentially employ them as biological markers as well as drugs to treat major diseases such as muscular dystrophies.

6. Conclusions

Although in the majority of cases the etiology of muscular dystrophies is not ascribed to functional non-coding RNA molecules (with the exception of FSHD), they appear as powerful regulators of

several key-pathways and show how actively they can contribute to the progression of disease. This reflects the strong ability of miRNAs and lncRNAs in the modulation of the phenotype of dystrophic affected individuals via fine regulatory pathways that can lead to increased transcript stability, mRNA splicing control, enhanced protein production, posttranslational protein modification and other mechanisms. These versatile roles support the idea to use regulatory RNAs as novel targeted molecules acting as enhancers or inhibitors in well-established therapeutic strategies (based both on drugs and gene therapies). After all, the general goal is to ameliorate the final output of the specific treatments.

Acknowledgments

The BIO-NMD EU project (N. 241665 to A.F.) is acknowledged.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M.C.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C.; *et al.* The transcriptional landscape of the mammalian genome. *Science* **2005**, *309*, 1559–1563.
2. Mattick, J.S.; Makunin, I.V. Non-coding RNA. *Hum. Mol. Genet.* **2006**, *15*, R17–R29.
3. Kapranov, P.; Cheng, J.; Dike, S.; Nix, D.A.; Duttagupta, R.; Willingham, A.T.; Stadler, P.F.; Hertel, J.; Hackermuller, J.; Hofacker, I.L.; *et al.* RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **2007**, *316*, 1484–1488.
4. Wang, X.Q.; Crutchley, J.L.; Dostie, J. Shaping the genome with non-coding RNAs. *Curr. Genomics* **2011**, *12*, 307–321.
5. Ponting, C.P.; Oliver, P.L.; Reik, W. Evolution and functions of long noncoding RNAs. *Cell* **2009**, *136*, 629–641.
6. Cirak, S.; Arechavala-Gomeza, V.; Guglieri, M.; Feng, L.; Torelli, S.; Anthony, K.; Abbs, S.; Garralda, M.E.; Bourke, J.; Wells, D.J.; *et al.* Exon skipping and dystrophin restoration in patients with duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: An open-label, phase 2, dose-escalation study. *Lancet* **2011**, *378*, 595–605.
7. Goemans, N.M.; Tulinius, M.; van den Akker, J.T.; Burm, B.E.; Ekhart, P.F.; Heuvelmans, N.; Holling, T.; Janson, A.A.; Platenburg, G.J.; Sipkens, J.A.; *et al.* Systemic administration of pro051 in duchenne's muscular dystrophy. *N. Engl. J. Med.* **2011**, *364*, 1513–1522.
8. Evers, M.M.; Pepers, B.A.; van Deutekom, J.C.; Mulders, S.A.; den Dunnen, J.T.; Aartsma-Rus, A.; van Ommen, G.J.; van Roon-Mom, W.M. Targeting several cag expansion diseases by a single antisense oligonucleotide. *PLoS One* **2011**, *6*, e24308.
9. Ferlini, A.; Neri, M.; Gualandi, F. The medical genetics of dystrophinopathies: Molecular genetic diagnosis and its impact on clinical practice. *Neuromuscul. Disord. NMD* **2013**, *23*, 4–14.
10. Mitsuhashi, S.; Kang, P.B. Update on the genetics of limb girdle muscular dystrophy. *Semin. Pediatr. Neurol.* **2012**, *19*, 211–218.

11. Pegoraro, E.; Hoffman, E.P. Limb-girdle Muscular Dystrophy Overview. In *Genereviews*; Pagon, R.A., Adam, M.P., Bird, T.D., Dolan, C.R., Fong, C.T., Stephens, K., Eds.; University of Washington: Seattle, WA, USA, 1993.
12. Davies, K.E.; Nowak, K.J. Molecular mechanisms of muscular dystrophies: Old and new players. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 762–773.
13. Johnson, N.E.; Heatwole, C.R. Myotonic dystrophy: From bench to bedside. *Semin. Neurol.* **2012**, *32*, 246–254.
14. Bovolenta, M.; Erriquez, D.; Valli, E.; Brioschi, S.; Scotton, C.; Neri, M.; Falzarano, M.S.; Gherardi, S.; Fabris, M.; Rimessi, P.; *et al.* The dmd locus harbours multiple long non-coding RNAs which orchestrate and control transcription of muscle dystrophin mRNA isoforms. *PLoS One* **2012**, *7*, e45328.
15. Graves, P.; Zeng, Y. Biogenesis of mammalian micro RNAs: A global view. *Genomics Proteomics Bioinforma.* **2012**, *10*, 239–245.
16. Bentzinger, C.F.; von Maltzahn, J.; Rudnicki, M.A. Extrinsic regulation of satellite cell specification. *Stem Cell Res. Ther.* **2010**, *1*, 27.
17. Bentzinger, C.F.; Wang, Y.X.; Rudnicki, M.A. Building muscle: Molecular regulation of myogenesis. *Cold Spring Harb Perspect Biol.* **2012**, *4*, doi:10.1101/cshperspect.a008342.
18. Bismuth, K.; Relaix, F. Genetic regulation of skeletal muscle development. *Exp. Cell Res.* **2010**, *316*, 3081–3086.
19. Bryson-Richardson, R.J.; Currie, P.D. The genetics of vertebrate myogenesis. *Nat. Rev. Genet.* **2008**, *9*, 632–646.
20. Punch, V.G.; Jones, A.E.; Rudnicki, M.A. Transcriptional networks that regulate muscle stem cell function. *Wiley Interdiscip Rev. Syst. Biol. Med.* **2009**, *1*, 128–140.
21. Kassar-Duchossoy, L.; Gayraud-Morel, B.; Gomes, D.; Rocancourt, D.; Buckingham, M.; Shinin, V.; Tajbakhsh, S. Mrf4 determines skeletal muscle identity in myf5:Myod double-mutant mice. *Nature* **2004**, *431*, 466–471.
22. Nabeshima, Y.; Hanaoka, K.; Hayasaka, M.; Esumi, E.; Li, S.; Nonaka, I. Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature* **1993**, *364*, 532–535.
23. Charge, S.B.; Rudnicki, M.A. Cellular and molecular regulation of muscle regeneration. *Physiol. Rev.* **2004**, *84*, 209–238.
24. Parker, M.H.; Seale, P.; Rudnicki, M.A. Looking back to the embryo: Defining transcriptional networks in adult myogenesis. *Nat. Rev. Genet.* **2003**, *4*, 497–507.
25. Pownall, M.E.; Gustafsson, M.K.; Emerson, C.P., Jr. Myogenic regulatory factors and the specification of muscle progenitors in vertebrate embryos. *Annu. Rev. Cell Dev. Biol.* **2002**, *18*, 747–783.
26. Buckingham, M. Skeletal muscle formation in vertebrates. *Curr. Opin. Genet. Dev.* **2001**, *11*, 440–448.
27. Hasty, P.; Bradley, A.; Morris, J.H.; Edmondson, D.G.; Venuti, J.M.; Olson, E.N.; Klein, W.H. Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. *Nature* **1993**, *364*, 501–506.

28. Yokoyama, S.; Asahara, H. The myogenic transcriptional network. *Cell Mol. Life Sci.* **2011**, *68*, 1843–1849.
29. O'Rourke, J.R.; Georges, S.A.; Seay, H.R.; Tapscott, S.J.; McManus, M.T.; Goldhamer, D.J.; Swanson, M.S.; Harfe, B.D. Essential role for dicer during skeletal muscle development. *Dev. Biol.* **2007**, *311*, 359–368.
30. Zhao, Y.; Samal, E.; Srivastava, D. Serum response factor regulates a muscle-specific microRNA that targets hand2 during cardiogenesis. *Nature* **2005**, *436*, 214–220.
31. Rosenberg, M.I.; Georges, S.A.; Asawachaicharn, A.; Analau, E.; Tapscott, S.J. Myod inhibits fstl1 and utrn expression by inducing transcription of mir-206. *J. Cell Biol.* **2006**, *175*, 77–85.
32. Rao, P.K.; Kumar, R.M.; Farkhondeh, M.; Baskerville, S.; Lodish, H.F. Myogenic factors that regulate expression of muscle-specific microRNAs. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8721–8726.
33. Liu, N.; Williams, A.H.; Kim, Y.; McAnally, J.; Bezprozvannaya, S.; Sutherland, L.B.; Richardson, J.A.; Bassel-Duby, R.; Olson, E.N. An intragenic mef2-dependent enhancer directs muscle-specific expression of microRNAs 1 and 133. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 20844–20849.
34. Sun, Y.; Ge, Y.; Drnevich, J.; Zhao, Y.; Band, M.; Chen, J. Mammalian target of rapamycin regulates miRNA-1 and follistatin in skeletal myogenesis. *J. Cell Biol.* **2010**, *189*, 1157–1169.
35. Weintraub, H. The myod family and myogenesis: Redundancy, networks, and thresholds. *Cell* **1993**, *75*, 1241–1244.
36. Naya, F.J.; Olson, E. Mef2: A transcriptional target for signaling pathways controlling skeletal muscle growth and differentiation. *Curr. Opin. Cell Biol.* **1999**, *11*, 683–688.
37. Van Rooij, E.; Liu, N.; Olson, E.N. MicroRNAs flex their muscles. *Trends Genet. TIG* **2008**, *24*, 159–166.
38. Chen, J.F.; Callis, T.E.; Wang, D.Z. MicroRNAs and muscle disorders. *J. Cell Sci.* **2009**, *122*, 13–20.
39. Eisenberg, I.; Alexander, M.S.; Kunkel, L.M. miRNAs in normal and diseased skeletal muscle. *J. Cell. Mol. Med.* **2009**, *13*, 2–11.
40. Ge, Y.; Chen, J. MicroRNAs in skeletal myogenesis. *Cell Cycle* **2011**, *10*, 441–448.
41. Cacchiarelli, D.; Legnini, I.; Martone, J.; Cazzella, V.; D'Amico, A.; Bertini, E.; Bozzoni, I. miRNAs as serum biomarkers for duchenne muscular dystrophy. *EMBO Mol. Med.* **2011**, *3*, 258–265.
42. Mitchell, P.S.; Parkin, R.K.; Kroh, E.M.; Fritz, B.R.; Wyman, S.K.; Pogosova-Agadjanyan, E.L.; Peterson, A.; Noteboom, J.; O'Briant, K.C.; Allen, A.; *et al.* Circulating microRNAs as stable blood-based markers for cancer detection. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10513–10518.
43. Van Rooij, E.; Quiat, D.; Johnson, B.A.; Sutherland, L.B.; Qi, X.; Richardson, J.A.; Kelm, R.J., Jr.; Olson, E.N. A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. *Dev. Cell* **2009**, *17*, 662–673.
44. Sun, Q.; Zhang, Y.; Yang, G.; Chen, X.; Cao, G.; Wang, J.; Sun, Y.; Zhang, P.; Fan, M.; Shao, N.; *et al.* Transforming growth factor-beta-regulated mir-24 promotes skeletal muscle differentiation. *Nucleic Acids Res.* **2008**, *36*, 2690–2699.

45. Caretti, G.; Di Padova, M.; Micales, B.; Lyons, G.E.; Sartorelli, V. The polycomb ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. *Genes Dev.* **2004**, *18*, 2627–2638.
46. Dey, B.K.; Gagan, J.; Yan, Z.; Dutta, A. Mir-26a is required for skeletal muscle differentiation and regeneration in mice. *Genes Dev.* **2012**, *26*, 2180–2191.
47. Crist, C.G.; Montarras, D.; Pallafacchina, G.; Rocancourt, D.; Cumano, A.; Conway, S.J.; Buckingham, M. Muscle stem cell behavior is modified by microRNA-27 regulation of pax3 expression. *Proc. Natl. Acad. Sci. USA.* **2009**, *106*, 13383–13387.
48. Wang, H.; Garzon, R.; Sun, H.; Ladner, K.J.; Singh, R.; Dahlman, J.; Cheng, A.; Hall, B.M.; Qualman, S.J.; Chandler, D.S.; *et al.* Nf-kappab-yy1-mir-29 regulatory circuitry in skeletal myogenesis and rhabdomyosarcoma. *Cancer Cell* **2008**, *14*, 369–381.
49. Li, Z.; Hassan, M.Q.; Jafferji, M.; Aqeilan, R.I.; Garzon, R.; Croce, C.M.; van Wijnen, A.J.; Stein, J.L.; Stein, G.S.; Lian, J.B. Biological functions of mir-29b contribute to positive regulation of osteoblast differentiation. *J. Biol. Chem.* **2009**, *284*, 15676–15684.
50. Ge, Y.; Sun, Y.; Chen, J. Igf-ii is regulated by microRNA-125b in skeletal myogenesis. *J. Cell Biol.* **2011**, *192*, 69–81.
51. Erbay, E.; Park, I.H.; Nuzzi, P.D.; Schoenherr, C.J.; Chen, J. Igf-ii transcription in skeletal myogenesis is controlled by mtor and nutrients. *J. Cell Biol.* **2003**, *163*, 931–936.
52. Ge, Y.; Wu, A.L.; Warnes, C.; Liu, J.; Zhang, C.; Kawasome, H.; Terada, N.; Boppart, M.D.; Schoenherr, C.J.; Chen, J. Mtor regulates skeletal muscle regeneration *in vivo* through kinase-dependent and kinase-independent mechanisms. *Am. J. Physiol. Cell Physiol.* **2009**, *297*, C1434–C1444.
53. Seok, H.Y.; Tatsuguchi, M.; Callis, T.E.; He, A.; Pu, W.T.; Wang, D.Z. Mir-155 inhibits expression of the mef2a protein to repress skeletal muscle differentiation. *J. Biol. Chem.* **2011**, *286*, 35339–35346.
54. Naguibneva, I.; Ameyar-Zazoua, M.; Polesskaya, A.; Ait-Si-Ali, S.; Groisman, R.; Souidi, M.; Cuvellier, S.; Harel-Bellan, A. The microRNA mir-181 targets the homeobox protein hox-a11 during mammalian myoblast differentiation. *Nat. Cell Biol.* **2006**, *8*, 278–284.
55. Kuang, W.; Tan, J.; Duan, Y.; Duan, J.; Wang, W.; Jin, F.; Jin, Z.; Yuan, X.; Liu, Y. Cyclic stretch induced mir-146a upregulation delays c2c12 myogenic differentiation through inhibition of numb. *Biochem. Biophys. Res. Commun.* **2009**, *378*, 259–263.
56. Conboy, I.M.; Rando, T.A. The regulation of notch signaling controls satellite cell activation and cell fate determination in postnatal myogenesis. *Dev. Cell* **2002**, *3*, 397–409.
57. Flynt, A.S.; Li, N.; Thatcher, E.J.; Solnica-Krezel, L.; Patton, J.G. Zebrafish mir-214 modulates hedgehog signaling to specify muscle cell fate. *Nat. Genet.* **2007**, *39*, 259–263.
58. Cardinali, B.; Castellani, L.; Fasanaro, P.; Basso, A.; Alema, S.; Martelli, F.; Falcone, G. MicroRNA-221 and microRNA-222 modulate differentiation and maturation of skeletal muscle cells. *PLoS One* **2009**, *4*, e7607.
59. Sarkar, S.; Dey, B.K.; Dutta, A. Mir-322/424 and -503 are induced during muscle differentiation and promote cell cycle quiescence and differentiation by down-regulation of cdc25a. *Mol. Biol. Cell* **2010**, *21*, 2138–2149.

60. Dey, B.K.; Gagan, J.; Dutta, A. Mir-206 and -486 induce myoblast differentiation by downregulating pax7. *Mol. Cell Biol.* **2011**, *31*, 203–214.
61. Small, E.M.; O'Rourke, J.R.; Moresi, V.; Sutherland, L.B.; McAnally, J.; Gerard, R.D.; Richardson, J.A.; Olson, E.N. Regulation of pi3-kinase/akt signaling by muscle-enriched microRNA-486. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4218–4223.
62. Juan, A.H.; Kumar, R.M.; Marx, J.G.; Young, R.A.; Sartorelli, V. Mir-214-dependent regulation of the polycomb protein ezh2 in skeletal muscle and embryonic stem cells. *Mol. Cell* **2009**, *36*, 61–74.
63. Liu, J.; Luo, X.J.; Xiong, A.W.; Zhang, Z.D.; Yue, S.; Zhu, M.S.; Cheng, S.Y. MicroRNA-214 promotes myogenic differentiation by facilitating exit from mitosis via down-regulation of proto-oncogene n-ras. *J. Biol. Chem.* **2010**, *285*, 26599–26607.
64. Marrone, A.K.; Shcherbata, H.R. Dystrophin orchestrates the epigenetic profile of muscle cells via miRNAs. *Front. Genet.* **2011**, *2*, doi:10.3389/fgene.2011.0006.
65. Greco, S.; de Simone, M.; Colussi, C.; Zaccagnini, G.; Fasanaro, P.; Pescatori, M.; Cardani, R.; Perbellini, R.; Isaia, E.; Sale, P.; *et al.* Common micro-RNA signature in skeletal muscle damage and regeneration induced by duchenne muscular dystrophy and acute ischemia. *FASEB J.* **2009**, *23*, 3335–3346.
66. Cacchiarelli, D.; Martone, J.; Girardi, E.; Cesana, M.; Incitti, T.; Morlando, M.; Nicoletti, C.; Santini, T.; Sthandier, O.; Barberi, L.; *et al.* MicroRNAs involved in molecular circuitries relevant for the duchenne muscular dystrophy pathogenesis are controlled by the dystrophin/nnos pathway. *Cell Metab.* **2010**, *12*, 341–351.
67. Perbellini, R.; Greco, S.; Sarra-Ferraris, G.; Cardani, R.; Capogrossi, M.C.; Meola, G.; Martelli, F. Dysregulation and cellular mislocalization of specific miRNAs in myotonic dystrophy type 1. *Neuromuscul. Disord. NMD* **2011**, *21*, 81–88.
68. Rau, F.; Freyermuth, F.; Fugier, C.; Villemin, J.P.; Fischer, M.C.; Jost, B.; Dembele, D.; Gourdon, G.; Nicole, A.; Duboc, D.; *et al.* Misregulation of mir-1 processing is associated with heart defects in myotonic dystrophy. *Nat. Struct. Mol. Biol.* **2011**, *18*, 840–845.
69. Gambardella, S.; Rinaldi, F.; Lepore, S.M.; Viola, A.; Loro, E.; Angelini, C.; Vergani, L.; Novelli, G.; Botta, A. Overexpression of microRNA-206 in the skeletal muscle from myotonic dystrophy type 1 patients. *J. Transl. Med.* **2010**, *8*, doi:10.1186/1479-5876-8-48.
70. Cacchiarelli, D.; Incitti, T.; Martone, J.; Cesana, M.; Cazzella, V.; Santini, T.; Sthandier, O.; Bozzoni, I. Mir-31 modulates dystrophin expression: New implications for duchenne muscular dystrophy therapy. *EMBO Rep.* **2011**, *12*, 136–141.
71. Eisenberg, I.; Eran, A.; Nishino, I.; Moggio, M.; Lamperti, C.; Amato, A.A.; Lidov, H.G.; Kang, P.B.; North, K.N.; Mitrani-Rosenbaum, S.; *et al.* Distinctive patterns of microRNA expression in primary muscular disorders. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 17016–17021.
72. De Arcangelis, V.; Serra, F.; Cogoni, C.; Vivarelli, E.; Monaco, L.; Naro, F. Beta1-syntrophin modulation by mir-222 in mdx mice. *PLoS One* **2010**, *5*, e12098.
73. Greco, S.; Perfetti, A.; Fasanaro, P.; Cardani, R.; Capogrossi, M.C.; Meola, G.; Martelli, F. Deregulated microRNAs in myotonic dystrophy type 2. *PLoS One* **2012**, *7*, e39732.
74. Williams, A.H.; Valdez, G.; Moresi, V.; Qi, X.; McAnally, J.; Elliott, J.L.; Bassel-Duby, R.; Sanes, J.R.; Olson, E.N. MicroRNA-206 delays als progression and promotes regeneration of neuromuscular synapses in mice. *Science* **2009**, *326*, 1549–1554.

75. Durbeej, M.; Campbell, K.P. Muscular dystrophies involving the dystrophin-glycoprotein complex: An overview of current mouse models. *Curr. Opin. Genet. Dev.* **2002**, *12*, 349–361.
76. Ervasti, J.M.; Sonnemann, K.J. Biology of the striated muscle dystrophin-glycoprotein complex. *Int. Rev. Cytol.* **2008**, *265*, 191–225.
77. Matsumura, K.; Tome, F.M.; Collin, H.; Leturcq, F.; Jeanpierre, M.; Kaplan, J.C.; Fardeau, M.; Campbell, K.P. Expression of dystrophin-associated proteins in dystrophin-positive muscle fibers (revertants) in duchenne muscular dystrophy. *Neuromuscul. Disord. NMD* **1994**, *4*, 115–120.
78. McKinsey, T.A.; Zhang, C.L.; Lu, J.; Olson, E.N. Signal-dependent nuclear export of a histone deacetylase regulates muscle differentiation. *Nature* **2000**, *408*, 106–111.
79. Puri, P.L.; Iezzi, S.; Stiegler, P.; Chen, T.T.; Schiltz, R.L.; Muscat, G.E.; Giordano, A.; Kedes, L.; Wang, J.Y.; Sartorelli, V. Class i histone deacetylases sequentially interact with myod and prb during skeletal myogenesis. *Mol. Cell* **2001**, *8*, 885–897.
80. Brenman, J.E.; Chao, D.S.; Xia, H.; Aldape, K.; Bredt, D.S. Nitric oxide synthase complexed with dystrophin and absent from skeletal muscle sarcolemma in duchenne muscular dystrophy. *Cell* **1995**, *82*, 743–752.
81. Colussi, C.; Mozzetta, C.; Gurtner, A.; Illi, B.; Rosati, J.; Straino, S.; Ragone, G.; Pescatori, M.; Zaccagnini, G.; Antonini, A.;*et al.* Hdac2 blockade by nitric oxide and histone deacetylase inhibitors reveals a common target in duchenne muscular dystrophy treatment. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 19183–19187.
82. Brook, J.D.; McCurrach, M.E.; Harley, H.G.; Buckler, A.J.; Church, D.; Aburatani, H.; Hunter, K.; Stanton, V.P.; Thirion, J.P.; Hudson, T.; *et al.* Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (ctg) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **1992**, *68*, 799–808.
83. Fu, Y.H.; Pizzuti, A.; Fenwick, R.G., Jr.; King, J.; Rajnarayan, S.; Dunne, P.W.; Dubel, J.; Nasser, G.A.; Ashizawa, T.; de Jong, P.; *et al.* An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* **1992**, *255*, 1256–1258.
84. Mahadevan, M.; Tsilfidis, C.; Sabourin, L.; Shutler, G.; Amemiya, C.; Jansen, G.; Neville, C.; Narang, M.; Barcelo, J.; O'Hoy, K.; *et al.* Myotonic dystrophy mutation: An unstable ctg repeat in the 3' untranslated region of the gene. *Science* **1992**, *255*, 1253–1255.
85. Liquori, C.L.; Ricker, K.; Moseley, M.L.; Jacobsen, J.F.; Kress, W.; Naylor, S.L.; Day, J.W.; Ranum, L.P. Myotonic dystrophy type 2 caused by a cctg expansion in intron 1 of znf9. *Science* **2001**, *293*, 864–867.
86. Wojciechowska, M.; Krzyzosiak, W.J. Cellular toxicity of expanded RNA repeats: Focus on RNA foci. *Hum. Mol. Genet.* **2011**, *20*, 3811–3821.
87. Sicot, G.; Gomes-Pereira, M. RNA toxicity in human disease and animal models: From the uncovering of a new mechanism to the development of promising therapies. *Biochim. Biophys. Acta* **2013**, *1832*, 1390–1409.
88. Udd, B.; Krahe, R. The myotonic dystrophies: Molecular, clinical, and therapeutic challenges. *Lancet Neurol.* **2012**, *11*, 891–905.
89. Li, X.; Wu, Z.; Fu, X.; Han, W. Long noncoding RNAs: Insights from biological features and functions to diseases. *Med. Res. Rev.* **2013**, *33*, 517–553.

90. Cesana, M.; Cacchiarelli, D.; Legnini, I.; Santini, T.; Sthandier, O.; Chinappi, M.; Tramontano, A.; Bozzoni, I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **2011**, *147*, 358–369.
91. Watts, R.; Johnsen, V.L.; Shearer, J.; Hittel, D.S. Myostatin-induced inhibition of the long noncoding RNA malat1 is associated with decreased myogenesis. *Am. J. Physiol. Cell Physiol.* **2013**, *304*, C995–C1001.
92. Sunwoo, H.; Dinger, M.E.; Wilusz, J.E.; Amaral, P.P.; Mattick, J.S.; Spector, D.L. Men epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res.* **2009**, *19*, 347–359.
93. Sasaki, Y.T.; Ideue, T.; Sano, M.; Mituyama, T.; Hirose, T. Menepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 2525–2530.
94. Clemson, C.M.; Hutchinson, J.N.; Sara, S.A.; Ensminger, A.W.; Fox, A.H.; Chess, A.; Lawrence, J.B. An architectural role for a nuclear noncoding RNA: Neat1 RNA is essential for the structure of paraspeckles. *Mol. Cell* **2009**, *33*, 717–726.
95. Hube, F.; Velasco, G.; Rollin, J.; Furling, D.; Francastel, C. Steroid receptor RNA activator protein binds to and counteracts sra RNA-mediated activation of myod and muscle differentiation. *Nucleic Acids Res.* **2011**, *39*, 513–525.
96. Caretti, G.; Schiltz, R.L.; Dilworth, F.J.; Di Padova, M.; Zhao, P.; Ogryzko, V.; Fuller-Pace, F.V.; Hoffman, E.P.; Tapscott, S.J.; Sartorelli, V. The RNA helicases p68/p72 and the noncoding RNA sra are coregulators of myod and skeletal muscle differentiation. *Dev. Cell* **2006**, *11*, 547–560.
97. Caretti, G.; Lei, E.P.; Sartorelli, V. The dead-box p68/p72 proteins and the noncoding RNA steroid receptor activator sra: Eclectic regulators of disparate biological functions. *Cell Cycle* **2007**, *6*, 1172–1176.
98. Willingham, A.T.; Orth, A.P.; Batalov, S.; Peters, E.C.; Wen, B.G.; Aza-Blanc, P.; Hogenesch, J.B.; Schultz, P.G. A strategy for probing the function of noncoding RNAs finds a repressor of nfat. *Science* **2005**, *309*, 1570–1573.
99. Sharma, S.; Findlay, G.M.; Bandukwala, H.S.; Oberdoerffer, S.; Baust, B.; Li, Z.; Schmidt, V.; Hogan, P.G.; Sacks, D.B.; Rao, A. Dephosphorylation of the nuclear factor of activated t cells (nfat) transcription factor is regulated by an RNA-protein scaffold complex. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11381–11386.
100. Niland, C.N.; Merry, C.R.; Khalil, A.M. Emerging roles for long non-coding RNAs in cancer and neurological disorders. *Front. Genet.* **2012**, *3*, doi:10.3389/fgene.2012.00025.
101. Cabianca, D.S.; Casa, V.; Bodega, B.; Xynos, A.; Ginelli, E.; Tanaka, Y.; Gabellini, D. A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in fshd muscular dystrophy. *Cell* **2012**, *149*, 819–831.
102. Schmidt, L.H.; Spieker, T.; Koschmieder, S.; Schaffers, S.; Humberg, J.; Jungen, D.; Bulk, E.; Hascher, A.; Wittmer, D.; Marra, A.; *et al.* The long noncoding malat-1 RNA indicates a poor prognosis in non-small cell lung cancer and induces migration and tumor growth. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* **2011**, *6*, 1984–1992.
103. Tripathi, V.; Ellis, J.D.; Shen, Z.; Song, D.Y.; Pan, Q.; Watt, A.T.; Freier, S.M.; Bennett, C.F.; Sharma, A.; Bubulya, P.A.; *et al.* The nuclear-retained noncoding RNA malat1 regulates

alternative splicing by modulating sr splicing factor phosphorylation. *Mol. Cell* **2010**, *39*, 925–938.

104. Wilusz, J.E.; JnBaptiste, C.K.; Lu, L.Y.; Kuhn, C.D.; Joshua-Tor, L.; Sharp, P.A. A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(a) tails. *Genes Dev.* **2012**, *26*, 2392–2407.
105. Gutschner, T.; Hammerle, M.; Diederichs, S. Malat1—A paradigm for long noncoding RNA function in cancer. *J. Mol. Med.* **2013**, *91*, 791–801.
106. Langley, B.; Thomas, M.; Bishop, A.; Sharma, M.; Gilmour, S.; Kambadur, R. Myostatin inhibits myoblast differentiation by down-regulating myod expression. *J. Biol. Chem.* **2002**, *277*, 49831–49840.
107. Rios, R.; Carneiro, I.; Arce, V.M.; Devesa, J. Myostatin is an inhibitor of myogenic differentiation. *Am. J. Physiol. Cell Physiol.* **2002**, *282*, C993–C999.
108. Zhang, B.; Arun, G.; Mao, Y.S.; Lazar, Z.; Hung, G.; Bhattacharjee, G.; Xiao, X.; Booth, C.J.; Wu, J.; Zhang, C.; *et al.* The lncRNA malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Rep.* **2012**, *2*, 111–123.
109. Eissmann, M.; Gutschner, T.; Hammerle, M.; Gunther, S.; Caudron-Herger, M.; Gross, M.; Schirmacher, P.; Rippe, K.; Braun, T.; Zornig, M.; *et al.* Loss of the abundant nuclear non-coding RNA malat1 is compatible with life and development. *RNA Biol.* **2012**, *9*, 1076–1087.
110. Prasanth, K.V.; Prasanth, S.G.; Xuan, Z.; Hearn, S.; Freier, S.M.; Bennett, C.F.; Zhang, M.Q.; Spector, D.L. Regulating gene expression through RNA nuclear retention. *Cell* **2005**, *123*, 249–263.
111. Ahn, A.H.; Kunkel, L.M. The structural and functional diversity of dystrophin. *Nat. Genet.* **1993**, *3*, 283–291.
112. Muntoni, F.; Torelli, S.; Ferlini, A. Dystrophin and mutations: One gene, several proteins, multiple phenotypes. *Lancet Neurol.* **2003**, *2*, 731–740.
113. Torelli, S.; Ferlini, A.; Obici, L.; Sewry, C.; Muntoni, F. Expression, regulation and localisation of dystrophin isoforms in human foetal skeletal and cardiac muscle. *Neuromuscul. Disord. NMD* **1999**, *9*, 541–551.
114. Brioschi, S.; Gualandi, F.; Scotton, C.; Armaroli, A.; Bovolenta, M.; Falzarano, M.S.; Sabatelli, P.; Selvatici, R.; D'Amico, A.; Pane, M.; *et al.* Genetic characterization in symptomatic female dmd carriers: Lack of relationship between x-inactivation, transcriptional dmd allele balancing and phenotype. *BMC Med. Genet.* **2012**, *13*, doi:10.1186/1471-2350-13-73.
115. Tran, T.H.; Zhang, Z.; Yagi, M.; Lee, T.; Awano, H.; Nishida, A.; Okinaga, T.; Takeshima, Y.; Matsuo, M. Molecular characterization of an x(p21.2;q28) chromosomal inversion in a duchenne muscular dystrophy patient with mental retardation reveals a novel long non-coding gene on xq28. *J. Hum. Genet.* **2013**, *58*, 33–39.
116. Lee, J.T. Lessons from x-chromosome inactivation: Long ncRNA as guides and tethers to the epigenome. *Genes Dev.* **2009**, *23*, 1831–1842.
117. Rinn, J.L.; Kertesz, M.; Wang, J.K.; Squazzo, S.L.; Xu, X.; Brugmann, S.A.; Goodnough, L.H.; Helms, J.A.; Farnham, P.J.; Segal, E.; *et al.* Functional demarcation of active and silent chromatin domains in human hox loci by noncoding RNAs. *Cell* **2007**, *129*, 1311–1323.
118. Rukov, J.L.; Shomron, N. MicroRNA pharmacogenomics: Post-transcriptional regulation of drug response. *Trends Mol. Med.* **2011**, *17*, 412–423.

119. Mishra, P.J. The miRNA-drug resistance connection: A new era of personalized medicine using noncoding RNA begins. *Pharmacogenomics* **2012**, *13*, 1321–1324.
120. Giovannetti, E.; van der Velde, A.; Funel, N.; Vasile, E.; Perrone, V.; Leon, L.G.; de Lio, N.; Avan, A.; Caponi, S.; Pollina, L.E.; *et al.* High-throughput microRNA (mirRNA) arrays unravel the prognostic role of mir-211 in pancreatic cancer. *PLoS One* **2012**, *7*, e49145.
121. Gagan, J.; Dey, B.K.; Dutta, A. MicroRNAs regulate and provide robustness to the myogenic transcriptional network. *Curr. Opin. Pharmacol.* **2012**, *12*, 383–388.
122. Twayana, S.; Legnini, I.; Cesana, M.; Cacchiarelli, D.; Morlando, M.; Bozzoni, I. Biogenesis and function of non-coding RNAs in muscle differentiation and in duchenne muscular dystrophy. *Biochem. Soc. Trans.* **2013**, *41*, 844–849.

Reprinted from *IJMS*. Cite as: Deng, G.; Sui, G. Noncoding RNA in Oncogenesis: A New Era of Identifying Key Players. *Int. J. Mol. Sci.* **2013**, *14*, 18319-18349.

Review

Noncoding RNA in Oncogenesis: A New Era of Identifying Key Players

Guorui Deng [1,2] and Guangchao Sui [1,*]

[1] Department of Cancer Biology and Comprehensive Cancer Center, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA

[2] Hypertension and Vascular Research Center, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA; E-Mail: gdeng@wakehealth.edu

* Author to whom correspondence should be addressed; E-Mail: gsui@wakehealth.edu; Tel.: +1-336-713-0052; Fax: +1-336-716-0255.

Received: 8 July 2013; in revised form: 23 August 2013 / Accepted: 30 August 2013 / Published: 5 September 2013

Abstract: New discoveries and accelerating progresses in the field of noncoding RNAs (ncRNAs) continuously challenges our deep-rooted doctrines in biology and sometimes our imagination. A growing body of evidence indicates that ncRNAs are important players in oncogenesis. While a stunning list of ncRNAs has been discovered, only a small portion of them has been examined for their biological activities and very few have been characterized for the molecular mechanisms of their action. To date, ncRNAs have been shown to regulate a wide range of biological processes, including chromatin remodeling, gene transcription, mRNA translation and protein function. Dysregulation of ncRNAs contributes to the pathogenesis of a variety of cancers and aberrant ncRNA expression has a high potential to be prognostic in some cancers. Thus, a new cancer research era has begun to identify novel key players of ncRNAs in oncogenesis. In this review, we will first discuss the function and regulation of miRNAs, especially focusing on the interplay between miRNAs and several key cancer genes, including p53, PTEN and c-Myc. We will then summarize the research of long ncRNAs (lncRNAs) in cancers. In this part, we will discuss the lncRNAs in four categories based on their activities, including regulating gene expression, acting as miRNA decoys, mediating mRNA translation, and modulating protein activities. At the end, we will also discuss recently unraveled activities of circular RNAs (circRNAs).

Keywords: noncoding RNA; microRNA; gene expression; cancer; tumor suppressor; oncogene

1. Introduction

The central dogma of biology dictates that genetic information flows in a unidirectional fashion of DNA-mRNA-proteins. Thus, the majority of the early research in molecular and cellular biology was focused on the protein-coding genes and their transcripts, messenger RNAs (mRNAs). However, in the past two decades, noncoding RNAs (ncRNAs) have garnered increased appreciation for their important roles in regulating various biological processes. In the whole human genome, more than 90% of the DNA sequence can be transcribed, but only about 2% of it encodes proteins [1–5]. Recent studies reveal that most of these excess transcripts are not transcriptional noise, but rather serve as functional ncRNAs regulating chromatin modifications, gene transcription, mRNA translation and protein function [5–7].

The discovery of the first human oncogene RAS [8–10] and the first tumor suppressor Retinoblastoma (Rb) in the 1980s invigorated the field of cancer research [11,12]. The fact that a simple point mutation in the RAS gene could transform a normal growth-regulating gene into a cancer-causing factor inspired cancer researchers to seek out additional cancer-regulating genes. In the decades that followed, the search for novel oncogenes and tumor suppressors has dominated the field of cancer research, leading to the discovery of many well-characterized tumor suppressors, such as p53, and oncogenes, such as c-Myc. With the discovery of ncRNAs and their involvement in oncogenesis, a new surge of research in determining the oncogenic and tumor suppressive roles of these noncoding transcripts has taken place. To date, multiple classes of ncRNAs have been demonstrated to regulate different biological or physiological processes. It is possible that some ncRNAs may not have a discernible function; however, the enormous number of the identified noncoding transcripts and their differential expression profiles in cancers indicate that we are still at the beginning of this exciting era to explore the functions and regulatory mechanisms of ncRNAs in cancers.

NcRNAs can be divided into three categories based on length or number of nucleotides (nts) [13]. (1) Short ncRNAs: these RNAs are 17–30 nts in length and include microRNAs (miRNAs), piwi-interacting RNAs (piRNAs), and transcription initiation RNAs (tiRNAs). MiRNAs will be one of the major focuses of this review; (2) Middle-size ncRNAs: these RNAs have variable sizes but are typically between 20 and 200 nts in length. The small nucleolar RNAs (snoRNAs) belong to this category; (3) Long ncRNAs (lncRNAs): these RNAs are over 200 nts. This category includes several well-characterized ncRNA, such as MALAT1 and HOTAIR. Circular RNAs (circRNAs) have recently been discovered and can be grouped here based on their sizes [14]. Many members in this category were identified in the last few years, which will be extensively discussed in this review.

Due to rapid expanding of the ncRNA field, it is difficult to cover all aspects of ncRNA-related studies in a single review article. There are many excellent reviews discussing the activities of ncRNAs from different perspectives. In this review, we will discuss the biological function of miRNAs and lncRNAs in cancers, with particular attention paid to their interplay with p53, c-Myc and PTEN.

We will also summarize recent studies of circRNAs that have returned to the arena with a recently unraveled function since their discovery many years ago.

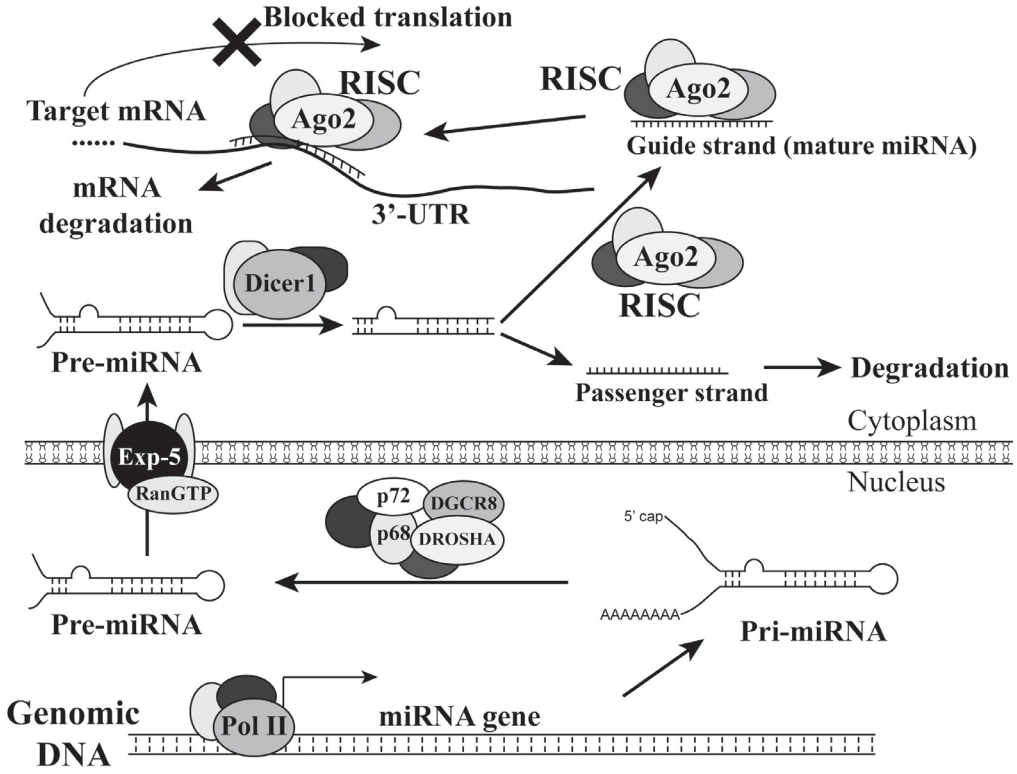
2. MicroRNAs

2.1. Introduction of MicroRNAs

MiRNAs are small and evolutionarily conserved noncoding RNAs that are typically 18 to 25 nts in length. The biogenesis of miRNAs begins with the synthesis of the primary transcripts (pri-miRNAs) by RNA polymerase II in the nucleus (Figure 1). Like the transcripts of protein-coding genes, pri-miRNAs contain a 5' cap structure, a poly(A) tail, and sometimes intron sequences [15,16]. A stem loop structure can be formed by regions of partial complementary sequences in each pri-miRNA. The nuclear ribonuclease DROSHA and its partner DGCR8 recognize this stem-loop structure, and subsequently crop the pri-miRNA to generate the pre-miRNA intermediate. After the pre-miRNA is exported to the cytoplasm by Exportin-5/Ran-GTP, it is further processed by DICER1, another ribonuclease, to generate a double-stranded miRNA molecule. Both strands can act as mature miRNAs; however, if only one strand (guide strand) becomes a functional miRNA, the other strand (passenger strand) will be quickly degraded [17,18].

A mature miRNA can associate with Argonaute proteins to form a RNA-induced silencing complex (RISC), in which the miRNA guides the complex to the 3'-untranslated region (3'-UTR) of the target mRNA to block translational protein synthesis and/or cause its degradation (Figure 1). Each miRNA contains a seed sequence (7 nts; from the nucleotides 2 to 8) at its 5'-end with conserved complementarity to perfectly pair the ‘seed-match’ sequence at the 3'-UTR of its target mRNA [19,20]. Since only a 7-nucleotide match between a target 3'-UTR and a miRNA seed sequence will theoretically make the mRNA a potential target of the miRNA, miRNAs do not have a high specificity for their targets like small interference RNAs (siRNAs) have. MiRNAs can be divided into different families. A miRNA family consists miRNAs that share the same seed sequence and thus may target the same set of genes [19]. MiRNA coding regions, or their genes, can be located in either protein-coding or noncoding regions of transcription units in the human genome. Each of these regions may encode one or a cluster of mRNAs. A miRNA cluster is a set of two or more miRNAs that are transcribed from physically adjacent miRNA genes. Thus, the miRNAs in a cluster are transcribed by the same promoter, in the same direction, and as an unseparated transcription unit. One cluster typically contains two or three miRNAs, but large clusters do exist, such as the miRNA-17-92 cluster consisting of seven miRNAs. Interestingly, some members in one miRNA family are encoded by one miRNA gene cluster, such as miR-15a/miR-16-1 and miR-34b/miR-34c. Different miRNA clusters may belong to the same miRNA family, such as miR-15a/miR-16-1 and miR-497/miR-195.

Figure 1. MicroRNA biogenesis and their role in inhibiting gene expression. A miRNA is transcribed by RNA polymerase II (Pol II) mediated by other transcription factors, such as p53 and c-Myc. The generated pri-miRNA is processed by the DROSHA complex to become a pre-miRNA, which is transported from nucleus to cytoplasm with a complex consisting of Exportin-5 (Exp-5) and Ran-GTP. In cytoplasm, the pre-miRNA is further processed by a Dicer-1 complex to become a duplex that consists of a guide strand (to become a mature miRNA) and a passenger strand. The guide strand associates with the RNA-induced silencing complex (RISC) and guides it to the target site on the 3'-UTR of an mRNA to inhibit the mRNA translation and cause mRNA degradation. For some miRNA-duplexes post Dicer-1 processing, both strands can become mature miRNAs; the one at the 5'-end of pre-miRNA is suffixed by "-5p", while the 3'-end one is suffixed by "-3", such as miR-17-5p and miR-17-3p.



To date, over 2,500 human microRNAs have been identified based on the microRNA database (miRBase) at the Sanger Institute (http://www.mirbase.org/). Bioinformatical analyses suggest that miRNAs may regulate over 5,300 human genes, which represents 30% of human genes [19] and each miRNA regulates about 200 genes [21]. Thus, altered expression of microRNAs in cancer can cause significant perturbation of gene expression with a profound effect on malignant transformation and cancer progression [20].

There are much debate and investigation regarding the mechanisms of miRNA-repressed protein expression. Initially, miRNA-mediated gene silencing was portrayed as a different mechanism from that of siRNA-mediated mRNA degradation, based on the observation that the lin-4 miRNA decreased lin-14 protein expression without affecting its mRNA levels in *C. elegans* [22,23]. This was supported by multiple studies showing that miRNA-mediated gene silencing could be achieved with no or minor change of the target mRNA levels in *Drosophila* [24], *Arabidopsis* [25] and mammalian cells [26,27].

However, several later studies suggested that microRNAs can both block translation and trigger target mRNA degradation [28,29]. Notably, a recent report by Djuranovic *et al.* provided new insights into the kinetics of miRNA-mediated gene silencing in *Drosophila* cells. The authors demonstrated that miR-9b and miR-279 first repress the translation of the target mRNAs and then cause mRNA deadenylation and degradation [30]. Whether this mechanism is applicable to all miRNAs or other species remains to be determined.

Through the ability to repress the expression of multiple genes, miRNAs play an important role in regulating many cellular activities such as proliferation, differentiation and apoptosis. Calin *et al.* provided the evidence showing the deletion of the miR-15/miR-16 cluster at 13q14 and its downregulation in patients with B cell chronic lymphocytic leukemia [31]. This was the first study suggesting that noncoding genes correlate with and may even contribute to oncogenesis. The same group further investigated the loci of 186 miRNAs in the human genome to evaluate their potential involvement in cancer pathogenesis [32]. They discovered that over 50% of these miRNA genes are present in the genomic regions with reported alterations in cancers. These are highly instable loci, including fragile sites, minimal heterozygous deletion regions, frequently amplified sections and common breakpoints. However, the instability of these miRNA-coding regions does not generate frequent miRNA somatic mutations within their seed sequences in cancers but rather changes their expression. A more recent study indicated that the genes encoding oncogenic miRNAs are mainly located in the amplified regions in human cancers, whereas the majority of genes for tumor suppressive miRNAs are in the deleted regions [33]. Interestingly, many oncogenes can produce alternative mRNA isoforms with shorter 3'-UTR sequences through a mechanism involved in alternative cleavage and polyadenylation [34]. A short mRNA isoform of an oncogene can avoid miRNA-mediated inhibition, which consequently increases its stability and typically produces ten-fold more protein. Overall, miRNA expression is globally reduced in tumors compared to their matched normal tissues [35]. Thus, numerous studies have demonstrated the potential of using the expression profiles of single or multiple miRNAs as biomarkers to classify tumor origins, stages and clinical outcomes [36–39]. Currently, microRNA expression detection is not used clinically in cancer diagnosis and prognosis, but it bears great promise in multiple applications of cancer therapies. Especially, due to the high stability of miRNAs and advance of RNA purification techniques, miRNAs can be extracted from not only tumors samples exposed to variable treatments, including formalin-fixing and paraffin-embedding [40], but also serum and urine [41–44].

2.2. MicroRNAs in Oncogenesis

Many miRNAs have been shown to regulate cell survival and proliferation, angiogenesis, and epithelial-mesenchymal transition (EMT). The miRNAs associated with oncogenesis are also known as “oncomirs”. Depending on their major targets, oncomirs can be classified into the oncogenic and tumor suppressive miRNA groups. As the classification implies, tumor suppressive miRNAs repress protein-coding oncogenes whereas oncogenic miRNAs repress protein-coding tumor suppressors. Some miRNAs display both oncogenic and tumor suppressive activities, depending on the tissue and tumor contexts. Wang *et al.* demonstrated that oncogenic and tumor suppressor miRNAs show clearly

different patterns in many aspects, including evolutionary rates, expression patterns, chromosome distribution, molecule sizes and of course targets [33].

An example of miRNAs that have opposing effects is the miRNA-17-92 cluster that has been implicated in regulating cellular survival. The polycistronic miRNA-17-92 cluster (miR-17-92) is located on chromosome 13 open reading frame 25 (c13orf25) in the human genome and contains seven miRNAs (*miR-92-1*, *miR-19a*, *miR-20a*, *miR-19b*, *miR-18a, miR-17-5p and miR-17-3p*) [45]. Ota *et al.* reported that the chromosomal region 13q31-q32, where the miR-17-92 cluster resides, is amplified in malignant lymphoma [46]. He *et al.* observed that five miRNAs (*miR-92-1*, *miR-19a*, *miR-20a*, *miR-19b and miR-17-5p*) from this cluster are up-regulated in human B-cell lymphoma and cell lines [47]. Ectopic expression of this miRNA cluster promotes c-Myc-induced lymphomagenesis in mice suggesting the cooperation between mir-17-19b and c-Myc in oncogenesis. In another study, O'Donnell *et al.* identified an additional miRNA, miR-18a, at the miRNA-17-92 locus and demonstrated that c-Myc transcriptionally activates the expression of these 6 miRNAs (*miR-92-1*, *miR-19a*, *miR-20a*, *miR-19b, miR-17-5p* and miR-18a) [48]. Interestingly, miR17-5p and miR-20a negatively regulate E2F1 expression, which is also activated by c-Myc. E2F1 is a crucial driver of cell cycle progression from the G1 to S phase, but its expression can induce either pro-metastatic activity or cell apoptosis depending on its molecular contexts [49,50]. Thus, the ectopic miR-17-92 expression in a c-Myc background confers an overall pro-survival trait to the cells through eliminating E2F1-mediated apoptotic potential. This is corroborated by the aforementioned studies of He *et al.*, in which c-Myc and miR-19b overexpression promoted murine lymphomas without causing any indicative sign of apoptosis [47]. The presence of this interactive regulatory network suggests a tight control of c-Myc-mediated proliferation signal during oncogenesis.

Data from several other groups also suggest tumor suppressive activities of the miRNAs in this cluster. Hossain *et al.* reported that miR-17-5p acts as a tumor suppressor through repressing the expression of AIB1 (or NCOA3) and consequently reducing the proliferation of breast cancer cells [51]. AIB1 was named for its "amplified in breast cancer" and has been suggested to play an oncogenic role in mammary oncogenesis [52,53]. Interestingly, AIB1 functions as a coactivator of E2F1 to promote breast cancer cell proliferation [52]. Thus, through repressing both AIB1 and E2F1, miR-17-5p may "kill two birds with one stone" in reducing breast cancer cell proliferation.

Yu *et al.* reported a negative feedback regulation between miR-17-5p/miR-20a and cyclin D1 [54], an oncoprotein in breast cancer. Cyclin D1 inversely correlated with miR-17-5p/miR-20a levels in breast cancer samples and cell lines. In a negative regulatory loop, cyclin D1 induces miR-17-5p/miR-20a expression, while these two miRNAs target the cyclin D1 3'-UTR to limit its proliferative activities. The same group also reported that miR-17-5p/miR-20a repress cytokeratin 8 through inhibiting cyclin D1 [55]. Additionally, these two miRNAs target the 3'-UTR of the interleukin-8 mRNA. Thus, miR-17-5p/miR-20a regulates both cellular secretion and tumor microenvironment to block migration and invasion of neighboring cells in breast cancer [55]. The seventh member of this cluster, miR-17-3p, is the "passenger strand" of miR-17-5p that is also processed into a mature miRNA. Recent studies indicate that it represses the expression of vimentin and Mdm2, suggesting a tumor suppressive role in oncogenesis [56,57].

Recent studies revealed a tumor suppressive role of miR-101 in a variety of cancers, including prostate, bladder, breast and gastric cancers [58–62]. MiR-101 is frequently downregulated in these

cancers compared to their normal adjacent tissues. On the other hand, ectopic expression of miR-101 in various cancer cell lines resulted in decreased cellular proliferation, motility and invasiveness, indicating its antitumoral activities. The tumor suppressive function of miR-101 has been corroborated by its inhibitory effect on EZH2, an oncogene of many solid tumors [63]. Varambally *et al.* reported the miR-101 loss resulted in increased EZH2 expression [62]. Importantly, genomic loss of the miR-101 locus correlates with EZH2 overexpression in solid tumors. Cao *et al.* further demonstrated that miR-101 negatively regulates EZH2 expression in prostate cancer cells, while miR-101 expression is modulated by androgen receptor and HIF-1α/HIF-1β [58]. The negative regulation of EZH2 by miR-101 has been confirmed by the studies from several other groups [64–66]. Consistent with these observations, a recent report suggested a positive regulation of E-cadherin by miR-101 [67]. E-cadherin is frequently downregulated in aggressive cancers and its loss increases cell dissemination and cancer cell invasiveness. In the study by Qazi *et al.*, ectopic miR-101 restored E-cadherin expression in pancreatic ductal adenocarcinoma cells by reducing EZH2-mediated histone H3-K27 methylation. Overall, these studies suggest that miR-101 primarily targets oncogene EZH2 and its deletion contributes to oncogenesis through aberrant epigenetics caused by EZH2 overexpression.

The let-7 family was one of the first mammalian miRNAs to be discovered and consists of 13 members located on different genomic loci that are often lost in human cancers [68]. Reduced expression of the let-7 members is associated with more dedifferentiated and aggressive cancers [69]. These observations suggested a possible tumor suppressive role of the let-7 family members. Recent studies indicated that let-7 can target oncogenes RAS and c-Myc that have multiple potential binding sites of let-7 miRNAs in their 3'-UTRs. Johnson *et al.* detected reduced let-7 expression in tumor tissues compared to normal adjacent tissues in lung cancer samples [70]. Interestingly, let-7 miRNA levels inversely correlated with RAS protein expression but not its mRNA levels, suggesting a mechanism of let-7-mediated translational inhibition without mRNA degradation. Overall, let-7 miRNAs may target several key oncogenes, such as RAS and c-Myc, to suppress the proliferative signals from these two oncogenes.

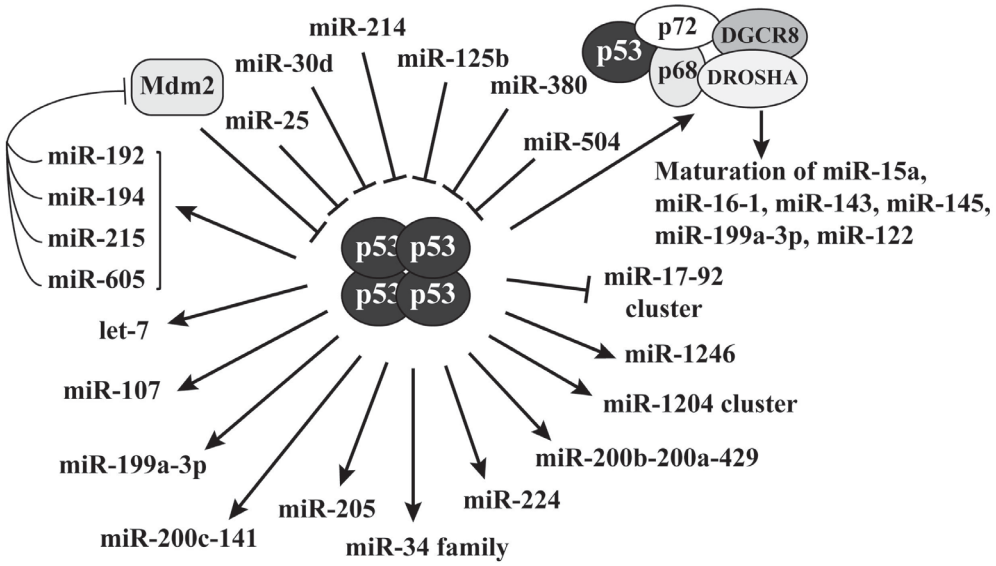
Many other miRNAs have been demonstrated to play a role in cancer development and progression. It is noteworthy that reports from different groups may present paradoxical results. These discrepancies could result from different experimental settings and also reflect the complexity of miRNA-regulated network in cancers. As discussed above, each miRNA potentially target hundreds genes and its ectopic overexpression may perturb multiple cellular processes leading to artificial phenotypic changes. Thus, studies only using miRNA overexpression or reporter assay without any miRNA depletion or tumor sample correlation studies may not truly represent physiological relevance. Nevertheless, currently available literature unequivocally indicates the prognostic potential and biological activities of miRNAs in different cancers.

2.3. P53 Is a Key Regulator of MicroRNAs Biogenesis

The tumor suppressor p53 is one of the best-studied proteins in the field of cancer research. Owing its well-characterized regulation, the role of p53 in the miRNA network has been extensively explored in the past decade. As a transcription factor, p53 forms tetramers to bind to its consensus sites on target genes and mediates their transcription. Increasing evidence shows that p53 exerts its antiproliferative

activities at least partially through the transcriptional regulation of miRNA expression (Figure 2), in addition to its canonical tumor suppressive role [20,71]. The best characterized p53 target is the miR-34 family [72]. The miR-34 family includes miR-34a, -34b and -34c. MiR-34a is transcribed by chromosome 1, and miR-34b and -34c are transcribed by two proximal loci on chromosome 11 and controlled by the same promoter. These miRNAs exert tumor suppressive activities through repressing proliferative genes, such as c-Myc and BCL2, and their overall function is inducing apoptosis, cell cycle arrest or senescence [73–75]. Consistently, they are frequently silenced by promoter methylation in tumors [76]. Thus, the p53-promoted expression of the miR-34 family extended its activated tumor suppressive network.

Figure 2. The interplay between miRNAs and p53. While p53 regulates the gene expression of many miRNAs, its expression is also inhibited by miRNAs. When cells are exposed to genotoxic stresses, p53 activates four miRNAs that repress Mdm2 expression, which leads to p53 accumulation and cell cycle arrest or apoptosis. P53 protein also associates with the DROSHA complex to directly regulate miRNA maturation.



P53 also stimulates the expression of many other miRNAs that have antiproliferative activities. For example, p53-mediated miRNA expression plays a role in hypoxia. During hypoxia of tumor cells, the upregulated hypoxia inducing factor (HIF)-1α forms a heterodimer with HIF-1β to become a transcription factor HIF-1 that activates pro-angiogenic genes such as vascular endothelial growth factor A (VEGFA), to promote angiogenesis for tumor growth and metastasis [77,78]. Yamakuchi *et al.* reported that miR-107 reduces hypoxia signaling by inhibiting HIF-1α expression in human colon cancer cells [79]. As a transactivator, p53 promotes the expression of miR-107 to reduce HIF-1α levels, which quenches the hypoxic signal to block tumor angiogenesis. Among other p53-activated miRNAs, miR-145 represses c-Myc expression [80], and miR-200c/miR-141 and miR-200b/miR-200a/miR-429 inhibit EMT through downregulating ZEB1 and ZEB2 [81,82]. As discussed above, miR-17-92 cluster exhibits proliferative activities based on most literature. P53 represses the expression of the miR-17-92 cluster, opposite to its transactivating effects on most other miRNAs (Figure 2), and this regulation likely contributes to the p53-induced apoptosis [83].

In addition to transcriptional regulation, p53 is directly involved in the maturation process of miRNAs. DROSHA is a major component of the complex that processes pri-miRNA into pre-miRNA. P53 associates with DROSHA in a RNA dependent manner and facilitates DROSHA-mediated pri-miRNA processing [84]. Interestingly, although this process is independent of its transcriptional activity, transcriptionally inactive p53 mutants do not show this capability. Actually, these p53 mutants interfere with the assembly of the DROSHA complex and consequently attenuate miRNA processing. Currently, p53 has been reported to regulate the maturation of at least 6 miRNAs (miR-15a, miR-16-1, miR-143, miR-145, miR-199a, and miR-122) [84–87]. Whether p53 is involved in the processing of other miRNAs remains to be determined.

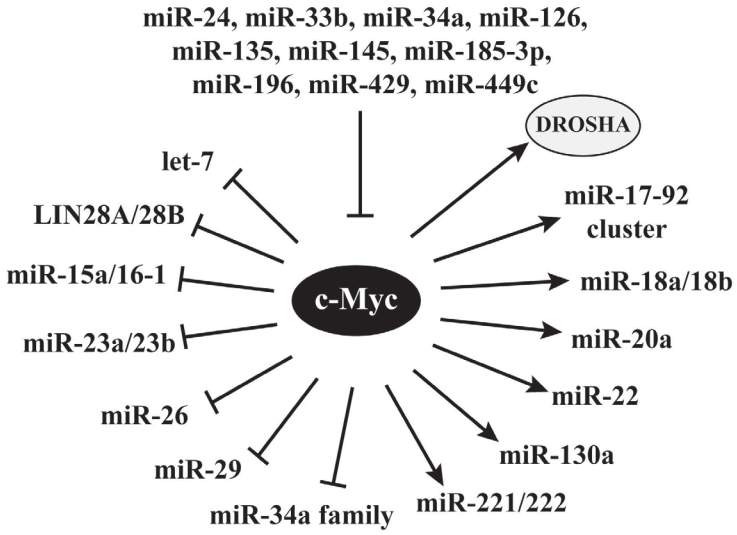
Conversely, p53 expression can also be regulated by miRNAs (Figure 2). The functions of these miRNAs are not uniformly oncogenic; thus, whether their role in repressing p53 expression is physiologically significant is unclear. There is a negative feedback loop between p53 and the ubiquitin E3 ligase Mdm2. While Mdm2 stimulates p53 ubiquitination and degradation, p53 activates Mdm2 gene expression. This feedback regulation is important to maintain p53 homeostasis in normal cells, which is disrupted when exposed to genotoxic stresses [88]. Recent studies reveal that some miRNAs can break this negative feedback loop leading to p53 accumulation. Pichiorri *et al.* reported that p53 activates the expression of miR-192, miR-194 and miR-215 and these miRNAs target the Mdm2 3'-UTR to repress its expression; thus, downregulation of p53-inducible miR-192, miR-194 and miR-215 causes aberrant Mdm2 increase and p53 downregulation in multiple myeloma [89]. Consistently, miR-192, miR-194 and miR-215 are downregulated in multiple myeloma and renal cancers [89–91]. Xiao *et al.* demonstrated that p53 activates miR-605 that also represses Mdm2 expression [92]. Thus, p53 activates these four miRNAs to downregulate its negative regulator Mdm2 leading to p53 accumulation in response to stress.

2.4. C-Myc Regulates the Synthesis of miRNAs

The oncogene c-Myc is a transcription factor and mediates target gene expression through recruiting various chromatin modifiers [93]. Recent studies reveal its role in regulating miRNA synthesis. Interestingly, the effects of c-Myc-mediated transcription of many miRNAs are opposite to those regulated by p53.

C-Myc represses the expression of multiple miRNAs (Figure 3). As discussed above, p53 transactivates the miR-34 family that represses c-Myc expression [72,75]. On the other hand, c-Myc represses the expression of the miRNAs in the miR-34a family [94]. C-Myc also recruits HDAC3 to downregulate the expression of miR-15a/miR-16-1 [95] that block the expression of multiple oncogenes, such as BCL2 and cyclin D1 [27,96]. Similarly, c-Myc also recruits HDAC3 and EZH2 to silence the expression of miR-29 [97] that has been defined as a tumor suppressor [98]. Among other c-Myc repressed miRNAs, let-7 inhibits androgen receptor and KRAS [68,99], miR-23a/miR-23b block glutaminase [100], and miR-26 targets EZH2 and cyclin D2 [58]. Thus, c-Myc-mediated repression of these miRNAs can release the expression of multiple oncogenes or proliferative genes to promote oncogenesis.

Figure 3. The interplay between miRNAs and c-Myc. While c-Myc regulates the expression of multiple miRNAs, its expression is inhibited by different miRNAs. C-Myc also activates the gene expression of DROSHA to directly promote miRNA procession.



c-Myc also activates the expression of multiple miRNAs. Similar to its role in antagonizing p53-mediated miR-34 miRNA expression, c-Myc promotes the transcription of miR-17-92 cluster [101] that is downregulated by p53. Kim *et al.* carried out a combined analysis of mRNA and miRNA expression profiles that revealed multiple c-Myc-induced miRNAs and their downstream targets [102]. Among their discovered miRNAs, miR-221 and miR-222 have been shown to target p27, p57 and PTEN, and exhibit proliferative activities [103–106]. For the other c-Myc-activated miRNAs, the authors predicted that miR-20a targets p21, RB1, PTEN and interferon regulatory factor 1 (IRF-1), and miR-130a inhibits tuberous sclerosis 1 (TSC1) and CYLD [102]. While p53 protein direct associates with the miRNA processing machinery, c-Myc activates DROSHA expression through binding to the E-box in its promoter and consequently facilitates miRNA processing [107].

While c-Myc acts as a regulator of miRNAs, its expression has been reported to be modulated by a number of miRNAs (Figure 3). Among them, miR-34a, miR-126 and miR-145 have been documented as tumor suppressors [33].

3. Long Noncoding RNA

Long noncoding RNAs (lncRNAs) were previously defined as RNA molecules longer than 200 nucleotides that are not translated into proteins [6,108]. Recently, Spizzo *et al.* amended this definition linking to the biological functions and described that lncRNAs are a class of RNA molecules that do not fit into any known class of small and structural RNAs, and possess regulatory roles in their primary or spliced form [109]. The GENCODE Consortium is a part of the ENCODE (ENCyclopedia Of DNA Elements) project and aims to identify all gene features in the human genome. The GENCODE 7 release in 2012 indicates that the human genome contains at least 9,640 long noncoding RNA loci that can potentially encode 15,512 transcripts [110]. LncRNA production is regulated by mechanisms similar to these of protein-coding genes, such as histone modifications and RNA splicing, and their expression shows tissue-specific patterns [111]. Most lncRNAs are localized in nucleus and

associated with chromatin, and some of them are preferentially processed into small RNAs [111]. Banfai *et al.* demonstrated that lncRNAs are rarely translated in two tested human cell lines, suggesting that ribosomes can differentiate the coding and noncoding transcripts for translation [112].

Recent studies continue elucidating novel and unpredicted biological activities of the lncRNAs, which were previously ascribed as protein functions. Based on the regulatory mechanisms, the lncRNAs can be divided into four categories, including the lncRNAs that (1) regulate gene expression, (2) act as miRNA decoys to free target mRNAs, (3) regulate mRNA translation, and (4) regulate protein activities.

3.1. LncRNAs Regulating Gene Expression

The X-inactive-specific transcript (Xist) was one of the first lncRNAs discovered in mammals [113]. Xist is encoded by the inactive X chromosome (Xi) and the genomic locus can transcribe a 17- to 20-kb RNA. This lncRNA binds the Xi *in cis* to induce chromosome X silencing by recruiting the polycomb repressive complex 2 (PRC2) that induces histone H3-K27 methylation, a hallmark of gene inactivation [114,115]. In this process, the transcription repressor Yin Yang 1 (YY1) confers allele-specific binding of Xist to the Xi with the involvement of two other noncoding RNAs, Jpx [116] and Ftx [117]. There is another lncRNA, Tsix, that is transcribed at the same locus of Xist but in the reverse direction and thus antisense to Xist [118]. Tsix regulates imprinted and random X inactivation in development [119]. An additional Xist-related RNA transcript is Xite that also plays a role in the X chromosome inactivation [120].

The association of epigenetic silencing complexes with Xist to induce transcriptional silencing has been extended to several recently characterized lncRNAs, which also associate with the PRC2 and other chromatin repressive complexes [121,122]. One of these lncRNAs is HOX Antisense Intergenic RNA (HOTAIR) that is a 2.2-kb transcript located at the HOXC gene cluster on chromosome 12 [122]. HOTAIR also modulates gene expression through epigenetic regulation. Unlike Xist that acts *in cis*, HOTAIR functions *in trans* to recruit the PRC2 to the *HOXD* locus on chromosome 2 to induce transcriptional silencing [122]. In addition to associating with PRC2, HOTAIR also interacts with the LSD1/CoREST/REST histone modification complex, leading to both histone H3-K27 methylation and H3-K4 demethylation [123]. Since HOTAIR plays an important role in the epigenetic regulation of its target genes, it is not surprising that its deregulation has been observed in different types of cancers. Recent studies suggest that HOTAIR overexpression is positively associated with increased tumor cell malignancy. Gupta *et al.* reported that HOTAIR is overexpressed in both primary and metastatic breast cancer tissues, and its levels in the primary tumors could be used as a significant predictor of subsequent tumor metastasis and survival of the patients [124]. Ectopically expressed HOTAIR could confer the breast epithelial cancer cells with invasive and metastatic potential while its depletion in breast cancer cells abrogated these activities [124]. The role of HOTAIR in promoting oncogenesis has also been reported in other cancers. Yang *et al.* compared HOTAIR levels between tumorigenic and adjacent non-tumorigenic tissues of hepatocellular carcinoma (HCC) samples and found that this lncRNA was expressed at higher levels in malignant tissues [125]. In addition, the survival analysis of a cohort consisting of 60 HCC patients revealed that high HOTAIR expression could serve as an independent prognostic marker for disease recurrence and reduced patient survival [125]. Another

HCC-related study suggested that HOTAIR expression is a potential biomarker for lymph node metastasis from the primary tumors [126]. Furthermore, HOTAIR upregulation was also observed in a cohort of patients diagnosed with the stage IV colorectal cancer (CRC) [127]. In this study, Kogo *et al.* indicated that high HOTAIR expression showed significantly positive correlation with the liver metastasis and poor patient outcome, further supporting that HOTAIR expression is a potential prognostic marker of multiple cancers. In pancreatic cancer, HOTAIR levels were also increased in tumorigenic tissues compared to the non-tumorigenic tissues, and associated with a more aggressive phenotype [128]. The oncogenic role of HOTAIR in pancreatic cancer cell invasion was validated by its siRNA-mediated knockdown and overexpression studies [128], consistent with the observations in the other aforementioned cancers. Interestingly, gene array studies showed only a small overlap of HOTAIR-regulated genes between pancreatic cancer and breast cancer [128], suggesting that HOTAIR may regulate different sets of target genes in a cell type-specific manner.

EZH2 is a core component of the PRC2. EZH2 knockdown followed by chromatin immunoprecipitation demonstrated that HOTAIR-mediated gene repression could be either PRC2-depedent or -independent in pancreatic cancer cells, although the PRC2 is necessary for HOTAIR target gene repression in breast cancer cell [124,128]. This discrepancy between the two cancer types suggests the presence of yet unidentified epigenetic mechanisms regulating HOTAIR-mediated transcriptional silencing.

Wang *et al.* identified the lincRNA HOTTIP that is transcribed from the 5' tip of the HOXA locus and modulates the activity of the WDR5-MLL complex, in which the WD40-repeat protein WDR5 binds the MLL complex to activate its histone H3-K4 methyltransferase activity [129]. Chromosomal looping can make HOTTIP stay in the vicinity of its target genes and let it bind WDR5 to promote the WDR5/MLL complexes-mediated histone H3-K4 methylation, which leads to target gene activation. Thus, HOTTIP serves as a key intermediate to transmit information from higher order chromosomal looping into chromatin modifications [129]. Another HOXA-related lncRNA is HOX antisense intergenic RNA myeloid 1 (HOTAIRM1) that is transcribed at a direction antisense to the HOXA gene [130]. The knockdown of HOTAIRM1 reduced the expression of HOXA1 and HOXA4 during the myeloid differentiation in promyelocytic leukemia cells. Whether this lncRNA acts as a miRNA decoy to promote the expression of these HOX genes remains to be determined.

Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), also known as nuclear-enriched abundant transcript 2 (NEAT2), is one of the first identified cancer-associated lncRNAs [131]. MALAT1 is a highly conserved noncoding transcript of over 8000 nts encoded by a locus on chromosome 11. It was initially recognized as a prognostic marker of increased metastatic risk for the patients of non-small cell lung carcinoma (NSCLC). Subsequent studies revealed that MALAT1 is localized in nuclear structures enriched with splicing and transcription factors, known as nuclear speckles, suggesting that this lncRNA may modulate alternative splicing of target genes [132,133]. Using RNAi-mediated depletion for the components of the nuclear speckles, Miyagawa *et al.* demonstrated that the pre-mRNA splicing activator RNPS1, the splicing coactivator SRm160 and the spliceosomal intron binding protein IBP160 promote MALAT1 localization to the nuclear speckles [134]. Furthermore, MALAT1 depletion and delocalization from the nuclear speckles resulted in downregulation of two interferon-induced genes (OASL and IFI44) and a potential celiac disease susceptibility gene, SPINK4 [134].

Despite these studies suggesting a role of MALAT1 in RNA splicing, a recent study by Gutschner *et al.* showed that this lncRNA regulates gene expression but not alternative splicing in lung cancer cells [135]. MALAT1 knockout was achieved by genomic integration of RNA destabilizing elements using zinc finger nucleases, leading to 1000-fold MALAT1 reduction. Using these MALAT1-knockout cells, the authors demonstrated that MALAT1 represses anti-metastatic genes and activates pro-metastatic genes in lung cancer cells, but does not affect genes regulating cell growth. In a xenograft mouse model, the MALAT1-depleted lung cancer cells showed reduced tumor formation compared to the cells with the intact MALAT1. This observation is in contrast to the findings reported by Yang *et al.* showing that MALAT1 cooperates with Polycomb 2 protein (Pc2) in regulating the activation of the growth-control gene program in 293T cells [136]. The different cell types employed in their experiments might contribute to the discrepancy between the two studies. Consistent with this prediction, MALAT1 is widely expressed in most normal human tissues, such as pancreas and lung, but absent in several other tissues including skin, stomach, bone marrow and uterus [131]. This suggests that this lncRNA may possess tissue specific functions. In addition to its role in regulating lung cancer metastasis, MALAT1 is upregulated in uterine endometrial stromal sarcoma [137], cervical cancer [4] and hepatocellular carcinoma [138], while its expression in the corresponding healthy tissues is undetectable or intermediate [131]. Although the oncogenic role of MALAT1 in different cancers has been demonstrated by correlational and functional studies, the molecular mechanisms underlying its activities in regulating gene expression and RNA splicing remain undetermined. To date, several studies suggested the essential role of the 3' end sequence and structure to its metastasis-promoting function, nuclear localization and stability [139–142].

ANRIL is a large antisense ncRNA of the INK4b/ARF/INK4a locus [143]. Yap *et al.* demonstrated that ANRIL binds chromobox 7 (CBX7), a component of the polycomb repressive complex 1. This interaction contributes to the role of CBX in promoting EZH2-mediated H3-K27 methylation at the INK4b/ARF/INK4a locus and consequently represses the tumor suppresser INK4a gene. Consistently, both CBX7 and ANRIL are increasingly expressed in prostate cancer [144].

GAS5 (growth arrest-specific transcript 5) is a lncRNA regulating growth arrest of T-cells and lymphocytes [145]. Ectopic GAS5 increases apoptosis and reduces cell cycle progression. Consistently, its downregulation inhibits apoptosis and promotes cell cycle. Kino *et al.* investigated the mechanism underlying the growth suppressive activities of GAS5 and discovered its role in blocking gene expression mediated by glucocorticoid receptor (GR) [146]. GAS5 binds to the DNA-binding domain of GR and thus prevents its association with the glucocorticoid response element of the GR target genes with anti-apoptotic activities, such as inhibitor of apoptosis 2 (cIAP2). Thus, abundantly expressed GAS5 during starvation can sensitize the cells to apoptosis. The nonsense-mediated mRNA decay (NMD) is a system that controls the quality of gene transcripts and reduces errors in gene expression by eliminating RNAs with premature stop codons [147]. Meanwhile, this mechanism also regulates the abundance of cellular transcripts, including ncRNAs. Recently, Zhang *et al.* demonstrated a reciprocally negative regulation between GAS5 and miR-21 [148]. While miR-21 represses GAS5 by targeting a sequence encoded by its exon 4, GAS5 inhibits miR-21 expression. Thus, GAS5 antagonizes the oncogenic activity of miR-21 [149] through reducing its cellular levels. Consistently, miR-21 and GAS5 showed negative correlation in breast cancer specimens [148]. The tumor suppressive role of GAS5 is supported by the identification of genetic susceptibility of its genomic

locus, 1q25, to several cancers, including melanoma [150], prostate cancer [151], breast cancer [152,153], colorectal cancer [154] and B-cell lymphoma [155]. Additionally, Tani *et al.* indicated that GAS5 can be stabilized with the depletion of UPF1, an essential component of NMD or during starvation [156] and the GAS5 introns encode multiple snoRNAs [157–160].

Several other lncRNAs have also been demonstrated to regulate chromatin remodeling and gene transcription. PTENP1 is the PTEN pseudogene and encodes two antisense RNA (asRNA) transcripts, asRNA α and β [161]. The α asRNA isoform can recruit DNMT3A, EZH2 and G9A to the PTEN promoter and repress its transcription. This will be further discussed below with other regulatory mechanisms of PTENP1. Evf2 is a polyadenylated lncRNA identified in embryonic brain cells [162]. This lncRNA regulates the transcription of homeodomain transcription factors DLX5 and DLX6 through recruiting DLX and MECP2 to the DNA regulatory elements in the intergenic region of these two genes. As a p53 transactivated lncRNA, lincRNA-p21 is a key mediator of p53-dependent gene repression through a mechanism of recruiting heterogeneous nuclear ribonucleoprotein K (HNRNP) to these p53 target genes [163]. Thus, inhibition of lincRNA-p21 affects the expression of a number of p53 repressed genes. Sheik *et al.* identified an Oct4-activated lncRNA, AK028326, and discovered that this lncRNA regulates pluripotency in mouse embryonic stem cells [164]. Interestingly, AK028326 activates Oct4 expression in a regulatory feedback loop.

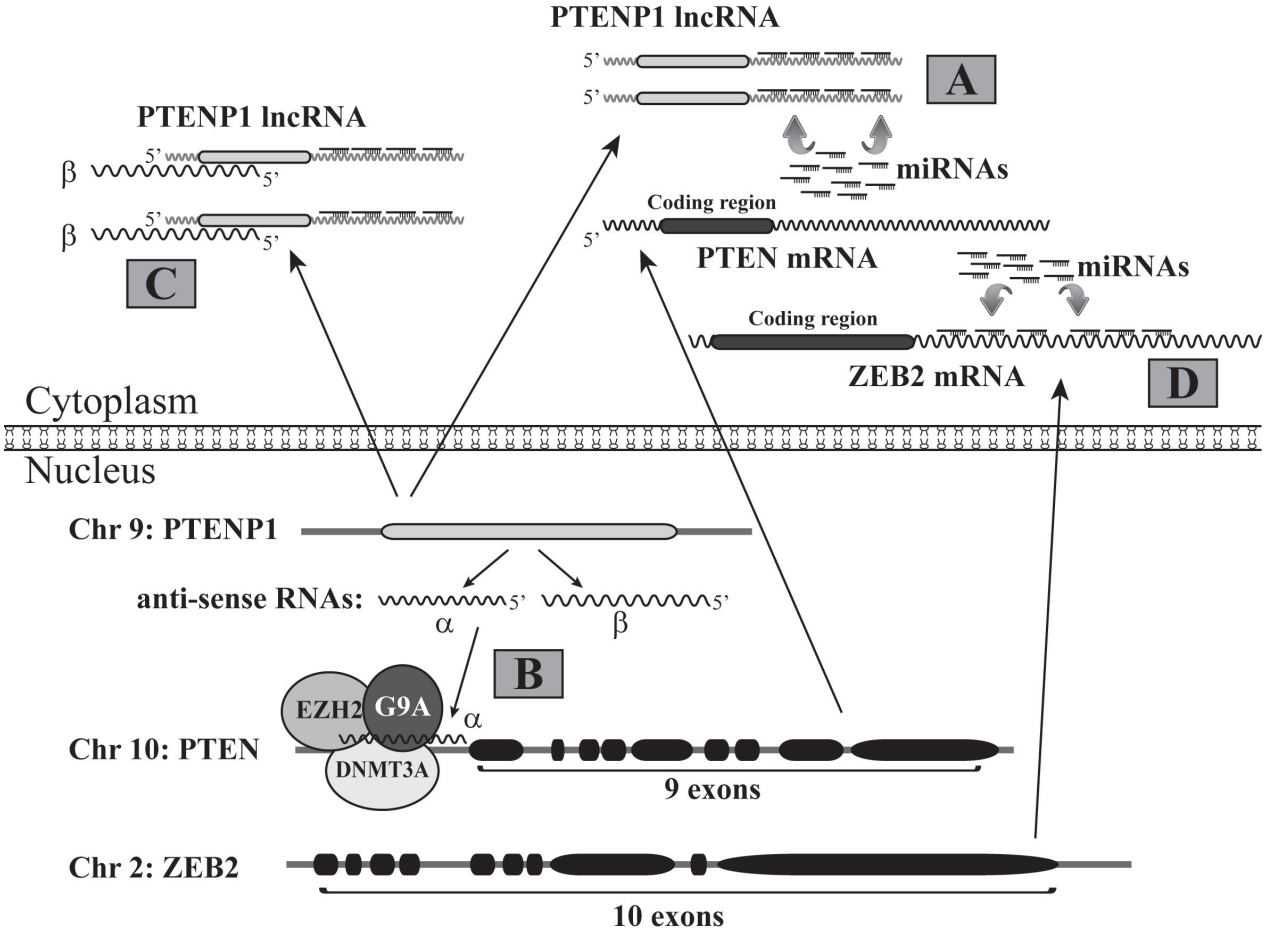
3.2. LncRNAs Acting as miRNA Decoys to Free Target mRNAs

Since the regulation of gene expression by miRNAs was revealed, researchers have been using miRNA sponges, RNA molecules containing the target sequence or reverse complementary sequence of a miRNA to be sponged, as a tool to inhibit the function of miRNAs and release their target gene expression [165]. Recent studies suggest that this approach naturally exists in cancers to modulate tumor suppressor and oncogene levels. In some literature, these decoy RNAs are also named as competitive endogenous RNAs (ceRNAs).

PTEN (phosphatase and tensin homolog) is a well characterized tumor suppressor with phosphatase activity. It is encoded at the 10q23.3 locus on chromosome 10 and frequently inactivated through diverse mechanisms in human cancers, highlighting its crucial role in oncogenesis. The 3'-UTR of the PTEN mRNA has 3,329 nts, markedly longer than the average 3'-UTR length (740 nts) of eukaryotic mRNAs [166], implicating its vulnerability as a target of miRNAs. Thus, while PTEN inactivation can be achieved by gene deletion and epigenetic silencing in cancers, its expression is also regulated by multiple miRNAs. For example, the PTEN 3'-UTR contains potential binding sites for over 10 miRNAs overexpressed in glioblastoma multiforme, which is more than 2 times higher than any other tumor suppressor [167]. PTEN expression can be repressed by miR-21, miR-221 and miR-222 [106,168]. Recent studies revealed novel mechanisms regulating PTEN expression through its pseudogene PTENP1 (also called PTH2 or ψPTEN). Pseudogenes are dysfunctional relatives of their cognate genes but have lost the protein-coding ability due to premature stop codons, deletions/insertions or frameshift mutations, and thus cannot be translated into functional proteins [169]. The PTEN pseudogene PTENP1 is highly transcribed in certain tissues and cells, suggesting that this lncRNA may have biological activities [170]. The functional relationship between the PTEN and its pseudogene was first discovered by Poliseno *et al.* [171]. In their study, the authors demonstrated that PTENP1 modulates endogenous PTEN transcript levels by acting as a molecular sponge for

PTEN-targeting miRNAs; thus, the PTENP1 transcript serves as an effective decoy and exerts tumor suppressive functions (Figure 4). This novel regulatory role of the PTENP1 can be extended to KRAS1P, the pseudogene of the oncogene KRAS. In a study by Poliseno *et al.*, the overexpression of KRAS1P led to increased KRAS mRNA levels in prostate cancer DU145 cells through a mechanism of sequestering KRAS-targeting microRNAs, and consequently promoted cell proliferation [171].

Figure 4. PTEN expression is regulated by multiple ncRNAs. (**A**) The PTEN pseudogene, PTENP1, can transcribe into the PTENP1 lncRNA that acts as a decoy to sponge the miRNAs targeting at the 3'-UTR of the PTEN mRNA; (**B**,**C**) The locus of the PTEN pseudogene can also transcribe from the reverse direction to make two antisense RNA (asRNA) isoforms, α and β; (**B**) The α asRNA isoform binds the PTEN promoter and recruits DNMT3A, EZH2 and G9A, which causes epigenetic silencing of the PTEN gene; (**C**) The β asRNA isoform associates with the PTENP1 lncRNA to increase its stability and miRNA decoy activity; (**D**) The ZEB2 mRNA has a long 3'-UTR with many binding sites of miRNAs that also potentially target the PTEN 3'-UTR. Thus, the ZEB2 mRNA acts as a decoy to sponge many miRNAs and consequently promotes PTEN expression.



As briefly discussed above, a recent study by Johnsson *et al.* provided evidence of another regulatory mechanism of PTEN by its pseudogene PTENP1 (Figure 4). The authors discovered that the PTENP1 locus can be transcribed from a reverse direction to create two isoforms of an antisense RNA (asRNA), α and β [161]. The α isoform of the asRNA binds the PTEN promoter and recruits

DNMT3A and two histone methyltransferases EZH2 and G9A, which mediate the methylation of histone H3-K27 and H3-K9, respectively, two well-characterized markers of gene repression [172]; thus, the PTENP1 α asRNA negatively regulates PTEN gene expression through promoting the epigenetic silencing of its promoter. The β asRNA isoform interacts with the PTENP1 lncRNA through RNA-RNA pairing, which can maintain stable PTENP1 lncRNA levels in cytoplasm, increase its stability and facilitate its role as a microRNA sponge; thus, the PTENP1 β asRNA activates PTEN expression through facilitating the decoy activity of the PTENP1 lncRNA. Overall, the PTENP1 α and β asRNAs exhibit oncogenic and tumor suppressive roles, respectively, based on their effects on PTEN expression (Figure 4). Whether the two asRNA isoforms are differentially expressed in human cancers remains to be determined.

Interestingly, PTEN expression can also be regulated at the translational level by the sponge effect of the transcript from another gene. The full length ZEB2 mRNA has a very long 3'-UTR (over 5,000 nts) and contains multiple potential binding sites of miRNAs that can also target the PTEN 3'-UTR, such as miR-181a and miR-200/miR-141 [173]. Thus, the ZEB2 mRNA serves as a decoy or ceRNA of the PTEN mRNA (Figure 4) and reduced ZEB2 expression activates the PI3K/AKT pathway through downregulating PTEN.

There are several other examples showing that lncRNAs act as decoys to stabilize mRNAs. BACE-AS regulates the expression of β-secretase-1 (BACE1), a crucial enzyme in Alzheimer's disease pathophysiology [174]. Linc-MD1 is a muscle-specific lncRNA and activates the expression of MAML1 and MEF2C through its decoy role for miR-133 [175].

3.3. LncRNAs Regulating mRNA Translation

As discussed above, MALAT1 was initially demonstrated to possess activities in regulating alternative splicing. MALAT1 associates with serine/arginine (SR) splicing factors at the nuclear speckle domains of nucleus and is involved in the process of alternative splicing [133]. Thus, MALAT1 depletion or ectopic SR protein expression affects the alternative splicing of a similar set of pre-mRNAs. Additionally, MALAT1 alters the phosphorylation of the SR splicing factors, which is essential to their activities in regulating alternative splicing. However, this regulation may only be present in specific cell types or under particular conditions, because MALAT1-knockout mice were viable and fertile, and showed regularly localized nuclear speckle markers [176]. Whether this regulatory mechanism contributes to the activity of MALAT1 in promoting tumor metastasis has not been determined.

A recent study from Zhang *et al.* demonstrated the lncRNA, lincRNA-RoR, acts as a strong negative regulator of p53 [177]. LincRNA-RoR reduces p53 expression in cells exposed to DNA damage stress, but not in unstressed cells, through directly binding to the heterogeneous nuclear ribonucleoprotein I (hnRNPI) and repressing p53 mRNA translation. As an autoregulatory feedback regulation, p53 transcriptionally induces lincRNA-RoR expression. The functional interplay between lincRNA-RoR and p53 may serve as additional surveillance to mediate cell response to genotoxic stresses.

As discussed above, another p53-transactivated lncRNA, lincRNA-p21, is involved in p53-dependent gene transrepression [163]. A recent study from Yoon *et al.* demonstrated that lincRNA-p21 also modulates translation [178]. This lncRNA can associate with JunB and β-catenin mRNAs to

reduce their translation rates. A RNA-binding protein, HuR, can recruit let-7/AGO2 to lincRNA-p21 to reduce its stability; thus, elevated HuR releases lincRNA-p21-mediated repression of JunB and β-catenin expression. Since lincRNA-p21 is transactivated by p53, the negative regulation of JunB and β-catenin by lincRNA-p21 is consistent to the tumor suppressive role of p53.

3.4. LncRNAs Regulating Protein Activities

Telomeres are the DNA-protein complexes at the end of eukaryotic chromosomes and essential to chromosome stability. As a frequently activated reverse transcriptase, telomerase adds DNA sequence repeats ("TTAGGG" in vertebrates) to the telomere regions of chromosome to maintain the telomere length. Azzalin *et al.* discovered that telomeres can be transcribed into telomeric repeat-containing RNA (TERRA) [179]. These molecules have different lengths and were predicted to play a role in the maintenance of telomere integrity through an unclear mechanism. A later study from Redon *et al.* provided a possible mechanism underlying this activity of TERRA [180]. The authors demonstrated that TERRA binds telomerase and thus acts as a potent competitive inhibitor for the telomeric DNA. Consistently, TERRA expression is significantly downregulated in multiple tumor cell lines, which plays a role in telomere maintenance and cell immortalization of cancer cells.

Wang *et al.* demonstrated an indirect regulation of the histone acetyltransferase activities of CBP and p300 by ncRNAs [181]. The 5'-regulatory region of the cyclin D1 gene encodes at least four ncRNAs. In response to ionizing radiation, these ncRNAs bind TLS (translocated in liposarcoma) protein at the chromatin of the cyclin D1 promoter region and cause an allosteric effect on this protein. This TLS conformational change promotes its activity of inhibiting CBP/p300-mediated histone acetylation and consequently silences cyclin D1 gene. These data indicate that ncRNAs encoded by a promoter can act as selective ligands to modulate the activities of transcription cofactors in response to genotoxic stresses.

LncRNAs can also regulate protein activity through altering their subcellular localization. Nuclear factor of activated T-cells (NFAT) represents a family of transcription factors regulating immune response. Some of NFAT proteins, such as NFAT1 and NFAT5, contribute to tumor metastasis and cell motility [182,183]. The noncoding repressor of NFAT (NRON) is a ncRNA associated with multiple proteins [184]. NRON binds phosphorylated NFAT1 to sequester NFAT in the cytoplasm. The depletion of NRON leads to NFAT dephosphorylation and nuclear import. Thus, ncRNAs can be a part of a scaffold to trap a latent transcription factor [185] and regulate the expression of its target genes.

4. Circular RNAs

Recent advances in high-throughput sequencing coupled with powerful computational analyses of expression data have allowed researchers to identify and characterize new RNA species, one of which is a new class of RNA molecules, circular RNAs (circRNAs), present in humans and animals. CircRNAs were first discovered in plants and considered as viroids due to their predicted role as subviral agents [186]. These covalently linked and single-stranded ncRNAs have molecular weights of over 100,000 Da and are highly thermal stable. Initially, most circRNA molecules were dismissed as transcription noise or by-products generated by cellular splicing machineries. However, this notion has been challenged by emerging studies suggesting that some circRNAs are evolutionary conserved in

human and mice [14,187]. The exact mechanism of circRNA biogenesis remains to be elucidated. Salzman *et al.* indicated that circRNAs are produced by a non-canonical mode of RNA splicing [188]. Several other studies suggested that circRNA may be at least partially contributed by exon-skipping events, which can create an exon-containing lariat and then possibly undergo internal splicing to generate an exon circle [187,189,190].

In 2011, Hansen *et al.* reported the miRNA decoy function of a noncoding circRNA [191]. In this study, the authors demonstrated that miR-671 directly targets and cleaves a circular antisense transcript of the cerebellar degeneration-related protein 1 (CDR1) in an AGO2-dependent manner. In 2013, thousands of well-expressed and stable circRNAs with tissue- and developmental stage-specific expression were identified [14]. The antisense to the CDR1 transcript (CDR1as) is densely bound by Argonaute proteins and contains 63 conserved binding sites of miR-7, suggesting that this circRNA functions as a miRNA sponge of miR-7 to release its target gene expression [14,190]. Indeed, using a zebrafish model, Memczak *et al.* demonstrated that introduction of human CDR1as resulted in a phenotype similar to that of miR-7 knockdown, while injection of the miR-7 precursor partially reversed this phenotype, further implicating CDR1as as an antagonist of miR-7 [14]. The results of this study provided evidence that circRNAs may function as regulators of gene expression at the post-transcriptional level. In addition to these studies, Jeck *et al.* recently discovered over 25,000 circRNAs in human fibroblasts [187]. Due to the lack of exposed 5' and 3' ends, circRNAs are predicted to be resistant to degradation by cellular enzymes such as ribonucleases and thus more stable than linear RNAs. The increased stability of circRNAs makes these novel RNA molecules act as efficient miRNA sponges in modulating gene expression.

5. Conclusions

NcRNAs have gained intensive and growing attention for their potential as both regulators and biomarkers of cancers in the past two decades. We are now witnessing the beginning of a new era of identifying key players and determining their underlying mechanisms during oncogenesis. In this review, we are only able to summarize some studies of miRNAs, lncRNAs and circRNAs. We did not intentionally ignore other excellent reports, but just could not include them due to the enormous amounts of ncRNA-related studies and the limitation of this article. Many review papers have summarized the roles of ncRNAs in different diseases including cancers and some of them also discussed several other types of ncRNAs, such as snoRNA and small nuclear RNAs (snRNA).

Currently, many thousands of ncRNAs have been identified and their differential expression profiles in a variety of cancers or between normal and tumorigenic specimens have been demonstrated; however, only a small number of ncRNAs, especially lncRNAs, have been well characterized. Thus, what we have discovered is just a tip of the iceberg in this area. Understanding the functions and regulatory mechanisms of ncRNAs in cancer pathogenesis remains a fertile research field to be explored at least in next decade. Future studies are needed to dissect the upstream mechanisms regulating ncRNA expression and processing, and the downstream proteins or pathways mediated by ncRNAs at different stages of cancer development. These efforts may lead to the discovery of entirely novel regulatory mechanisms or advance our understanding for the currently recognized signaling pathways. The achievement of these ncRNA studies in cancers will definitely provide insights into

discovering new biomarkers for cancer diagnosis and prognosis, unraveling novel therapeutic targets, and developing unconventional therapeutic modalities to reduce cancer-related mortality.

Acknowledgments

We thank Steven J Kridel for critical reading of the manuscript. GD is supported by the Randi B. Weiss Cancer Fund. GS is supported by American Cancer Society (116403-RSG-09-082-01-MGO) and NIH R01 (5R01CA106314).

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Bertone, P.; Stolc, V.; Royce, T.E.; Rozowsky, J.S.; Urban, A.E.; Zhu, X.; Rinn, J.L.; Tongprasit, W.; Samanta, M.; Weissman, S.; *et al.* Global identification of human transcribed sequences with genome tiling arrays. *Science* **2004**, *306*, 2242–2246.
2. Birney, E.; Stamatoyannopoulos, J.A.; Dutta, A.; Guigo, R.; Gingeras, T.R.; Margulies, E.H.; Weng, Z.; Snyder, M.; Dermitzakis, E.T.; Thurman, R.E.; *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **2007**, *447*, 799–816.
3. Cheng, J.; Kapranov, P.; Drenkow, J.; Dike, S.; Brubaker, S.; Patel, S.; Long, J.; Stern, D.; Tammana, H.; Helt, G.; *et al.* Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* **2005**, *308*, 1149–1154.
4. Kapranov, P.; Cheng, J.; Dike, S.; Nix, D.A.; Duttagupta, R.; Willingham, A.T.; Stadler, P.F.; Hertel, J.; Hackermuller, J.; Hofacker, I.L.; *et al.* RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **2007**, *316*, 1484–1488.
5. Wilusz, J.E.; Sunwoo, H.; Spector, D.L. Long noncoding RNAs: Functional surprises from the RNA world. *Genes Dev.* **2009**, *23*, 1494–1504.
6. Mercer, T.R.; Dinger, M.E.; Mattick, J.S. Long non-coding RNAs: Insights into functions. *Nat. Rev. Genet.* **2009**, *10*, 155–159.
7. Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914.
8. Cooper, G.M. Cellular transforming genes. *Science* **1982**, *217*, 801–806.
9. Santos, E.; Tronick, S.R.; Aaronson, S.A.; Pulciani, S.; Barbacid, M. T24 human bladder carcinoma oncogene is an activated form of the normal human homologue of BALB- and Harvey-MSV transforming genes. *Nature* **1982**, *298*, 343–347.
10. Parada, L.F.; Tabin, C.J.; Shih, C.; Weinberg, R.A. Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus ras gene. *Nature* **1982**, *297*, 474–478.
11. Godbout, R.; Dryja, T.P.; Squire, J.; Gallie, B.L.; Phillips, R.A. Somatic inactivation of genes on chromosome 13 is a common event in retinoblastoma. *Nature* **1983**, *304*, 451–453.

12. Murphree, A.L.; Benedict, W.F. Retinoblastoma: Clues to human oncogenesis. *Science* **1984**, *223*, 1028–1033.
13. Esteller, M. Non-coding RNAs in human disease. *Nat. Rev. Genet.* **2011**, *12*, 861–874.
14. Memczak, S.; Jens, M.; Elefsinioti, A.; Torti, F.; Krueger, J.; Rybak, A.; Maier, L.; Mackowiak, S.D.; Gregersen, L.H.; Munschauer, M.; *et al.* Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **2013**, *495*, 333–338.
15. Kim, V.N. MicroRNA biogenesis: Coordinated cropping and dicing. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 376–385.
16. Czech, B.; Hannon, G.J. Small RNA sorting: Matchmaking for Argonautes. *Nat. Rev. Genet.* **2011**, *12*, 19–31.
17. Bartel, D.P. MicroRNAs: Target recognition and regulatory functions. *Cell* **2009**, *136*, 215–233.
18. Guo, L.; Lu, Z. The fate of miRNA* strand through evolutionary analysis: Implication for degradation as merely carrier strand or potential regulatory molecule? *PLoS One* **2010**, *5*, e11387.
19. Lewis, B.P.; Burge, C.B.; Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **2005**, *120*, 15–20.
20. Lujambio, A.; Lowe, S.W. The microcosmos of cancer. *Nature* **2012**, *482*, 347–355.
21. Krek, A.; Grun, D.; Poy, M.N.; Wolf, R.; Rosenberg, L.; Epstein, E.J.; MacMenamin, P.; da Piedade, I.; Gunsalus, K.C.; Stoffel, M.; *et al.* Combinatorial microRNA target predictions. *Nat. Genet.* **2005**, *37*, 495–500.
22. Lee, R.C.; Feinbaum, R.L.; Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **1993**, *75*, 843–854.
23. Wightman, B.; Ha, I.; Ruvkun, G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell* **1993**, *75*, 855–862.
24. Brennecke, J.; Hipfner, D.R.; Stark, A.; Russell, R.B.; Cohen, S.M. Bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. *Cell* **2003**, *113*, 25–36.
25. Chen, X. A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science* **2004**, *303*, 2022–2025.
26. Poy, M.N.; Eliasson, L.; Krutzfeldt, J.; Kuwajima, S.; Ma, X.; Macdonald, P.E.; Pfeffer, S.; Tuschl, T.; Rajewsky, N.; Rorsman, P.; *et al*. A pancreatic islet-specific microRNA regulates insulin secretion. *Nature* **2004**, *432*, 226–230.
27. Cimmino, A.; Calin, G.A.; Fabbri, M.; Iorio, M.V.; Ferracin, M.; Shimizu, M.; Wojcik, S.E.; Aqeilan, R.I.; Zupo, S.; Dono, M.; *et al.* miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13944–13949.
28. Valencia-Sanchez, M.A.; Liu, J.; Hannon, G.J.; Parker, R. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev*. **2006**, *20*, 515–524.
29. Huntzinger, E.; Izaurralde, E. Gene silencing by microRNAs: Contributions of translational repression and mRNA decay. *Nat. Rev. Genet.* **2011**, *12*, 99–110.
30. Djuranovic, S.; Nahvi, A.; Green, R. miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. *Science* **2012**, *336*, 237–240.
31. Calin, G.A.; Dumitru, C.D.; Shimizu, M.; Bichi, R.; Zupo, S.; Noch, E.; Aldler, H.; Rattan, S.; Keating, M.; Rai, K.; *et al*. Frequent deletions and down-regulation of micro- RNA genes miR15

and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15524–15529.

32. Calin, G.A.; Sevignani, C.; Dumitru, C.D.; Hyslop, T.; Noch, E.; Yendamuri, S.; Shimizu, M.; Rattan, S.; Bullrich, F.; Negrini, M.; *et al*. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 2999–3004.
33. Wang, D.; Qiu, C.; Zhang, H.; Wang, J.; Cui, Q.; Yin, Y. Human microRNA oncogenes and tumor suppressors show significantly different biological patterns: From functions to targets. *PLoS One* **2010**, *5*, e13067.
34. Mayr, C.; Bartel, D.P. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **2009**, *138*, 673–684.
35. Diederichs, S.; Haber, D.A. Sequence variations of microRNAs in human cancer: Alterations in predicted secondary structure do not affect processing. *Cancer Res.* **2006**, *66*, 6097–6104.
36. Lu, J.; Getz, G.; Miska, E.A.; Alvarez-Saavedra, E.; Lamb, J.; Peck, D.; Sweet-Cordero, A.; Ebert, B.L.; Mak, R.H.; Ferrando, A.A.; *et al*. MicroRNA expression profiles classify human cancers. *Nature* **2005**, *435*, 834–838.
37. Murakami, Y.; Yasuda, T.; Saigo, K.; Urashima, T.; Toyoda, H.; Okanoue, T.; Shimotohno, K. Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. *Oncogene* **2006**, *25*, 2537–2545.
38. Yanaihara, N.; Caplen, N.; Bowman, E.; Seike, M.; Kumamoto, K.; Yi, M.; Stephens, R.M.; Okamoto, A.; Yokota, J.; Tanaka, T.; *et al*. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* **2006**, *9*, 189–198.
39. Rosenfeld, N.; Aharonov, R.; Meiri, E.; Rosenwald, S.; Spector, Y.; Zepeniuk, M.; Benjamin, H.; Shabes, N.; Tabak, S.; Levy, A.; *et al.* MicroRNAs accurately identify cancer tissue origin. *Nat. Biotechnol.* **2008**, *26*, 462–469.
40. Xi, Y.; Nakajima, G.; Gavin, E.; Morris, C.G.; Kudo, K.; Hayashi, K.; Ju, J. Systematic analysis of microRNA expression of RNA extracted from fresh frozen and formalin-fixed paraffin-embedded samples. *RNA* **2007**, *13*, 1668–1674.
41. Mitchell, P.S.; Parkin, R.K.; Kroh, E.M.; Fritz, B.R.; Wyman, S.K.; Pogosova-Agadjanyan, E.L.; Peterson, A.; Noteboom, J.; O'Briant, K.C.; Allen, A.; *et al.* Circulating microRNAs as stable blood-based markers for cancer detection. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10513–10518.
42. Chan, M.; Liaw, C.S.; Ji, S.M.; Tan, H.H.; Wong, C.Y.; Thike, A.A.; Tan, P.H.; Ho, G.H.; Lee, A.S. Identification of circulating microRNA signatures for breast cancer detection. *Clin. Cancer Res.* **2013**, *19*, 4477–4487.
43. Chen, J.; Yao, D.; Li, Y.; Chen, H.; He, C.; Ding, N.; Lu, Y.; Ou, T.; Zhao, S.; Li, L.; *et al*. Serum microRNA expression levels can predict lymph node metastasis in patients with early-stage cervical squamous cell carcinoma. *Int. J. Mol. Med.* **2013**, *32*, 557–567.
44. Xiao, Y.F.; Yong, X.; Fan, Y.H.; Lu, M.H.; Yang, S.M.; Hu, C.J. microRNA detection in feces, sputum, pleural effusion and urine: Novel tools for cancer screening (Review). *Oncol. Rep.* **2013**, *30*, 535–544.
45. Xiang, J.; Wu, J. Feud or Friend? The role of the miR-17-92 cluster in tumorigenesis. *Curr. Genomics* **2010**, *11*, 129–135.

46. Ota, A.; Tagawa, H.; Karnan, S.; Tsuzuki, S.; Karpas, A.; Kira, S.; Yoshida, Y.; Seto, M. Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma. *Cancer Res.* **2004**, *64*, 3087–3095.
47. He, L.; Thomson, J.M.; Hemann, M.T.; Hernando-Monge, E.; Mu, D.; Goodson, S.; Powers, S.; Cordon-Cardo, C.; Lowe, S.W.; Hannon, G.J.; *et al*. A microRNA polycistron as a potential human oncogene. *Nature* **2005**, *435*, 828–833.
48. O'Donnell, K.A.; Wentzel, E.A.; Zeller, K.I.; Dang, C.V.; Mendell, J.T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **2005**, *435*, 839–843.
49. Engelmann, D.; Putzer, B.M. The dark side of E2F1: In transit beyond apoptosis. *Cancer Res.* **2012**, *72*, 571–575.
50. Biswas, A.K.; Johnson, D.G. Transcriptional and nontranscriptional functions of E2F1 in response to DNA damage. *Cancer Res.* **2012**, *72*, 13–17.
51. Hossain, A.; Kuo, M.T.; Saunders, G.F. Mir-17-5p regulates breast cancer cell proliferation by inhibiting translation of AIB1 mRNA. *Mol. Cell. Biol.* **2006**, *26*, 8191–8201.
52. Louie, M.C.; Zou, J.X.; Rabinovich, A.; Chen, H.W. ACTR/AIB1 functions as an E2F1 coactivator to promote breast cancer cell proliferation and antiestrogen resistance. *Mol. Cell. Biol.* **2004**, *24*, 5157–5171.
53. Qin, L.; Liao, L.; Redmond, A.; Young, L.; Yuan, Y.; Chen, H.; O'Malley, B.W.; Xu, J. The AIB1 oncogene promotes breast cancer metastasis by activation of PEA3-mediated matrix metalloproteinase 2 (MMP2) and MMP9 expression. *Mol. Cell. Biol.* **2008**, *28*, 5937–5950.
54. Yu, Z.; Wang, C.; Wang, M.; Li, Z.; Casimiro, M.C.; Liu, M.; Wu, K.; Whittle, J.; Ju, X.; Hyslop, T.; *et al*. A cyclin D1/microRNA 17/20 regulatory feedback loop in control of breast cancer cell proliferation. *J. Cell Biol.* **2008**, *182*, 509–517.
55. Yu, Z.; Willmarth, N.E.; Zhou, J.; Katiyar, S.; Wang, M.; Liu, Y.; McCue, P.A.; Quong, A.A.; Lisanti, M.P.; Pestell, R.G. microRNA 17/20 inhibits cellular invasion and tumor metastasis in breast cancer by heterotypic signaling. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 8231–8236.
56. Li, H.; Yang, B.B. Stress response of glioblastoma cells mediated by miR-17-5p targeting PTEN and the passenger strand miR-17-3p targeting MDM2. *Oncotarget* **2012**, *3*, 1653–1668.
57. Shan, S.W.; Fang, L.; Shatseva, T.; Rutnam, Z.J.; Yang, X.; Du, W.; Lu, W.Y.; Xuan, J.W.; Deng, Z.; Yang, B.B. Mature miR-17-5p and passenger miR-17-3p induce hepatocellular carcinoma by targeting PTEN, GalNT7 and vimentin in different signal pathways. *J. Cell Sci.* **2013**, *126*, 1517–1530.
58. Cao, P.; Deng, Z.; Wan, M.; Huang, W.; Cramer, S.D.; Xu, J.; Lei, M.; Sui, G. MicroRNA-101 negatively regulates Ezh2 and its expression is modulated by androgen receptor and HIF-1alpha/HIF-1beta. *Mol. Cancer* **2010**, *9*, doi:10.1186/1476-4598-9-108.
59. Wang, H.J.; Ruan, H.J.; He, X.J.; Ma, Y.Y.; Jiang, X.T.; Xia, Y.J.; Ye, Z.Y.; Tao, H.Q. MicroRNA-101 is down-regulated in gastric cancer and involved in cell migration and invasion. *Eur. J. Cancer* **2010**, *46*, 2295–2303.
60. Hu, Z.; Lin, Y.; Chen, H.; Mao, Y.; Wu, J.; Zhu, Y.; Xu, X.; Li, S.; Zheng, X.; Xie, L. MicroRNA-101 suppresses motility of bladder cancer cells by targeting c-Met. *Biochem. Biophys. Res. Commun.* **2013**, *435*, 82–87.

61. Wang, R.; Wang, H.B.; Hao, C.J.; Cui, Y.; Han, X.C.; Hu, Y.; Li, F.F.; Xia, H.F.; Ma, X. MiR-101 is involved in human breast carcinogenesis by targeting Stathmin1. *PLoS One* **2012**, *7*, e46173.
62. Varambally, S.; Cao, Q.; Mani, R.S.; Shankar, S.; Wang, X.; Ateeq, B.; Laxman, B.; Cao, X.; Jing, X.; Ramnarayanan, K.; *et al.* Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. *Science* **2008**, *322*, 1695–1699.
63. Sellers, W.R.; Loda, M. The EZH2 polycomb transcriptional repressor—A marker or mover of metastatic prostate cancer? *Cancer Cell* **2002**, *2*, 349–350.
64. Friedman, J.M.; Liang, G.; Liu, C.C.; Wolff, E.M.; Tsai, Y.C.; Ye, W.; Zhou, X.; Jones, P.A. The putative tumor suppressor microRNA-101 modulates the cancer epigenome by repressing the polycomb group protein EZH2. *Cancer Res.* **2009**, *69*, 2623–2629.
65. Banerjee, R.; Mani, R.S.; Russo, N.; Scanlon, C.S.; Tsodikov, A.; Jing, X.; Cao, Q.; Palanisamy, N.; Metwally, T.; Inglehart, R.C.; *et al.* The tumor suppressor gene rap1GAP is silenced by miR-101-mediated EZH2 overexpression in invasive squamous cell carcinoma. *Oncogene* **2011**, *30*, 4339–4349.
66. Kottakis, F.; Polytarchou, C.; Foltopoulou, P.; Sanidas, I.; Kampranis, S.C.; Tsichlis, P.N. FGF-2 regulates cell proliferation, migration, and angiogenesis through an NDY1/KDM2B-miR-101-EZH2 pathway. *Mol. Cell* **2011**, *43*, 285–298.
67. Qazi, A.M.; Gruzdyn, O.; Semaan, A.; Seward, S.; Chamala, S.; Dhulipala, V.; Sethi, S.; Ali-Fehmi, R.; Philip, P.A.; Bouwman, D.L.; *et al*. Restoration of E-cadherin expression in pancreatic ductal adenocarcinoma treated with microRNA-101. *Surgery* **2012**, *152*, 704–711, discussion 711–713.
68. Nadiminty, N.; Tummala, R.; Lou, W.; Zhu, Y.; Zhang, J.; Chen, X.; eVere White, R.W.; Kung, H.J.; Evans, C.P.; Gao, A.C. MicroRNA let-7c suppresses androgen receptor expression and activity via regulation of Myc expression in prostate cancer cells. *J. Biol. Chem.* **2012**, *287*, 1527–1537.
69. Boyerinas, B.; Park, S.M.; Hau, A.; Murmann, A.E.; Peter, M.E. The role of let-7 in cell differentiation and cancer. *Endocr. Relat. Cancer* **2010**, *17*, F19–F36.
70. Johnson, S.M.; Grosshans, H.; Shingara, J.; Byrom, M.; Jarvis, R.; Cheng, A.; Labourier, E.; Reinert, K.L.; Brown, D.; Slack, F.J. RAS is regulated by the let-7 microRNA family. *Cell* **2005**, *120*, 635–647.
71. Hermeking, H. MicroRNAs in the p53 network: Micromanagement of tumour suppression. *Nat. Rev. Cancer* **2012**, *12*, 613–626.
72. Hunten, S.; Siemens, H.; Kaller, M.; Hermeking, H. The p53/microRNA network in cancer: Experimental and bioinformatics approaches. *Adv. Exp. Med. Biol*. **2013**, *774*, 77–101.
73. Yamamura, S.; Saini, S.; Majid, S.; Hirata, H.; Ueno, K.; Chang, I.; Tanaka, Y.; Gupta, A.; Dahiya, R. MicroRNA-34a suppresses malignant transformation by targeting c-Myc transcriptional complexes in human renal cell carcinoma. *Carcinogenesis* **2012**, *33*, 294–300.
74. Yamamura, S.; Saini, S.; Majid, S.; Hirata, H.; Ueno, K.; Deng, G.; Dahiya, R. MicroRNA-34a modulates c-Myc transcriptional complexes to suppress malignancy in human prostate cancer cells. *PLoS One* **2012**, *7*, e29722.
75. Cole, K.A.; Attiyeh, E.F.; Mosse, Y.P.; Laquaglia, M.J.; Diskin, S.J.; Brodeur, G.M.; Maris, J.M. A functional screen identifies miR-34a as a candidate neuroblastoma tumor suppressor gene. *Mol. Cancer Res*. **2008**, *6*, 735–742.

76. Hermeking, H. The miR-34 family in cancer and apoptosis. *Cell Death Differ.* **2010**, *17*, 193–199.
77. Wong, C.S.; Sceneay, J.; House, C.M.; Halse, H.M.; Liu, M.C.; George, J.; Hunnam, T.C.; Parker, B.S.; Haviv, I.; Ronai, Z.; *et al*. Vascular normalization by loss of Siah2 results in increased chemotherapeutic efficacy. *Cancer Res.* **2012**, *72*, 1694–1704.
78. Keith, B.; Johnson, R.S.; Simon, M.C. HIF1alpha and HIF2alpha: Sibling rivalry in hypoxic tumour growth and progression. *Nat. Rev. Cancer* **2012**, *12*, 9–22.
79. Yamakuchi, M.; Lotterman, C.D.; Bao, C.; Hruban, R.H.; Karim, B.; Mendell, J.T.; Huso, D.; Lowenstein, C.J. P53-induced microRNA-107 inhibits HIF-1 and tumor angiogenesis. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 6334–6339.
80. Sachdeva, M.; Zhu, S.; Wu, F.; Wu, H.; Walia, V.; Kumar, S.; Elble, R.; Watabe, K.; Mo, Y.Y. p53 represses c-Myc through induction of the tumor suppressor miR-145. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 3207–3212.
81. Chang, C.J.; Chao, C.H.; Xia, W.; Yang, J.Y.; Xiong, Y.; Li, C.W.; Yu, W.H.; Rehman, S.K.; Hsu, J.L.; Lee, H.H.; *et al*. p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. *Nat. Cell Biol.* **2011**, *13*, 317–323.
82. Kim, T.; Veronese, A.; Pichiorri, F.; Lee, T.J.; Jeon, Y.J.; Volinia, S.; Pineau, P.; Marchio, A.; Palatini, J.; Suh, S.S.; *et al*. p53 regulates epithelial-mesenchymal transition through microRNAs targeting ZEB1 and ZEB2. *J. Exp. Med.* **2011**, *208*, 875–883.
83. Yan, H.L.; Xue, G.; Mei, Q.; Wang, Y.Z.; Ding, F.X.; Liu, M.F.; Lu, M.H.; Tang, Y.; Yu, H.Y.; Sun, S.H. Repression of the miR-17–92 cluster by p53 has an important function in hypoxia-induced apoptosis. *EMBO J.* **2009**, *28*, 2719–2732.
84. Suzuki, H.I.; Yamagata, K.; Sugimoto, K.; Iwamoto, T.; Kato, S.; Miyazono, K. Modulation of microRNA processing by p53. *Nature* **2009**, *460*, 529–533.
85. Kohlstedt, K.; Trouvain, C.; Boettger, T.; Shi, L.; Fisslthaler, B.; Fleming, I. AMP-activated protein kinase regulates endothelial cell angiotensin-converting enzyme expression via p53 and the post-transcriptional regulation of microRNA-143/145. *Circ. Res*. **2013**, *112*, 1150–1158.
86. Wang, J.; He, Q.; Han, C.; Gu, H.; Jin, L.; Li, Q.; Mei, Y.; Wu, M. p53-facilitated miR-199a-3p regulates somatic cell reprogramming. *Stem Cells* **2012**, *30*, 1405–1413.
87. Manfe, V.; Biskup, E.; Rosbjerg, A.; Kamstrup, M.; Skov, A.G.; Lerche, C.M.; Lauenborg, B.T.; Odum, N.; Gniadecki, R. miR-122 regulates p53/Akt signalling and the chemotherapy-induced apoptosis in cutaneous T-cell lymphoma. *PLoS One* **2012**, *7*, e29541.
88. Michael, D.; Oren, M. The p53-Mdm2 module and the ubiquitin system. *Semin. Cancer Biol.* **2003**, *13*, 49–58.
89. Pichiorri, F.; Suh, S.S.; Rocci, A.; De Luca, L.; Taccioli, C.; Santhanam, R.; Zhou, W.; Benson, D.M., Jr.; Hofmainster, C.; Alder, H.; *et al.* Downregulation of p53-inducible microRNAs 192, 194, and 215 impairs the p53/MDM2 autoregulatory loop in multiple myeloma development. *Cancer Cell* **2010**, *18*, 367–381.
90. Khella, H.W.; Bakhet, M.; Allo, G.; Jewett, M.A.; Girgis, A.H.; Latif, A.; Girgis, H.; von Both, I.; Bjarnason, G.A.; Yousef, G.M. miR-192, miR-194 and miR-215: A convergent microRNA network suppressing tumor progression in renal cell carcinoma. *Carcinogenesis* **2013**, doi:10.1093/carcin/bgt184.

91. Senanayake, U.; Das, S.; Vesely, P.; Alzoughbi, W.; Frohlich, L.F.; Chowdhury, P.; Leuschner, I.; Hoefler, G.; Guertl, B. miR-192, miR-194, miR-215, miR-200c and miR-141 are downregulated and their common target ACVR2B is strongly expressed in renal childhood neoplasms. *Carcinogenesis* **2012**, *33*, 1014–1021.
92. Xiao, J.; Lin, H.; Luo, X.; Wang, Z. miR-605 joins p53 network to form a p53:miR-605:Mdm2 positive feedback loop in response to stress. *EMBO J*. **2011**, *30*, 524–532.
93. Pelengaris, S.; Khan, M.; Evan, G. c-MYC: More than just a matter of life and death. *Nat. Rev. Cancer* **2002**, *2*, 764–776.
94. Craig, V.J.; Cogliatti, S.B.; Imig, J.; Renner, C.; Neuenschwander, S.; Rehrauer, H.; Schlapbach, R.; Dirnhofer, S.; Tzankov, A.; Muller, A. Myc-mediated repression of microRNA-34a promotes high-grade transformation of B-cell lymphoma by dysregulation of FoxP1. *Blood* **2011**, *117*, 6227–6236.
95. Zhang, X.; Chen, X.; Lin, J.; Lwin, T.; Wright, G.; Moscinski, L.C.; Dalton, W.S.; Seto, E.; Wright, K.; Sotomayor, E.; *et al*. Myc represses miR-15a/miR-16–1 expression through recruitment of HDAC3 in mantle cell and other non-Hodgkin B-cell lymphomas. *Oncogene* **2012**, *31*, 3002–3008.
96. Bonci, D.; Coppola, V.; Musumeci, M.; Addario, A.; Giuffrida, R.; Memeo, L.; D'Urso, L.; Pagliuca, A.; Biffoni, M.; Labbaye, C.; *et al*. The miR-15a-miR-16-1 cluster controls prostate cancer by targeting multiple oncogenic activities. *Nat. Med*. **2008**, *14*, 1271–1277.
97. Zhang, X.; Zhao, X.; Fiskus, W.; Lin, J.; Lwin, T.; Rao, R.; Zhang, Y.; Chan, J.C.; Fu, K.; Marquez, V.E.; *et al*. Coordinated silencing of MYC-mediated miR-29 by HDAC3 and EZH2 as a therapeutic target of histone modification in aggressive B-Cell lymphomas. *Cancer Cell* **2012**, *22*, 506–523.
98. Schmitt, M.J.; Margue, C.; Behrmann, I.; Kreis, S. MiRNA-29: A microRNA family with tumor-suppressing and immune-modulating properties. *Curr. Mol. Med*. **2013**, *13*, 572–585.
99. Chin, L.J.; Ratner, E.; Leng, S.; Zhai, R.; Nallur, S.; Babar, I.; Muller, R.U.; Straka, E.; Su, L.; Burki, E.A.; *et al.* A SNP in a let-7 microRNA complementary site in the KRAS 3' untranslated region increases non-small cell lung cancer risk. *Cancer Res*. **2008**, *68*, 8535–8540.
100. Dang, C.V.; Le, A.; Gao, P. MYC-induced cancer cell energy metabolism and therapeutic opportunities. *Clin. Cancer Res*. **2009**, *15*, 6479–6483.
101. Jung, Y.J.; Kim, J.W.; Park, S.J.; Min, B.Y.; Jang, E.S.; Kim, N.Y.; Jeong, S.H.; Shin, C.M.; Lee, S.H.; Park, Y.S.; *et al*. c-Myc-mediated overexpression of miR-17-92 suppresses replication of hepatitis B virus in human hepatoma cells. *J. Med. Virol*. **2013**, *85*, 969–978.
102. Kim, J.W.; Mori, S.; Nevins, J.R. Myc-induced microRNAs integrate Myc-mediated cell proliferation and cell fate. *Cancer Res*. **2010**, *70*, 4820–4828.
103. Medina, R.; Zaidi, S.K.; Liu, C.G.; Stein, J.L.; van Wijnen, A.J.; Croce, C.M.; Stein, G.S. MicroRNAs 221 and 222 bypass quiescence and compromise cell survival. *Cancer Res*. **2008**, *68*, 2773–2780.
104. Visone, R.; Russo, L.; Pallante, P.; de Martino, I.; Ferraro, A.; Leone, V.; Borbone, E.; Petrocca, F.; Alder, H.; Croce, C.M.; *et al*. MicroRNAs (miR)-221 and miR-222, both overexpressed in human thyroid papillary carcinomas, regulate p27Kip1 protein levels and cell cycle. *Endocr. Relat. Cancer* **2007**, *14*, 791–798.

105. Fornari, F.; Gramantieri, L.; Ferracin, M.; Veronese, A.; Sabbioni, S.; Calin, G.A.; Grazi, G.L.; Giovannini, C.; Croce, C.M.; Bolondi, L.; *et al*. MiR-221 controls CDKN1C/p57 and CDKN1B/p27 expression in human hepatocellular carcinoma. *Oncogene* **2008**, *27*, 5651–5661.
106. Garofalo, M.; di Leva, G.; Romano, G.; Nuovo, G.; Suh, S.S.; Ngankeu, A.; Taccioli, C.; Pichiorri, F.; Alder, H.; Secchiero, P.; *et al.* miR-221&222 regulate TRAIL resistance and enhance tumorigenicity through PTEN and TIMP3 downregulation. *Cancer Cell* **2009**, *16*, 498–509.
107. Wang, X.; Zhao, X.; Gao, P.; Wu, M. c-Myc modulates microRNA processing via the transcriptional regulation of Drosha. *Sci. Rep.* **2013**, *3*, doi:10.1038/srep01942.
108. Goodrich, J.A.; Kugel, J.F. Non-coding-RNA regulators of RNA polymerase II transcription. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 612–616.
109. Spizzo, R.; Almeida, M.I.; Colombatti, A.; Calin, G.A. Long non-coding RNAs and cancer: A new frontier of translational research? *Oncogene* **2012**, *31*, 4577–4587.
110. Harrow, J.; Frankish, A.; Gonzalez, J.M.; Tapanari, E.; Diekhans, M.; Kokocinski, F.; Aken, B.L.; Barrell, D.; Zadissa, A.; Searle, S.; *et al.* GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res.* **2012**, *22*, 1760–1774.
111. Derrien, T.; Johnson, R.; Bussotti, G.; Tanzer, A.; Djebali, S.; Tilgner, H.; Guernec, G.; Martin, D.; Merkel, A.; Knowles, D.G.; *et al.* The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.* **2012**, *22*, 1775–1789.
112. Banfai, B.; Jia, H.; Khatun, J.; Wood, E.; Risk, B.; Gundling, W.E., Jr.; Kundaje, A.; Gunawardena, H.P.; Yu, Y.; Xie, L.; *et al.* Long noncoding RNAs are rarely translated in two human cell lines. *Genome Res.* **2012**, *22*, 1646–1657.
113. Brown, C.J.; Hendrich, B.D.; Rupert, J.L.; Lafreniere, R.G.; Xing, Y.; Lawrence, J.; Willard, H.F. The human XIST gene: Analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **1992**, *71*, 527–542.
114. Clemson, C.M.; McNeil, J.A.; Willard, H.F.; Lawrence, J.B. XIST RNA paints the inactive X chromosome at interphase: Evidence for a novel RNA involved in nuclear/chromosome structure. *J. Cell Biol.* **1996**, *132*, 259–275.
115. Zhao, J.; Sun, B.K.; Erwin, J.A.; Song, J.J.; Lee, J.T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **2008**, *322*, 750–756.
116. Tian, D.; Sun, S.; Lee, J.T. The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell* **2010**, *143*, 390–403.
117. Chureau, C.; Chantalat, S.; Romito, A.; Galvani, A.; Duret, L.; Avner, P.; Rougeulle, C. Ftx is a non-coding RNA which affects Xist expression and chromatin structure within the X-inactivation center region. *Hum. Mol. Genet.* **2011**, *20*, 705–718.
118. Lee, J.T.; Davidow, L.S.; Warshawsky, D. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat. Genet.* **1999**, *21*, 400–404.
119. Lee, J.T. Disruption of imprinted X inactivation by parent-of-origin effects at Tsix. *Cell* **2000**, *103*, 17–27.
120. Ogawa, Y.; Lee, J.T. Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol. Cell* **2003**, *11*, 731–743.
121. Khalil, A.M.; Guttman, M.; Huarte, M.; Garber, M.; Raj, A.; Rivea Morales, D.; Thomas, K.; Presser, A.; Bernstein, B.E.; van Oudenaarden, A.; *et al*. Many human large intergenic

noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11667–11672.

122. Rinn, J.L.; Kertesz, M.; Wang, J.K.; Squazzo, S.L.; Xu, X.; Brugmann, S.A.; Goodnough, L.H.; Helms, J.A.; Farnham, P.J.; Segal, E.; *et al*. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **2007**, *129*, 1311–1323.
123. Tsai, M.C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J.K.; Lan, F.; Shi, Y.; Segal, E.; Chang, H.Y. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **2010**, *329*, 689–693.
124. Gupta, R.A.; Shah, N.; Wang, K.C.; Kim, J.; Horlings, H.M.; Wong, D.J.; Tsai, M.C.; Hung, T.; Argani, P.; Rinn, J.L.; *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **2010**, *464*, 1071–1076.
125. Yang, Z.; Zhou, L.; Wu, L.M.; Lai, M.C.; Xie, H.Y.; Zhang, F.; Zheng, S.S. Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. *Ann. Surg. Oncol.* **2011**, *18*, 1243–1250.
126. Geng, Y.J.; Xie, S.L.; Li, Q.; Ma, J.; Wang, G.Y. Large intervening non-coding RNA HOTAIR is associated with hepatocellular carcinoma progression. *J. Int. Med. Res.* **2011**, *39*, 2119–2128.
127. Kogo, R.; Shimamura, T.; Mimori, K.; Kawahara, K.; Imoto, S.; Sudo, T.; Tanaka, F.; Shibata, K.; Suzuki, A.; Komune, S.; *et al*. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res*. **2011**, *71*, 6320–6326.
128. Kim, K.; Jutooru, I.; Chadalapaka, G.; Johnson, G.; Frank, J.; Burghardt, R.; Kim, S.; Safe, S. HOTAIR is a negative prognostic factor and exhibits pro-oncogenic activity in pancreatic cancer. *Oncogene* **2013**, *32*, 1616–1625.
129. Wang, K.C.; Yang, Y.W.; Liu, B.; Sanyal, A.; Corces-Zimmerman, R.; Chen, Y.; Lajoie, B.R.; Protacio, A.; Flynn, R.A.; Gupta, R.A.; *et al*. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **2011**, *472*, 120–124.
130. Zhang, X.; Lian, Z.; Padden, C.; Gerstein, M.B.; Rozowsky, J.; Snyder, M.; Gingeras, T.R.; Kapranov, P.; Weissman, S.M.; Newburger, P.E. A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. *Blood* **2009**, *113*, 2526–2534.
131. Ji, P.; Diederichs, S.; Wang, W.; Boing, S.; Metzger, R.; Schneider, P.M.; Tidow, N.; Brandt, B.; Buerger, H.; Bulk, E.; *et al*. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* **2003**, *22*, 8031–8041.
132. Hutchinson, J.N.; Ensminger, A.W.; Clemson, C.M.; Lynch, C.R.; Lawrence, J.B.; Chess, A. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **2007**, *8*, 39.
133. Tripathi, V.; Ellis, J.D.; Shen, Z.; Song, D.Y.; Pan, Q.; Watt, A.T.; Freier, S.M.; Bennett, C.F.; Sharma, A.; Bubulya, P.A.; *et al*. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* **2010**, *39*, 925–938.
134. Miyagawa, R.; Tano, K.; Mizuno, R.; Nakamura, Y.; Ijiri, K.; Rakwal, R.; Shibato, J.; Masuo, Y.; Mayeda, A.; Hirose, T.; *et al*. Identification of cis- and trans-acting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles. *RNA* **2012**, *18*, 738–751.

135. Gutschner, T.; Hammerle, M.; Eissmann, M.; Hsu, J.; Kim, Y.; Hung, G.; Revenko, A.; Arun, G.; Stentrup, M.; Gross, M.; *et al*. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res*. **2013**, *73*, 1180–1189.
136. Yang, L.; Lin, C.; Liu, W.; Zhang, J.; Ohgi, K.A.; Grinstein, J.D.; Dorrestein, P.C.; Rosenfeld, M.G. ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. *Cell* **2011**, *147*, 773–788.
137. Guo, F.; Li, Y.; Liu, Y.; Wang, J.; Li, G. Inhibition of metastasis-associated lung adenocarcinoma transcript 1 in CaSki human cervical cancer cells suppresses cell proliferation and invasion. *Acta Biochim. Biophys. Sin. (Shanghai)* **2010**, *42*, 224–229.
138. Yamada, K.; Kano, J.; Tsunoda, H.; Yoshikawa, H.; Okubo, C.; Ishiyama, T.; Noguchi, M. Phenotypic characterization of endometrial stromal sarcoma of the uterus. *Cancer Sci*. **2006**, *97*, 106–112.
139. Wilusz, J.E.; Freier, S.M.; Spector, D.L. 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell* **2008**, *135*, 919–932.
140. Xu, C.; Yang, M.; Tian, J.; Wang, X.; Li, Z. MALAT-1: A long non-coding RNA and its important 3' end functional motif in colorectal cancer metastasis. *Int. J. Oncol*. **2011**, *39*, 169–175.
141. Wilusz, J.E.; JnBaptiste, C.K.; Lu, L.Y.; Kuhn, C.D.; Joshua-Tor, L.; Sharp, P.A. A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.* **2012**, *26*, 2392–2407.
142. Brown, J.A.; Valenstein, M.L.; Yario, T.A.; Tycowski, K.T.; Steitz, J.A. Formation of triple-helical structures by the 3'-end sequences of MALAT1 and MENbeta noncoding RNAs. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 19202–19207.
143. Pasmant, E.; Laurendeau, I.; Heron, D.; Vidaud, M.; Vidaud, D.; Bieche, I. Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: Identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res*. **2007**, *67*, 3963–3969.
144. Yap, K.L.; Li, S.; Munoz-Cabello, A.M.; Raguz, S.; Zeng, L.; Mujtaba, S.; Gil, J.; Walsh, M.J.; Zhou, M.M. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* **2010**, *38*, 662–674.
145. Mourtada-Maarabouni, M.; Hedge, V.L.; Kirkham, L.; Farzaneh, F.; Williams, G.T. Growth arrest in human T-cells is controlled by the non-coding RNA growth-arrest-specific transcript 5 (GAS5). *J. Cell Sci.* **2008**, *121*, 939–946.
146. Kino, T.; Hurt, D.E.; Ichijo, T.; Nader, N.; Chrousos, G.P. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci. Signal* **2010**, *3*, doi:10.1126/scisignal.2000568.
147. Conti, E.; Izaurralde, E. Nonsense-mediated mRNA decay: Molecular insights and mechanistic variations across species. *Curr. Opin. Cell Biol.* **2005**, *17*, 316–325.
148. Zhang, Z.; Zhu, Z.; Watabe, K.; Zhang, X.; Bai, C.; Xu, M.; Wu, F.; Mo, Y.Y. Negative regulation of lncRNA GAS5 by miR-21. *Cell Death Differ.* **2013**, doi:10.1038/cdd.2013.110.
149. Pan, X.; Wang, Z.X.; Wang, R. MicroRNA-21: A novel therapeutic target in human cancer. *Cancer Biol. Ther.* **2010**, *10*, 1224–1232.

150. Smedley, D.; Sidhar, S.; Birdsall, S.; Bennett, D.; Herlyn, M.; Cooper, C.; Shipley, J. Characterization of chromosome 1 abnormalities in malignant melanomas. *Genes Chromosomes Cancer* **2000**, *28*, 121–125.
151. Nupponen, N.N.; Carpten, J.D. Prostate cancer susceptibility genes: Many studies, many results, no answers. *Cancer Metastasis Rev*. **2001**, *20*, 155–164.
152. Stange, D.E.; Radlwimmer, B.; Schubert, F.; Traub, F.; Pich, A.; Toedt, G.; Mendrzyk, F.; Lehmann, U.; Eils, R.; Kreipe, H.; *et al*. High-resolution genomic profiling reveals association of chromosomal aberrations on 1q and 16p with histologic and genetic subgroups of invasive breast cancer. *Clin. Cancer Res*. **2006**, *12*, 345–352.
153. Morrison, L.E.; Jewell, S.S.; Usha, L.; Blondin, B.A.; Rao, R.D.; Tabesh, B.; Kemper, M.; Batus, M.; Coon, J.S. Effects of ERBB2 amplicon size and genomic alterations of chromosomes 1, 3, and 10 on patient response to trastuzumab in metastatic breast cancer. *Genes Chromosomes Cancer* **2007**, *46*, 397–405.
154. Peters, U.; Jiao, S.; Schumacher, F.R.; Hutter, C.M.; Aragaki, A.K.; Baron, J.A.; Berndt, S.I.; Bezieau, S.; Brenner, H.; Butterbach, K.; *et al.* Identification of genetic susceptibility loci for colorectal tumors in a genome-wide meta-analysis. *Gastroenterology* **2013**, *144*, 799–807.
155. Nakamura, Y.; Takahashi, N.; Kakegawa, E.; Yoshida, K.; Ito, Y.; Kayano, H.; Niitsu, N.; Jinnai, I.; Bessho, M. The GAS5 (growth arrest-specific transcript 5) gene fuses to BCL6 as a result of t(1;3)(q25;q27) in a patient with B-cell lymphoma. *Cancer Genet Cytogenet* **2008**, *182*, 144–149.
156. Tani, H.; Torimura, M.; Akimitsu, N. The RNA degradation pathway regulates the function of GAS5 a non-coding RNA in mammalian cells. *PLoS One* **2013**, *8*, e55684.
157. Smith, C.M.; Steitz, J.A. Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. *Mol. Cell. Biol.* **1998**, *18*, 6897–6909.
158. Tanaka, R.; Satoh, H.; Moriyama, M.; Satoh, K.; Morishita, Y.; Yoshida, S.; Watanabe, T.; Nakamura, Y.; Mori, S. Intronic U50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t(3;6)(q27;q15) of human B-cell lymphoma. *Genes Cells* **2000**, *5*, 277–287.
159. Raho, G.; Barone, V.; Rossi, D.; Philipson, L.; Sorrentino, V. The gas 5 gene shows four alternative splicing patterns without coding for a protein. *Gene* **2000**, *256*, 13–17.
160. Hirose, T.; Steitz, J.A. Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12914–12919.
161. Johnsson, P.; Ackley, A.; Vidarsdottir, L.; Lui, W.O.; Corcoran, M.; Grander, D.; Morris, K.V. A pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human cells. *Nat. Struct. Mol. Biol.* **2013**, *20*, 440–446.
162. Bond, A.M.; Vangompel, M.J.; Sametsky, E.A.; Clark, M.F.; Savage, J.C.; Disterhoft, J.F.; Kohtz, J.D. Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry. *Nat. Neurosci*. **2009**, *12*, 1020–1027.
163. Huarte, M.; Guttman, M.; Feldser, D.; Garber, M.; Koziol, M.J.; Kenzelmann-Broz, D.; Khalil, A.M.; Zuk, O.; Amit, I.; Rabani, M.; *et al*. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **2010**, *142*, 409–419.

164. Sheik Mohamed, J.; Gaughwin, P.M.; Lim, B.; Robson, P.; Lipovich, L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *RNA* **2010**, *16*, 324–337.
165. Ebert, M.S.; Neilson, J.R.; Sharp, P.A. MicroRNA sponges: Competitive inhibitors of small RNAs in mammalian cells. *Nat. Methods* **2007**, *4*, 721–726.
166. Pesole, G.; Liuni, S.; Grillo, G.; Saccone, C. Structural and compositional features of untranslated regions of eukaryotic mRNAs. *Gene* **1997**, *205*, 95–102.
167. Huse, J.T.; Brennan, C.; Hambardzumyan, D.; Wee, B.; Pena, J.; Rouhanifard, S.H.; Sohn-Lee, C.; le Sage, C.; Agami, R.; Tuschl, T.; *et al*. The PTEN-regulating microRNA miR-26a is amplified in high-grade glioma and facilitates gliomagenesis *in vivo*. *Genes Dev.* **2009**, *23*, 1327–1337.
168. Liu, Z.L.; Wang, H.; Liu, J.; Wang, Z.X. MicroRNA-21 (miR-21) expression promotes growth, metastasis, and chemo- or radioresistance in non-small cell lung cancer cells by targeting PTEN. *Mol. Cell. Biochem.* **2013**, *372*, 35–45.
169. Vanin, E.F. Processed pseudogenes: Characteristics and evolution. *Annu. Rev. Genet.* **1985**, *19*, 253–272.
170. Fujii, G.H.; Morimoto, A.M.; Berson, A.E.; Bolen, J.B. Transcriptional analysis of the PTEN/MMAC1 pseudogene, psiPTEN. *Oncogene* **1999**, *18*, 1765–1769.
171. Poliseno, L.; Salmena, L.; Zhang, J.; Carver, B.; Haveman, W.J.; Pandolfi, P.P. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **2010**, *465*, 1033–1038.
172. Lund, A.H.; van Lohuizen, M. Epigenetics and cancer. *Genes Dev.* **2004**, *18*, 2315–2335.
173. Karreth, F.A.; Tay, Y.; Perna, D.; Ala, U.; Tan, S.M.; Rust, A.G.; DeNicola, G.; Webster, K.A.; Weiss, D.; Perez-Mancera, P.A.; *et al. In vivo* identification of tumor- suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. *Cell* **2011**, *147*, 382–395.
174. Faghihi, M.A.; Modarresi, F.; Khalil, A.M.; Wood, D.E.; Sahagan, B.G.; Morgan, T.E.; Finch, C.E.; St Laurent, G., 3rd.; Kenny, P.J.; Wahlestedt, C. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat. Med*. **2008**, *14*, 723–730.
175. Cesana, M.; Cacchiarelli, D.; Legnini, I.; Santini, T.; Sthandier, O.; Chinappi, M.; Tramontano, A.; Bozzoni, I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **2011**, *147*, 358–369.
176. Nakagawa, S.; Ip, J.Y.; Shioi, G.; Tripathi, V.; Zong, X.; Hirose, T.; Prasanth, K.V. Malat1 is not an essential component of nuclear speckles in mice. *RNA* **2012**, *18*, 1487–1499.
177. Zhang, A.; Zhou, N.; Huang, J.; Liu, Q.; Fukuda, K.; Ma, D.; Lu, Z.; Bai, C.; Watabe, K.; Mo, Y.Y. The human long non-coding RNA-RoR is a p53 repressor in response to DNA damage. *Cell Res.* **2013**, *23*, 340–350.
178. Yoon, J.H.; Abdelmohsen, K.; Srikantan, S.; Yang, X.; Martindale, J.L.; De, S.; Huarte, M.; Zhan, M.; Becker, K.G.; Gorospe, M. LincRNA-p21 suppresses target mRNA translation. *Mol. Cell* **2012**, *47*, 648–655.
179. Azzalin, C.M.; Reichenbach, P.; Khoriauli, L.; Giulotto, E.; Lingner, J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **2007**, *318*, 798–801.

180. Redon, S.; Reichenbach, P.; Lingner, J. The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. *Nucleic Acids Res.* **2010**, *38*, 5797–5806.
181. Wang, X.; Arai, S.; Song, X.; Reichart, D.; Du, K.; Pascual, G.; Tempst, P.; Rosenfeld, M.G.; Glass, C.K.; Kurokawa, R. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* **2008**, *454*, 126–130.
182. Jauliac, S.; Lopez-Rodriguez, C.; Shaw, L.M.; Brown, L.F.; Rao, A.; Toker, A. The role of NFAT transcription factors in integrin-mediated carcinoma invasion. *Nat. Cell Biol.* **2002**, *4*, 540–544.
183. Yoeli-Lerner, M.; Yiu, G.K.; Rabinovitz, I.; Erhardt, P.; Jauliac, S.; Toker, A. Akt blocks breast cancer cell motility and invasion through the transcription factor NFAT. *Mol. Cell* **2005**, *20*, 539–550.
184. Willingham, A.T.; Orth, A.P.; Batalov, S.; Peters, E.C.; Wen, B.G.; Aza-Blanc, P.; Hogenesch, J.B.; Schultz, P.G. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **2005**, *309*, 1570–1573.
185. Sharma, S.; Findlay, G.M.; Bandukwala, H.S.; Oberdoerffer, S.; Baust, B.; Li, Z.; Schmidt, V.; Hogan, P.G.; Sacks, D.B.; Rao, A. Dephosphorylation of the nuclear factor of activated T cells (NFAT) transcription factor is regulated by an RNA-protein scaffold complex. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11381–11386.
186. Sanger, H.L.; Klotz, G.; Riesner, D.; Gross, H.J.; Kleinschmidt, A.K. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 3852–3856.
187. Jeck, W.R.; Sorrentino, J.A.; Wang, K.; Slevin, M.K.; Burd, C.E.; Liu, J.; Marzluff, W.F.; Sharpless, N.E. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **2013**, *19*, 141–157.
188. Salzman, J.; Gawad, C.; Wang, P.L.; Lacayo, N.; Brown, P.O. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* **2012**, *7*, e30733.
189. Capel, B.; Swain, A.; Nicolis, S.; Hacker, A.; Walter, M.; Koopman, P.; Goodfellow, P.; Lovell-Badge, R. Circular transcripts of the testis-determining gene Sry in adult mouse testis. *Cell* **1993**, *73*, 1019–1030.
190. Hansen, T.B.; Jensen, T.I.; Clausen, B.H.; Bramsen, J.B.; Finsen, B.; Damgaard, C.K.; Kjems, J. Natural RNA circles function as efficient microRNA sponges. *Nature* **2013**, *495*, 384–388.
191. Hansen, T.B.; Wiklund, E.D.; Bramsen, J.B.; Villadsen, S.B.; Statham, A.L.; Clark, S.J.; Kjems, J. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J.* **2011**, *30*, 4414–4422.

Reprinted from *IJMS*. Cite as: Lung, R.W.-M.; Tong, J.H.-M.; To, K.-F. Emerging Roles of Small Epstein-Barr Virus Derived Non-Coding RNAs in Epithelial Malignancy. *Int. J. Mol. Sci.* **2013**, *14*, 17378-17409.

Review

Emerging Roles of Small Epstein-Barr Virus Derived Non-Coding RNAs in Epithelial Malignancy

Raymond Wai-Ming Lung [1], Joanna Hung-Man Tong [1] and Ka-Fai To [1,2,*]

[1] Department of Anatomical and Cellular Pathology, State Key Laboratory in Oncology in South China, The Chinese University of Hong Kong, Hong Kong, China;
E-Mails: raymond_lung@cuhk.edu.hk (R.W.-M.L.); jtong@cuhk.edu.hk (J.H.-M.T.)

[2] Institute of Digestive Diseases, Li Ka-Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong, China

* Author to whom correspondence should be addressed; E-Mail: kfto@cuhk.edu.hk; Tel.: +852-2632-3334; Fax: +852-2637-6274.

Received: 8 July 2013; in revised form: 1 August 2013 / Accepted: 13 August 2013 / Published: 23 August 2013

Abstract: Latent Epstein-Barr virus (EBV) infection is an etiological factor in the progression of several human epithelial malignancies such as nasopharyngeal carcinoma (NPC) and a subset of gastric carcinoma. Reports have shown that EBV produces several viral oncoproteins, yet their pathological roles in carcinogenesis are not fully elucidated. Studies on the recently discovered of EBV-encoded microRNAs (ebv-miRNAs) showed that these small molecules function as post-transcriptional gene regulators and may play a role in the carcinogenesis process. In NPC and EBV positive gastric carcinoma (EBVaGC), 22 viral miRNAs which are located in the long alternative splicing EBV transcripts, named BamH1 A rightward transcripts (BARTs), are abundantly expressed. The importance of several miR-BARTs in carcinogenesis has recently been demonstrated. These novel findings enhance our understanding of the oncogenic properties of EBV and may lead to a more effective design of therapeutic regimens to combat EBV-associated malignancies. This article will review the pathological roles of miR-BARTs in modulating the expression of cancer-related genes in both host and viral genomes. The expression of other small non-coding RNAs in NPC and the expression pattern of miR-BARTs in rare EBV-associated epithelial cancers will also be discussed.

Keywords: Epstein-Barr virus (EBV); nasopharyngeal carcinoma (NPC); EBV-associated gastric carcinoma (EBVaGC); lymphoepithelioma-like carcinoma (LELC) of the lung; lymphoepithelioma-like cholangiocarcinoma; BART; miRNAs; v-snoRNA1

1. Introduction

The Epstein-Barr virus (EBV) is a ubiquitous gamma herpesvirus that was originally identified in Burkitt's lymphoma (BL) by Tony Epstein and his colleagues in 1964 [1]. It is also the first virus known to play an essential role in the induction of human malignancies [2]. Interestingly, EBV infection is common among most population groups worldwide. However, most infected individuals are asymptomatic, and for those who show signs and symptoms of acute EBV infection tend to recover without sequelae [3]. Interestingly, the virus will not be entirely eliminated and will ultimately establish life-long latent infection within memory B cell pool, in which the viral genome circulates as episome and undergoes replication using host cellular expression machinery [4,5]. Latent EBV infection is associated with a wide range of human lymphoid and epithelial malignancies, including Hodgkin disease, Burkitt's lymphoma (BL), nasopharyngeal carcinoma (NPC), and a subset of gastric carcinoma (GC).

While the EBV genome contains around 80 open reading frames (*ORF*s), only a few of viral proteins are selectively expressed during latent infection. Expression of these viral proteins forms four latency patterns [6]. Latency III infection is commonly observed in EBV-associated B-cell lymphoma in the setting of immunosuppression, in which all viral latent genes, including six EBV encoded nuclear antigen proteins (EBNA-1, EBNA-2, EBNA-3A, 3B, 3C and EBNA-LP) and three latent membrane proteins (LMP1, LMP2A and LMP2B), are expressed [7]. The spectrum of lymphomas in latency III includes post-transplant lymphoproliferative diseases (PTLD) and immunodeficiency associated lymphoproliferative diseases. It is likely that the immunosuppressive environment allows the expression of all latent proteins in the absence of host immune responses [8]. EBV protein expression in latency II is limited to EBNA-1 and the three LMPs. This latency type is commonly observed in NPC [9]. EBNA-1, an essential protein for virus persistence, is the only viral protein expressed in latency I. In a healthy carrier, EBV remains dormant in memory B-cells which circulate in the peripheral blood (latency 0). At this stage, expression of the EBNA-1 viral protein is found only during cell division, albeit the LMP transcript is detected at very low level in the peripheral blood [5,10]. Interestingly, two types of non protein-coding RNA, *EBV-encoded RNAs* (*EBER-1* and *EBER-2*) and long alternative splicing non-coding RNA at BamH1 A rightward transcripts (BARTs), are expressed in all forms of latency programs [11,12].

2. EBV Infection in Epithelial Carcinoma

2.1. Nasopharyngeal Carcinoma (NPC)

Nasopharyngeal carcinoma (NPC) is a distinctive type of head and neck cancer arising from the nasopharynx. This is a rare tumor in most part of the world (below 1 per 100,000 persons per year) [13]. However, the incidence and mortality rates of NPC are remarkably high in Southern China including

Hong Kong (about 19.5 and 7.7 per 100,000 persons, respectively) [14]. The World Health Organization (WHO) has classified NPC into three histological categories: type I is keratinizing squamous cell carcinoma (SCC), type II is non-keratinizing carcinoma, and type III is undifferentiated carcinoma. The prevalence of individual type of NPC is dependent on geographic regions. Types I and II NPC are generally found in western population. And type III, which is consistently associated with EBV latent infection, accounts the majority of NPC in Southern China [15]. Because of the unusual prominent lymphocytic infiltration, type III NPC is also commonly described as lymphoepithelioma of nasopharynx. It has been suggested that EBV particles from the infected lymphoid cells are transferred to the nasopharynx cells via cell-cell contact [16,17]. Clonal EBV genome is consistently detected in both high-grade dysplastic lesions of the nasophargnx and invasive NPC, suggesting that EBV may contribute to the neoplastic transformation of nasopharyngeal epithelial cells and facilitate the clonal expansion of malignant cells [13,18].

Although the crucial roles of viral LMPs in NPC pathogenesis have previously been reported [19,20], their expressions in NPC are variable and generally low when compared to those in lymphoid malignancies. In NPC, viral proteins are processed and presented on the cell surface, resulting in the stimulation of host immune surveillance [21,22]. Therefore, low expression of viral proteins is critical for maintenance viral latency. In this regard, EBV may contribute to cancer development by two abundantly expressed non protein-coding viral products, *EBERs* and *BARTs*. Currently, the abundant expressions of 22 *BARTs*-derived microRNAs (miR-BARTs) in NPC have been identified; it is possible that EBV augment cancer development through these viral-encoded miRNAs [23,24].

2.2. EBV-Associated Gastric Carcinoma (EBVaGC)

Gastric carcinoma (GC) is another epithelial cancer reported to be associated with EBV infection. In 1990, Burke *et al.* first reported the presence of EBV DNA in a rare type of gastric carcinoma, namely the lymphoepithelioma-like carcinomas (LELCs) of stomach [25]. Subsequently, Shibata's research team and others demonstrated that EBV is present in more than 80% of gastric LELCs [26,27]. In 1992, Shibta's team further identified the presence of EBV in up to 16% of conventional gastric adenocarcinoma [28]. They collectively described this subtype of gastric cancer as EBV-associated gastric carcinoma (EBVaGC). Unlike NPC, EBVaGC is not an endemic disease and is present in about 10% of all gastric cancer worldwide. The incidence rates are highly similar among North America (9.9%), Europe (9.2%) and Asia (8.3%) [27,29–32]. EBVaGC represents a distinct type of GC as it has special clinicopathological characteristics such as male predominance, predisposition to the proximal stomach, lace pattern in the early GC and more frequent occurrence in gastric remnant carcinoma (GRC). EBVaGC has been usually described as latency I infection although LMP2A was weakly detected in some of the cases [29,33,34]. Previous works performed by our team also demonstrated the importance of LMP2A in gastric cancer development through genome-wide methylation analyses [35–37]. In contrast to the low expression level of viral protein, abundant miR-BARTs expression in the EBVaGC may be implicated in carcinogenesis [38,39].

2.3. Lymphoepithelioma-Like Carcinomas (LELCs)

LELC is histologically defined, a poorly differentiated carcinoma with dense lymphocytic infiltration. This type of cancer can be found not only in type III NPC, but also in other anatomical sites. Interestingly, EBV has been consistently detected in the LELC derived from foregut organs such as salivary gland, lung, thymus, gastric and intra-hepatic biliary epithelium (cholangiocarcinoma). They are rare cancers mainly reported in endemic regions [40–43].

LELC of the lung was first reported in 1987, and fewer than 200 cases have been reported in the literature [44]. This rare tumor occupies less than 1% of the overall lung cancer incidence in southeast China and Japan [45]. Nevertheless, EBER is consistently detected in almost all LELCs of lung in Asia. Latency program of EBV in LELC of the lung is unclear as only ~50% of the cases showed the expression of LMP1 viral oncoprotein. EBV positive carcinomas have been rarely developed in liver, some cases of these carcinomas reported as LELCs but the majority of them are EBV positive cholangiocarcinoma (CCA) [41]. In this review, we will address the expression pattern of EBV miRNAs in LELC of the lung and CCA and will discuss the possible pathogenic roles of EBV in LELC.

3. EBV Non Protein-Coding RNAs

3.1. Epstein-Barr Virus-Encoded RNAs (EBERs)

The EBERs contain two species, EBER1 and EBER2 with sizes of 166 and 172 nucleotides, respectively [46]. These two genes are separated by 161 nucleotides in the EBV genome and transcribed individually in the same direction using RNA polymerase III [46–48]. They are known to be abundantly expressed in all latent EBV infected cells with a copy number of up to 5×10^6 per infected cell [46,49]. Faint EBER signals in the cytoplasm can be detected using high-resolution fluorescent *in situ* hybridization (ISH) technique with confocal laser scanning microscopy [50]. However, these transcripts are dominantly and stably confined to the nucleus, where they are associated with several ribonucleoproteins [46,51], including lupus antigen (LA) [46], ribosomal protein L22 (formerly EBER-associated protein EAR) [52–54], interferon-inducible, double-stranded RNA activated protein kinase R (PKR) [55] and retinoic acid-inducible gene I (RIG-I) [56]. As a result, EBER transcripts can be detected in all EBV positive cells by ISH and are routinely used as a probe to determine EBV infection in clinical samples. The functional role of EBERs in epithelial cancer development is not clear. Despite the demonstration of oncogenic properties of EBERs on B-cells in *in vitro* and *in vivo* experiments, EBERs can neither induce cellular transformation nor increase tumorigenicity in epithelial cell models. However, several reports indicated that EBERs can confer an apoptotic-resistant phenotype in immortalized epithelial cells and can support cell growth by stimulating insulin-like growth factor (IGF-1) secretion in EBV-positive gastric carcinoma and NPC cell lines [57–59] (for more details about EBER1 function, see [47,51]).

Early work from Steitz's research team has shown that the predicted folding structure of EBERs is comprised of several stem-loops which are highly similar to two non-coding adenovirus-associated (VA) RNAs, namely VAI and VAII [60,61]. Moreover, EBERs can completely substitute the critical role of VA RNAs during adenovirus replication [62,63]. Intriguingly, two research groups recently demonstrated that VA RNAs can be further processed into a small functional interference RNA [64,65].

Our preliminary work also identified EBER1 derived small RNA fragments (18–25 nt) from NPC small RNA library cloning [23]. The expressions of these fragments in NPC have been verified by Northern blot analysis (Figure S1A). Nonetheless, EBER1 is unlikely to undergo typical miRNA processing due to the following reasons. Firstly, the predicted stem-loop structure on EBER1 is too small for pre-miRNA processing by Drosha/DGCR8 complex [66]. Secondly, Dicer is unable to cleave EBER1 in an *in vitro* experiment [65]. Thirdly, EBER1 is mainly confined in the nucleus and does not associate with exportin-5, a protein that facilitates nucleus export of pre-mature miRNAs [67]. We further demonstrated that EBER1 and its derived small fragments do not bind to Argonaute2 (Ago2), an integral part in the functional RNA-induced silencing complex (RISC) complex (Figure S1B).

3.2. Bam HI A Rightward Transcripts (BARTs)

BARTs are multi-spliced transcripts originally discovered by cDNA library analysis in C15 NPC xenograft, and their expressions have later been reported in other EBV-infected individuals [68,69]. The contribution of BARTs is dispensable in B-cell transformation since mutant recombinant EBV carrying a null BART region maintains its transforming activity in primary B lymphocyte [70]. However, it has long been speculated that BARTs have important functional roles in EBV-associated epithelial malignancies as it is abundantly expressed in NPC and EBVaGC. In the early study of BARTs, several small ORFs in the spliced cDNA transcripts were the focuses of investigation. RPMS1 and A73 were revealed to be translated into oncogenic proteins in an *in vitro* study and were shown to be negative regulators of Notch and RACK1 signaling pathways, respectively. However, expressions of these two viral proteins have not been demonstrated in natural EBV-infected samples [71,72]. Increasing interest in the functional role of BARTs, has eventually led to the discovery of 22 viral miRNAs (miR-BARTs) and a small nucleolar RNA (snoRNA) in this region [23,73–76]. The details of miR-BARTs' function will be discussed later in this review.

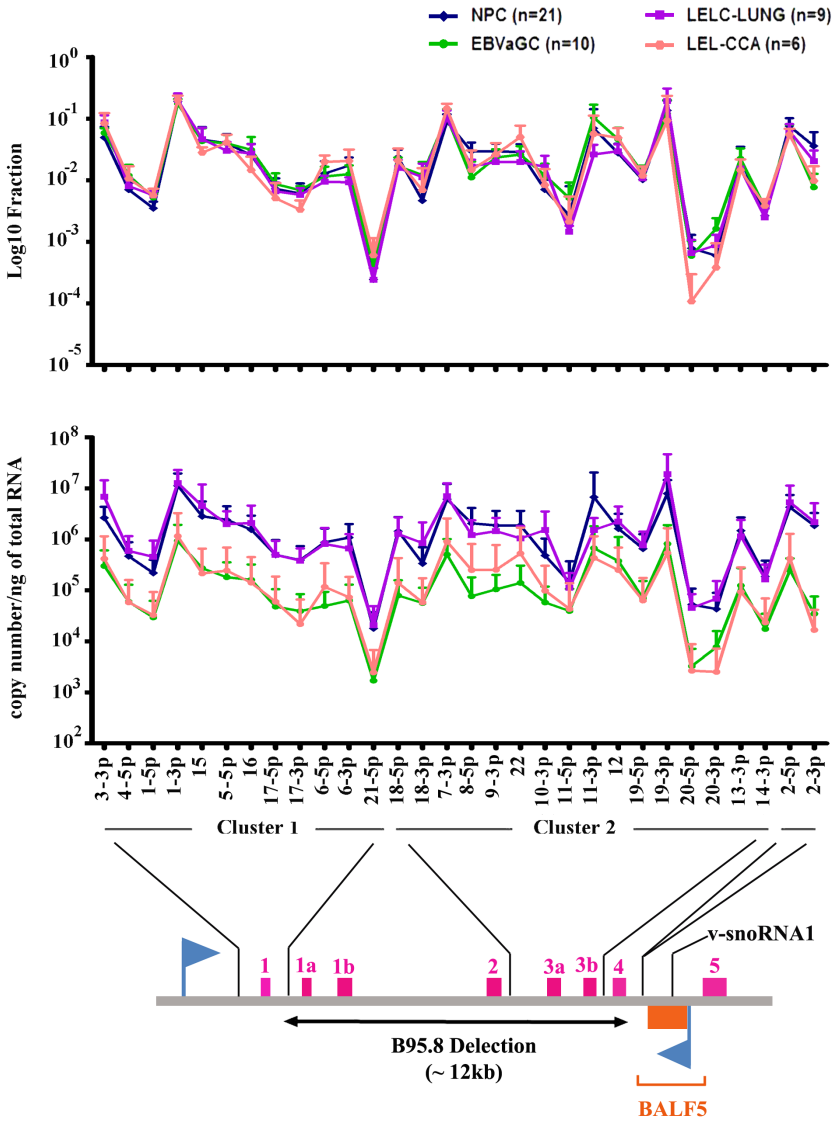
3.3. Discovery of BART-Encoded miRNAs

MicroRNAs (miRNAs), a group of 18–24 nucleotide non protein-coding RNAs, were produced by two endogenous enzymatic digestions (Drosha/DGCR8 complex and Dicer) from the hairpin structures of long RNA transcripts. A mature miRNA recruits a group of cellular proteins, including Argonaute2 (Ago2), to form a stable complex called RNA-induced silencing complex (RISC) [77], in which miRNA acts as a guide strand to negatively regulate gene expression through imperfect complementary sequence pairing to the 3' untranslated region (3'UTR) of the target gene [78]. One mature miRNA is expected to regulate expression of over 100 cellular genes by either repressing protein translation or inducing mRNA degradation [78]. It is estimated that about 30% of cellular protein-coding genes are regulated by miRNA [79].

EBV was the first virus identified to encode miRNA. In 2004, Thomas Tuschl and colleagues used molecular cloning methods to identify a total of 5 miRNAs (ebv-miR-BHRF1-1, -2, -3, ebv-miR-BART1 and 2) in human B-cells infected with laboratory isolated EBV strain B95.8, which carries a 12-kb deletion in the BART region of the EBV genome [73] (Figure 1). Eventually, subsequent work on other latently infected cell lines and biopsy materials identified a total of 25 EBV encoded miRNAs [23,24,74,75,80]. Among them, three miRNAs arising from the UTR of the distinct

Bcl2 homolog gene *BHRF1* (ebv-miR-BHRF1-1, -2 and -3) are predicted to be expressed in the EBV latency III program. The rest of the remaining 22 miRNAs (miR-BART1 to miR-BART22), which are located mainly in two clusters within the non-coding BART region, are generally highly expressed in most of the EBV infected epithelial cells (latency I and II) [20,39].

Figure 1. Expression of miR-BARTs in EBV positive epithelial carcinoma. The fraction of individual miRNA from all tested miR-BARTs is plotted in the upper panel. The middle panel shows the copy number of each miR-BART per nanogram of total RNA. The value represents mean + SD is plotted.



3.4. Discovery of Viral-Encoded snoRNA

Nucleoli contain a large population of conserved small stable RNAs known as small nucleolar RNA (snoRNA). They can mediate nucleotides modification of ribosomal RNAs and facilitate their production by forming a functional snoRNA:protein complex (snoRNPs) with a group of snoRNA core proteins [81]. Based on deep sequencing analysis of cDNA clones from B lymphocytes, Hutzinger *et al.* identified a snoRNA gene within the EBV genome. The viral snoRNA, mapped ~100 bp

downstream to the miR-BART2 transcripts, was 65 nucleotides (nt) in length and designated as v-snoRNA1. The authors also confirmed the expression of v-snoRNA1 in a panel of EBV positive BL and LCL cell lines. Given that v-snoRNA1 could associate with three canonical snoRNA proteins (fibrilarin, Nop56 and Nop58), it follows that v-snoRNA1 could be assembled into functional snoRNPs. On the other hand, a 24 nt RNA fragment from the v-snoRNA1 3' terminus (v-snoRNA^{24pp}) had been identified in cDNA library screening, but its expression was only detected during lytic induction on EBV infected 293 cells [76]. In view of the fact that some snoRNA species were reported as precursors for miRNA biogenesis [82,83], v-snoRNA1 might also be further processed into miRNA-like molecule, v-snoRNA^{24pp}. However, the experimental evidence to prove this hypothesis is not yet provided.

The miRNAs possess properties to modify the intracellular environment, a characteristic that is particularly useful and convenient for viral infection. Viral miRNAs were processed from short RNA transcripts (~200 nt) by using cellular transcriptional machinery. Thus, a number of miRNA precursors can be tightly packed into the relatively small viral genome. In addition, miRNAs are non-protein coding molecules that are presumably immunogenically inert, yet they can modulate several target genes that can enhance infected cell survival.

4. Expression of Viral miRNAs in Epithelial Cancers

4.1. Ebv-miR-BHRF1s

All three ebv-miR-BHRF1s are suggested to be derived from two latency III specific Cp/Wp-initiated transcripts, the primary EBNA transcript and the latent BHRF1 transcript. For this reason, BHRF1 miRNAs are mainly expressed in this specific latency program [38,73,74,84–86]. However, several research teams using QRT-PCR technique indicated the presence of BHRF1 miRNAs in latency type II EBV positive epithelial cancer cell lines [87,88]. The inconsistent observation may be the result of active EBV replication in a small portion of infected cells. During EBV replication, miR-BHRF1-2 and -3 may be generated from the lytic form of the BHRF1 transcripts that were produced from the alternative lytic promoter BHRF1p [86]. As a gold standard, Northern blot analysis is used for the detection of miRNA expression. Surprisingly, we could detect the expression of the lytic form of the BHRF1 transcript in NPC samples including cell lines, xenografts and half of the actual patient samples. However, no expression of BHRF1 miRNAs was detected in any of the tested samples (Figure S2A). In contrast to previous studies that reported that induction of lytic EBV replication could facilitate miR-BHRF1s expression in B-cells [74,84], induction of EBV lytic cycle in the native EBV positive NPC cell lines C666-1 did not activate miR-BHRF1 expression at detectable levels (Figure S2B). This finding aligned with the previous report that BHRF1 transcription does not correlate with the expression of mature miR-BHRF1s [86]. Failure of miR-BHRF1 biogenesis may be the result of the absence of their target transcript in NPCs, leading to the sharp degradation of the miRNAs [89]. In addition, while the level of miR-BHRF1s was negligible, miR-BARTs were detected from the biopsies of EBV positive lymphoepithelioma-like cholangiocarcinoma and LELC of the lung (Figure S2C). This is in line with the previous data that BHRF1-derived miRNAs are not processed in epithelial cancers when EBV infection is mainly in latency I/II programs [24,38,80,85].

4.2. Ebv-miR-BARTs

Viral BARTs are initiated from two constitutively active promoters P1 and P2 [86,90]. And miR-BARTs are the only known functional molecules produced from these transcripts. By using RT-PCR and small RNA library sequencing methods, expression profiles of EBV miRNAs in cancer have been extensively studied, albeit most of the studies focused on NPC [24,39,85–88,91,92]. It is obvious in NPC that miR-BARTs are abundantly expressed and constitute up to 23% of the total miRNAs [24]. However the expression of individual miR-BART in the cells is considerably different. The comprehensive miRNA profiling study by Qiu *et al.* revealed that several miR-BARTs that were not detected in normal infected cells (germinal center B cells: latency II and memory B cells: latency 0/I) were upregulated in various EBV-associated malignancies such as BL, EBVaGC, Hodgkin Disease and NPC. These critical findings designated the importance of miR-BARTs in the development of EBV-associated cancer [87].

Qiu *et al.* further concluded that the expression patterns of EBV miRNAs from two epithelial cancers (5 NPCs and 6 EBVaGCs) were extremely similar [87]. To gain insight into the expression pattern in other epithelial cancers, our team employed home-designed QRT-PCR assays to profile the expression of EBV miRNAs in two common and two rare EBV-associated epithelial malignancies in our local primary biopsies, including NPC (n = 23), EBVaGC (n = 10), LELC of the lung (n = 9) and lymphoepithelioma-like CCA (n = 6). The presence of EBV in the samples was confirmed by EBER *in situ* hybridization (ISH) on the paraffin-embedded tissue section. Since the tumor contents and transcription activities among individual samples might be different, we measured both the absolute miRNA copy number per nanogram of total RNA and the fraction of viral miRNAs. To obtain an absolute quantification of each miRNA in the sample, an in parallel experiment using synthetic DNA-RNA chimeric oligonucleotides of each mature miRNA sequence as a template for a standard curve was included (Table S1). On the whole, all miR-BARTs were favorably expressed in all the tested EBV associated malignancies except for miR-BART20 and miR-BART21, which showed low expression levels of 10^3 copies per nanogram of total RNA. The low expression levels of these two miRNAs aligned with our previous findings in a small-scale cloning analysis on NPC samples [23]. Whereas the average copy numbers of BART miRNAs expressed in NPC and LELC of the lung was around 5–10 fold higher than EBVaGC and CCA, the miR-BARTs fraction profile of four tumor types were vastly similar (Figure 1).

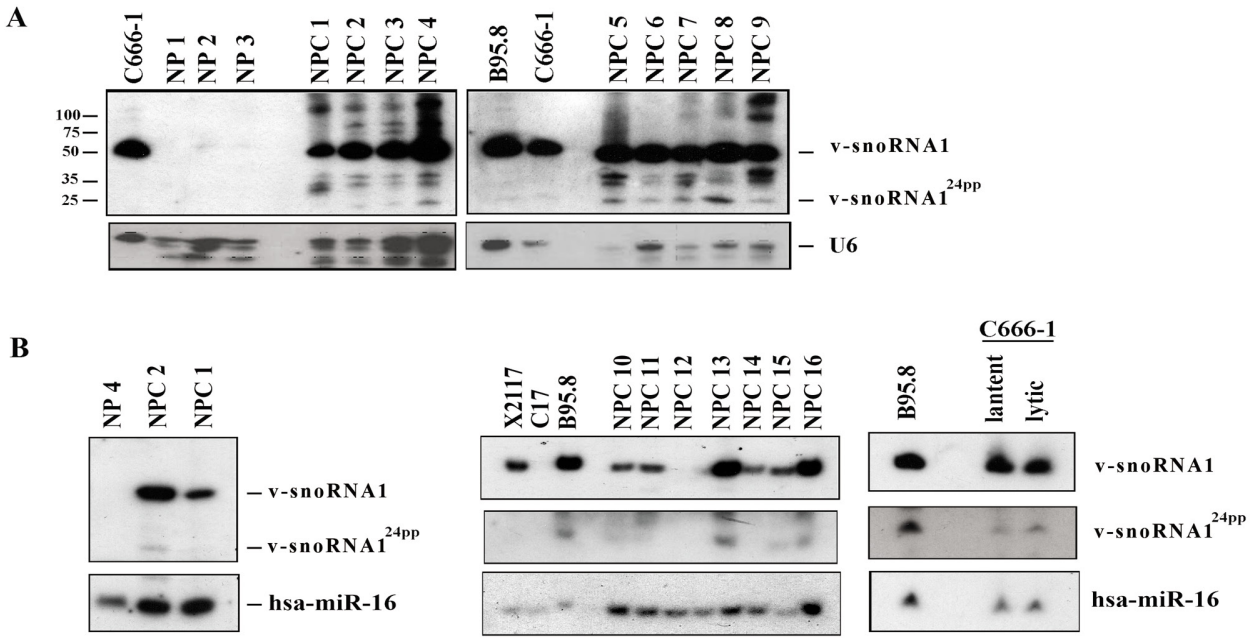
5. Recent Research Development of Viral Small Nucleolar RNA (v-snoRNA1)

5.1. Expression of v-snoRNA1 in NPC

Apart from the production of miRNAs, abundant BARTs expression in epithelial cancer may have another important role by generating v-snoRNA1 and v-snoRNA1^{24pp}. However, v-snoRNA1 expression is so far only described in EBV infected B-cells. In our recent study, we aimed to characterize the expression of v-snoRNA1 in NPC. We noticed in Northern blot analysis that v-snoRNA1 was highly expressed in almost all NPC samples including cell lines, xenografts and biopsies. Furthermore, v-snoRNA1^{24pp} was detected in a portion of the NPC samples (Figure 2A,B). Previous work by Hutzinger *et al.* demonstrated that inducing EBV lytic replication

could upregulate both v-snoRNA1 and v-snoRNA1^{24pp} production in wild type EBV infected 293 cells [76]. In contrast, we demonstrated that in native EBV infected C666-1 cells, induction of the lytic cycle could only slightly up-regulate v-snoRNA1^{24pp} expression and had no observable effect on v-snoRNA1 expression (Figure 2B).

Figure 2. Expression of v-snoRNA1 in NPC: (**A**) Expression of v-snoRNA1 and v-snoRNA1^{24pp} on NPC biopsies (NPC 1 to NPC 9) were analyzed by Northern blot analysis. Biopsies from noncancerous nasopharyngeal tissues (NP 1 to NP 3), NPC cell lines (C666-1) and B95.8 cells were included as controls. The membranes shown in Figure 2A were used previously in our publication [93]; (**B**) To confirm the result on Figure 2A, Northern blot analysis for samples from NPC 1, NPC 2 (*left panel*) and another 7 new NPC biopsies (*middle panel*) were performed on the newly prepared membranes. The effect of EBV replication on v-snoRNA1 expression in C666-1 cells is shown in the *right panel*.



5.2. Nucleotide Polymorphism Is Important for v-snoRNA1 Production

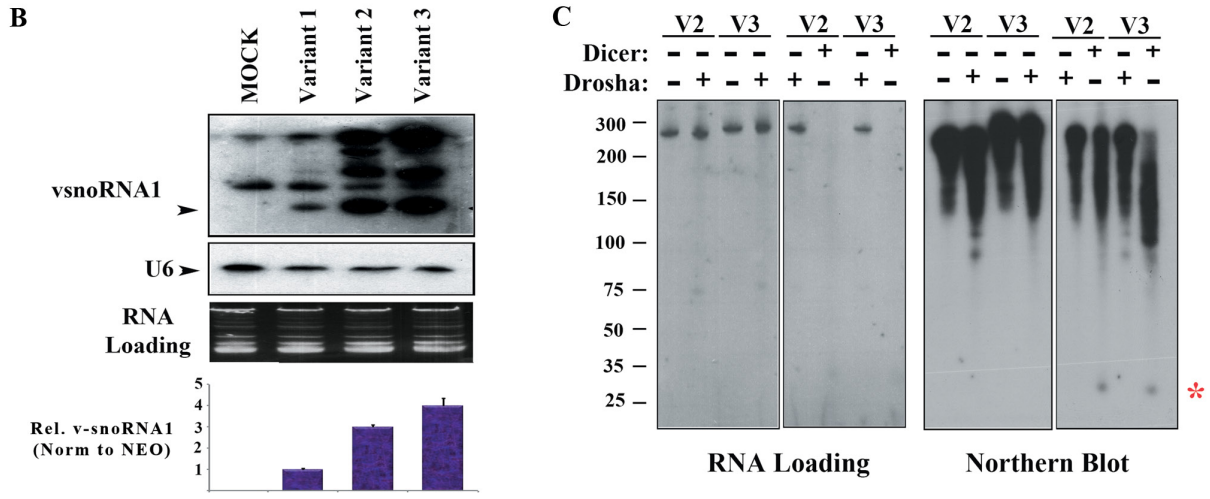
We next compared the v-snoRNA1 sequence in our available cell lines and NPC xenografts to the B95-8-EBVstrain by direct sequencing. Our finding was consistent with the previous observation by Hutzinger *et al.* that v-snoRNA1s from different EBV strains are highly conserved. However, several nucleotide polymorphisms at a position distal from the 3' end of v-snoRNA1 were observed. BC-1-EBV and B95-8-EBV strains had the same sequence polymorphism (variant 1), and all the other tested samples carried another sequence pattern as the C666-1-EBV strain (variant 2) (Figure 3A). To further analyze the variation pattern of EBV strains from 14 normal controls and 31 NPC patients in our population, we showed that the variant distribution of EBV strain was highly similar between NPC patients and normal controls, in which variant 2 accounted for ~70% in our population. Another variant (variant 3) with a three-nucleotide polymorphism to variant 2 accounted for 24% in our locality (Figure 3A). EBV nucleotide variation positioned distal from the hairpin structure was previously reported to affect miR-BART processing [23].

Figure 3. Nucleotide polymorphism of vsoRNA1 flanking sequence: (**A**) The three nucleotide polymorphism patterns (V1–V3) positioned distal from the 3' end of v-snoRNA1 were identified. The sequence shown is corresponded to 152925:152985 (AJ507799.2); (**B**) Expression level of v-snoRNA1 was analyzed by Northern blot using RNA from 293FT cells transiently transfected with expression vectors containing different v-snoRNA1 variants as indicated (AJ507799.2, 152761:153032); (**C**) *In vitro* Dicer and Drosha/DGCR8 processing on the long v-snoRNA1 containing fragment. The digestion products were visualized on 10% PAGE (*left panel*) and analyzed by Northern blot with v-snoRNA24pp complementary oligonucleotide probe (*right panel*). The signal of v-snoRNA24pp is showed as "*". The representative result is shown.

A

```
v-snoRNA1-3'CUGAUGAG-- ---------- ---------- ---------- ---------- ----------
B95.8 (V1)   CUGAUGAGAG GCUUGGGGGG GUCUUUGCCC UCAAGGCGAG GCUCCUUCUC CUAGGAAUGC
C666-1 (V2)  CUGAUGAGAG GCUUGGGGGG -UCUUUGCCC UCAAGGUGAG GCUCCUUCU- --AGGAAUGC
NPC (V3)     CUGAUGAGAG GCUUAAGGGG -UCUUUGCCC UCAAGGUGAG GCUCCUCCU- --AGGAATGC
```

	EBV Strain			
	Normal Control	NPC Patient	Overall	Tested EBV positive sample
Variant 1	1 (7%)	2 (6%)	3 (7%)	Cell lines: B95.8 (Prototype EBV), BC-1 (B-cell)
Variant 2	9 (64%)	22 (71%)	31 (69%)	Cell lines: Akata (B-cell), C666-1 (NPC), CB14022 (LCL), Jiyoye (B-cell); Local NPC xenograft: X666; X1915; X2117; X99186
Variant 3	4 (29%)	7 (23%)	11 (24%)	
Total	14	31	45	



To investigate whether the common variant would affect v-snoRNA1 production, we cloned different variants of v-snoRNA1 containing fragment (~270 nt) into the pcDNA 3.1 (+) expression vector and transiently expressed them into 293 cells. Subsequent Northern blot analysis on the RNA extracts revealed that both fragments from variants 2 and 3 displayed higher expression levels of v-snoRNA1 than the prototype EBV variant 1 (Figure 3B). Surprisingly, the predicted proximal RNA hairpin structures among these three variants are extremely similar to each other (Figure S3). In this sense, the Drosha/DGCR8 complex, which chopped off the hairpin structure (premature miRNA) from the long RNA transcripts, might not be responsible for viral v-snoRNA1 biogenesis. We further substantiated this hypothesis by performing additional *in vitro* enzymatic digestion on T7 transcribed RNA fragments from expression vectors that disabled Drosha/DGCR8 complex's activity to generate

v-snoRNA1 [23,94] (Figure 3C). Thus, the v-snoRNA1 or v-snoRNA^{24pp} biogenesis pathway might be dependent on other microprocessor complexes such as the Integrator complex, which was previously proved to take part in the viral pre-miRNA processing in primate herpesvirus [95]. On the other hand, we demonstrated that Dicer activity alone was sufficient to process the *in vitro* transcribed long RNA fragment into v-snoRNA1^{24pp}. Hence, v-snoRNA^{24pp} may be directly produced from v-snoRNA1 or it longer precursor with the help of Dicer activity (Figure 3C).

6. Functional Roles of miR-BARTs in Cancer Development

6.1. Methods for miRNA Targets Identification

Viral miRNA is expected to contribute to cancer development by silencing both the viral and cellular target transcripts in the post-transcriptional level. Thus, identification and validation of the downstream targets of miR-BARTs are critical to understand how EBV is involved in tumorigenesis. It is generally accepted that a perfect match between the miRNA seed sequence (nucleotides 2–8) and the target 3'UTRs is important in the interaction. However, the exact parameter to determine miRNA target recognition is not fully characterized. As a result, a huge amount of false positive targets are predicted by using *in silico* miRNA target prediction method. To avoid considering the elusive universal rule for target site binding, high throughput sequencing of RNAs isolated by cross-linking and immunoprecipitaion (HITS-CLIP) [96] and photoactivatable ribonucleoside enhanced CLIP (PAR-CLIP) [97,98] methods have recently been employed to identify the endogenous miR-BARTs associated targets [99–101]. These biochemical methods have been reported to identify a hundred of miR-BART cellular targets, indicating the heavy involvement of EBV in regulating the cellular transcriptome. Remarkably, most of the known targets of miR-BARTs are mainly responsible for the control of viral latency, host cell immunity and apoptosis [100,101].

6.2. Regulation of Viral Gene Expression

Although target identification is the most challenging part in miRNA study, the targets of some viral miRNAs are easy to identify because they are transcribed as an antisense sequence of the viral genes. The first EBV miRNA identified to target viral product was miR-BART2-5p, which was transcribed in the opposite direction to the 3'UTRs of the viral DNA polymerase gene, *BALF5* [73]. With the perfect complementarity between miR-BART2-5p and BALF5 transcript, miR-BART2-5p was finally proved to down-regulate BALF5 expression via siRNA cleavage at the binding site [102]. Similarly, vsnoRNA1^{24pp}, positioned distal from miR-BART2 and fully complementary to the BALF5 3'UTR, was indirectly proved to down-regulate BALF5 expression by cleaving the transcript [76]. Being a critical protein for DNA replication, down-regulation of BALF5 may facilitate EBV to enter latency.

Apart from the direct cleavage mechanism, EBV-miRNAs also regulate viral gene production by associating with their 3'UTRs. It had been revealed by our previous work in NPC that expression of LMP1 and LMP2A was downregulated by three BART cluster 1 miRNAs (miR-BART1-5p, 16 and 17-5p) and BART22, respectively [23,103]. Although both LMP1 and LMP2A can promote cell proliferation, migration and invasion, high expression of these two proteins in the infected cells may have destructive effects. LMP2A is the most immunogenic viral protein in latency I and II infection,

and circulating T-cells specific for LMP2A derived epitopes are frequently detected in healthy individuals [104,105]. Limiting LMP2A expression by miR-BART22 would have potential advantages for the cells to escape host immune surveillance. On the other hand, high LMP1 expression can suppress cell growth and guide the cells to commit apoptosis in response to several stimuli like TNF treatment and serum depletion [106,107]. Thus, LMP1 expression in cancer should be tightly regulated in order to maintain the balance between its growth promoting and suppressing effects [103]. Fascinatingly, more miR-BARTs have been identified to regulate LMP1 expression in lymphoma cell lines. Ranakrishnan *et al.* showed that miR-BART9 from the NK/T-cell lymphoma cell line SNK6 could increase viral LMP1 expression. This might be the key mechanism of EBV to promote cell growth [108]. However, the mechanism behind this regulation was not clearly demonstrated. Using Ago HITS-CLIP direct biochemical assay in EBV-transformed B cells (Jiyoye cells), Riley *et al.* further recognized that both miR-BART5 and miR-BART19-5p could down-regulate LMP1 expression [101]. While miR-BART5 and hsa-miR-18-5p share the same seed sequence, hsa-miR-18-5p is unable to target LMP1 [101]. This suggested that additional factors, in addition to the perfect Watson-Crick complementary between target 3'UTR and miRNA seed sequence, should also be considered for miRNA suppressive activity [23]. Riley *et al.* further uncovered that miR-BART10-3p could work together with miR-142-3p and miR-17-5p to co-repress viral anti-apoptotic gene *BHRF1* in latency III EBV infection cells. The means to target *BHRF1* was unknown, but blocking these miRNAs individually in Jiyoye cells definitely increased the cellular apoptotic rate [101]. Table 1 lists the confirmed viral targets of miR-BARTs.

Table 1. Function of EBV miRNAs in viral targets.

miRNA	Target	Function	miRNA effect	References
v-snoRNA^{24pp}	BALF5	DNA polymerase	Maintain latency	[76]
BART2-5p	BALF5	DNA polymerase	Maintain latency	[73,102]
BART1-5p/BART16/BART17-5p	LMP1	Transforming factor	Anti-apoptosis	[103]
BART19-5p/BART5	LMP1	Transforming factor	Immune evasion;	[101]
BART9	LMP1	Transforming factor	Promote cell growth (NK/T cells)	[108]
BART22	LMP2A	Signaling molecule	Immune evasion	[23]
BART10-3p	BHRF1	Homolog of Bcl2	Unknown	[101]

6.3. Regulation of Cellular Gene Expression

As a successful pathogen, EBV can develop life-long latency in the host with subsequent reactivation. For this reason, EBV needs to protect itself by tightly controlling viral gene expression and regulating host signaling cascades. It is not surprising to learn that miR-BARTs are the key regulators for host gene expression because only a few of viral proteins are expressed in carcinomas.

Until now, multiple lines of evidence have suggested that miR-BARTs target cellular genes mainly for preventing apoptosis and escaping the host immune system. The p53-up-regulated modulator of apoptosis (PUMA) is the first recognized cellular target of miR-BARTs and was targeted by miR-BART5 via the interaction on its 3'UTR [93]. The PUMA is one of the six members of the BH3-only group in the Bcl2 family. The BH3-only proteins, the essential initiators of apoptosis, are

responsible to control the release of cytochrome C from the mitochondrial inter-membrane [109]. Although PUMA is the immediate downstream target of the well-known tumor suppressor gene *p53*, it functions as a key pro-apoptotic protein under the influence of the p53 tumor suppressor and other apoptotic stimuli. With apoptotic stimulation, PUMA is believed to interact with Bcl-2 family proteins and helps to mediate mitochondrial dysfunction, followed by the activation of the caspase cascade and eventually induces apoptosis. The interaction of miR-BART5 and PUMA on cell apoptotic effects was clearly demonstrated in EBV-positive NPC cell lines, in which depletion of miR-BART5 or ectopic expression of PUMA could make the cells more susceptible to apoptosis-inducing drugs [93].

Another BH3-only group protein, Bcl-2 interacting mediator of cell death (Bim), was also shown to be the target of the multiple miRNAs encoded in the BART cluster 1 [110]. Interestingly, Bim's 3'UTR is not responsive to any of the individual miR-BARTs in stable transfectants. This observation indicates the importance of the co-operation of miR-BARTs in cluster 1 for Bim expression. In another study on EBV positive NK/T cell lymphoma, Lin *et al.* demonstrated by transient transfection and Western blot analysis that miR-BART20-5p could downregulate the T-box transcription factor TBX21 expression via the binding site located on the 3'UTR of the target mRNA [111]. They further revealed that inhibition of the TBX21 activity indirectly suppressed p53 function in NK/T lymphoma cell lines. This finding underlined the significance of miR-BART20-5p on preventing p53 dependent apoptosis. However, this miRNA is generally weakly expressed in epithelial cancers (Figure 1).

By using direct co-immunoprecipitated miRNA/mRNA containing RISC method, several studies identified a number of putative anti-apoptotic targets regulated by miR-BARTs. TOMM22 was identified and validated as a potential target of miR-BART16 [112]. This protein is a part of the mitochondrial membrane receptor involved in the association of the pro-apoptotic protein, Bcl-2-associated X (Bax). Treatment with antibodies against TOMM22 or siRNA knockdown has been shown to inhibit the association of Bax protein on mitochondria, thus preventing Bax-dependent apoptosis [113]. In Burkitt's lymphoma, BART miRNAs were found to exert anti-apoptotic effects by repressing CASP3 expression, in which the co-suppressive effects of miR-BART1-3p and -16 on CASP3 expression was confirmed by reporter assay [114]. Riley *et al.* also reported in HITS-CLIP result that both viral miR-BART13-3p and host miR-17 could inhibit Caprin-2 translation [101]. Despite Caprin-2's ability to promote activation of the oncogenic Wnt signaling pathway [115], over-expression of this protein would drastically induce apoptosis [116].

Surprisingly, not all miR-BARTs promote cell growth or inhibit apoptosis. Choi *et al.* demonstrated the tumor suppressive effect of miR-BART15 in gastric cancer [117]. The BRUCE protein is a member of the inhibitor of apoptosis (IAP) family that protects cells from apoptosis by ubiquitin-dependent degradation of caspase protein [118,119]. It was indicated in transient transfection assay that miR-BART15 could suppress cell growth and induced early apoptosis in AGS gastric carcinoma cell in part through targeting the *BRUCE* gene. By carrying a similar "seed sequence" with miR-223, miR-BART15 was also reported to repress NLRP3 expression via the miR-223 putative binding sites on the 3'UTR of NLRP3 [120]. NLRP3 is a NOD-like receptor protein that forms a multi-protein complex called the inflammasome. In response to pathogen invasion, NLRP3-inflammasome inside the cells can be activated to trigger the secretion of interleukin-1β (IL-1β), finally resulting in the stimulation of host inflammatory response. Thus, the presence of miR-BART15 may protect the infected cells by functioning as a NLRP3 blocker. Moreover, miR-223 has been proved as a tumor

suppressive miRNA in other epithelial cancers. For example, it can suppress oncogenic protein (stathmin1) expression in liver and gastric carcinomas [121,122]. Yet, the effect of miR-BART15 on the same oncogenic proteins has not been evaluated. Further studies may be needed to uncover the complexity of the miRNA network in carcinogenesis.

Recently, several miR-BARTs in infected cells were selectively secreted as an exosome and transported to the adjacent cells to exert their suppressive function [120,123,124]. In addition, the exosomal miR-BART15 from EBV-infected B cells was found to be sufficient for the inhibition activity of the NLRP3 inflammasome in the PMA-differentiated macrophage cell line THP-1 [120]. In this sense, miR-BART15 might indirectly contribute to cancer development by exercising its "anti-cancer" activities in the adjacent immune cells [117].

Besides BART15, two more miR-BARTs were also reported to augment cancer progression by evading host immunity system. For example, EBV miR-BART2-5p has been shown to regulate the stress-induced *MICB* gene [125]. Since MICB is a key cell surface ligand for natural killer (NK) cells recognition, down-regulation of MICB expression may allow them to escape direct killing by NK cells. The innate immunity related gene *importin 7* (*IPO7*) was a putative target recognized by miR-BART1-3p and BART3 [112,114,126]. The *importin 7* gene encodes a critical receptor protein responsible for importing transcription factors into the nucleus [127]. It is believed that IPO7 is involved in the innate immunity system because knocking down of this gene in macrophage reduces pro-inflammatory cytokine IL-6 secretion [112]. The function of IPO7 in epithelial cells is largely unknown. Therefore, miR-BART3 may mainly exert its function in neighboring immune related cells via exosome.

Aside from genes implicated in apoptosis and immune evasion, a few studies independently reported the function of miR-BARTs on regulating tumor suppressor genes and the gene responsible for latency maintenance. Although the direct interaction between miRNAs and the corresponding target sites were not assessed, Wong *et al.* identified in a transient transfection assay that miR-BARTs could extensively down-regulate the expression of several well-known Wnt inhibitory molecules, such as *WIF1* (miR-BART19-3p), *APC* (miR-BART19-3p, 7 and 17-5p) and *NLK* (miR-BART19-3p, 14 and 18-5p) [92]. On the other hand, miR-BART3* has been showed to down-regulate the expression of tumor suppressive gene *DICE1* in NPC cells [128]. Since inactivation of *DICE1* occurs frequently in non-EBV associated solid tumors by either allelic deletion or CpG promoter hypermethylation [129–131], this finding has highlighted the importance of miR-BART3 to downregulate DICE1 activity in EBV related carcinogenesis.

It has been shown in primary B cells that EBV infection could globally suppress the expression of cellular miRNAs, indicating the presence of strong miRNA repressors within the EBV genome [132]. The miR-BART6-5p may be one of such EBV-derived repressor since it can inhibit the expression of human Dicer, a critical enzyme for miRNA biogenesis, via four binding sites within the 3'UTR [133]. In addition, knocking down miR-BART6-5p activity in EBV positive cell lines with latency I or II program can enhance Dicer expression and indirectly re-activate the expression of both lytic and highly immunogenic viral proteins such as Zta, Rta and EBNA2 [133]. These findings suggested that the key function of miR-BART6-5p might assist infected cells to maintain latency.

7. Conclusions and Future Perspectives

In NPC and EBVaGC, the clonal EBV genome is consistently detected in all invasive carcinoma and high-grade dysplastic lesions. This observation underscores the importance of viral infection in the development of epithelial cancers. In this context, EBV should have the ability to amend several cellular signaling pathways for both host cell survival and to facilitate tumor progression. In contrast to the low viral protein levels, functional BART miRNAs are generally highly expressed in a number of epithelial cancers (Figure 1). Although the regulatory mechanism by which miRNAs affect tumor progression is not fully understood, some viral miRNAs have already been recognized to regulate expression of several key cancer-related proteins, including those responsible for facilitating latency maintenance, immune suppression and tumor promotion (Tables 1 and 2). For this reason, suppressing oncogenic viral miRNAs (viral onco-miR) activity in the cancer cell may provide a therapeutic benefit for patients.

Table 2. Function of EBV miRNAs in cellular target.

miRNA	Target	Function	miRNA effect	References
BART5-5p	*PUMA*	Proapoptotic protein	Anti-apoptosis	[93]
BARTs cluster 1	*Bim*	Proapoptotic protein	Anti-apoptosis	[110]
BART20-5p	*T-bet* (*TBX21*)	T-box transcription factor	Anti-apoptosis	[111]
BART16	*TOMM22*	Mitochondrial receptor for proapoptotic protein BAX	Anti-apoptosis	[112]
BART1-3p + BART16	*CASP3*	Proapoptotic protein	Anti-apoptosis	[114]
BART13-3p	*CAPRIN2*	Wnt-signaling enhancer	Anti-apoptosis	[101]
BART15	*BRUCE*	Anti-apoptotic protein		[117]
BART15	*NLPR3*	Regulation of inflammation	Immune evasion	[120]
BART2-5p	*MICB*	NK cell ligand	Immune evasion	[125]
BART3	*IPO7*	Nuclear importer protein	Immune evasion	[112]
BART3 + BART16	*IPO7*	Nuclear importer protein	Immune evasion	[114]
BART19-3p	*WIF1*	Wnt inhibitor	Oncogenic properties	[92]
BART19-3p, 7 and 17-5p	*APC*	Wnt inhibitor	Oncogenic properties	[92]
BART19-3p, 14 and 18-5p	*NLK*	Wnt inhibitor	Oncogenic properties	[92]
BART3*	*DICE1*	Tumor suppressor	Oncogenic properties	[128]
BART6-5p	*DICER*	miRNA biogenesis	Unknown	[133]

After the discovery of the first human miRNA 10 year ago, only the antagomir against miR-122 (SPC3649, Miravirsen) can enter into Phase II clinical trials for Hepatitis C treatment [134,135]. However, this therapeutic strategy is being continuously developed and improved, especially on the issues of delivery system and inhibitory efficiency of the therapeutic particles [136,137]. Currently, the schemes for targeting miR-BARTs are reviewed in great detail by Lo AK *et al.* [138]. In theory, EBV miRNAs are exogenous genes solely expressed in cancer cells, and display low sequence homology to known human miRNAs [139]. Advantage would be given to the development of miR-BART targeting therapeutic approach because of the tissue specific delivery and the potential off-target effect of the designed bullet. However, selecting suitable viral onco-miR candidates for targeting therapy is challenging. In this review, we revealed that the miR-BARTs expression patterns of four common

EBV-associated epithelial cancers are highly similar (Figure 1). This observation has suggested that most EBV-positive epithelial cancers employ similar regulatory mechanisms for controlling ebv-miRNA expression, thus a single formula of miRNA inhibitor may be possible for the treatment of those malignancies. Nevertheless, our understanding of viral miRNA function is still far from complete. Despite that several important cellular viral miRNA targets have been identified, their interactions were solely substantiated in cell culture experiments in which non-physiological levels of miRNAs were used. In addition, a single miRNA molecule might be possible to regulate numerous signaling cascades in response to different cellular environments. Therefore, all the identified miRNA targets should be tested in *in vivo* experiments. We believe that a comprehensive understanding of miR-BART function and expression pattern in epithelial cancers is crucial in designing effective therapeutic regimens to combat EBV-associated malignancies.

The involvement of EBV miRNAs on cellular interactions is even more complicated. Currently, certain EBV miRNAs were shown to be actively secreted from infected cells and transported to neighboring cells through a group of small endosomal-derived vesicles called "exosomes" [124,140–142]. It had been demonstrated in *in vitro* and *in vivo* experiments that secretory exosomal ebv-miRNAs can be internalized and exhibit biological functions in the surrounding uninfected recipient cells such as primary monocyte-derived dendritic cells (MoDC) [123,124,143,144]. Thus, EBV seems to utilize the exosome as a vehicle to extend their miRNA function on the neighboring cells for immune evasion. Being protected from enzymatic degradation by membrane bound microvesicles, exosomal miRNAs are efficiently recovered from biological fluids such as plasma, pleural fluid or sputum. This characteristic has lead scientists to consider exosomal viral miRNA as a non-invasive biomarker. However, the study of exosomal miRNAs is only at the initial stage, and the underlying mechanism to regulate its secretion remains largely unclear. For example, circulating miR-BART17 was consistently detected in plasma and serum samples from NPC patients [92,141,145], yet, the reason why the same miRNA was not enriched in the exosomal fraction of the same patients was unknown [145]. In addition, the exosomal miRNA content in hepatocytes has been reported to vary between different pathological and physiological conditions [146]. Whether these conditions may affect the composition of exosomal viral miRNAs in EBV positive cancer patients is not characterized. To this end, the present information is not sufficient to support the use of exosomal content as a biomarker. Nonetheless, Arroyo *et al.* found that the majority of circulating plasma miRNAs was protected by Ago2 complexes [147]; it would be interesting to know if the HITS-CLIP method can be used for biomarker identification.

Despite the uncertainty of the nature of exosomal miRNAs, some EBV miRNAs exhibit a higher expression in the blood from NPC patients compared to non-NPC controls [92,145,148]. The preliminary work from Courzones and colleagues further revealed that the plasma concentration of miR-BART17 in NPC patients is partly dependent on the tumor mass [145]. Surprisingly, there seemed to be no correlation between the concentration of plasma viral miRNAs and EBV DNA load in NPC patients [145,148]. While cell-free EBV DNA represents an effective biomarker for NPC diagnosis and prognosis, incorporating miR-BARTs in the analysis may provide additional information to enhance the assessment of the tumor status. We envisage that additional experimental data would warrant a more solid conclusion on this issue.

Acknowledgments

This research was supported by the Collaborative Research Fund from the Hong Kong Research Grants Council (CUHK04/CRF/08) and RFCID (11100162). The authors thank Elaine S. Chan (CUHK) and Y.K. Tong (CUHK) for critical reading of the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

Appendix

Materials and Methods

Patient Samples

All specimens in this study were recruited at the Prince of Wales Hospital, Hong Kong. They are included frozen biopsy tissues of 21 NPCs, one EBVaGC and two LELC of lungs, formalin fixed paraffin embedded (FFPE) specimens from 9 cases of EBVaGC, 7 cases of LELC of lung and 7 cases of LELC of CCA. The present of EBV in the samples were confirmed by EBV-encoded small RNA (EBER) *in situ* hybridization (ISH) on the paraffin-embedded tissue sectioned. All participants provided written informed consent for the collection of samples and subsequent analysis. Ethics approval was obtained from the Joint Chinese University of Hong Kong-New Territories East Cluster Clinical Research Ethics Committee, Hong Kong.

RNA Extraction and Quantitative Reverse Transcription PCR (QRT-PCR)

Total RNA from biopsies and FFPE samples were prepared by TRIzol reagent (Invitrogen, Carlsbad, CA, USA) and Recover All Total nucleic Acid Isolation kit (Ambion Inc, Austin, TX, USA) according to the manufacturer's recommendations respectively. The concentration and quality of all RNA was determined by using a NanoDrop-1000 UV-VIS Spectrophotometer (NanoDrop Company, Wilmington, DE, USA). One microgram of total RNA was reverse-transcribed in 20 μL of reaction with miScript Reversed Transcription Kit (Qiagen, Hilden, Germany) and diluted 5-fold for individual QRT-PCR assay. MiScript SYBR green PCR kit (Qiagen, Hilden, Germany) with kit provided universal primer and in-house designed miRNA-specific forward primers were used for QRT-PCR reaction with the following thermal cycling condition: 95 °C for 15 min, followed by 40 cycles at 94 °C for 15 s, annealing with specific temperature for 30 s and 70 °C for 30 s. The PCR was carried out in LightCycler® 480 Real-Time PCR System (Roche Diagnostic) and the crossing point (Cp) was analyzed by second derivative method. To eliminate the non-specific SYBR green signal, PCR result will be normalized with the signal generated by EBV negative epithelial cell line NP69. Standard curves were included in each experiment for absolute quantification of each miRNA per nanogram of total RNA using serially-diluted synthetic DNA-RNA chimeric oligonucleotides (Integrated DNA Technologies, Coralville, IA, USA) with mature miRNA sequences as RNA template (from 10^7 to 10^3 copies; 10 μM = 6.02×10^{12} copies per micro-liter). It had been demonstrated by other team

that standard curve generated from single-stranded oligonucleotides is similar to that obtained by using T7-transcribed RNA [149]. Analysis of each sample was performed in triplicate. The oligonucleotide sequences and specific annealing temperature used for QRT-PCR are shown in Table S1.

Expression of BHRF1 and Northern Blot

Expression of BHRF1 mRNA was analyzed by an RT-PCR based assay as previous described [150]. In brief, 1 μg of RNA was reversed transcribed and used to perform PCR reaction with BHRF1 specific primers (5'-GTC AAG GTT TCG TCT GTG TG-3' and 5'-TTC TCT TGC TGC TAG CTC CA-3') and actin specific primers (5'-TAA GGA GAA GCT GTG CTA CGT C-3' and 5'-GGA GTT GAA GGT AGT TTC GTG G-3). PCR products were analyzed in 2% agarose gel and transferred onto Nytran SPC membrane (Whatman Inc, Piscataway, NJ) for Southern blot analysis using γ-32P ATP end-labeled synthetic oligonucleotides complementary to the lytic form of BHRF1 (5'-ATG CAC ACG ACT GTC CCG TAT ACA C-3'). Northern blot was performed as described previously [23,122]. The sequence of synthetic oligonucleotides used as probes for EBER1-5p and 3p were EBER1-5p: 5'-TAG GGC AGC GTA GGT CCT-3' and EBER1-3p: 5'-AAA CAT GCG GAC CAC CAG CTG G-3'. As a positive control, 5 μg of RNA from EBV-positive marmoset cells line B95-8 was included for Northern blot analysis.

AGO2 Co-Immunoprecipitation

The AGO2 co-immunoprecipitation method was described previously [94].

Induction of Viral Lytic Cycle in NPC Cell Lines

C666-1 cells were treated with 100 ng/mL of Gemcitabine (GEM, Eli Lilly and Company, IN, USA) and 100 ng/mL Trichostatin A (TSA, Sigma-Aldrich, Saint Louis, MO, USA) for 48 days. The induction of viral lytic cycle in C666-1 cells was confirmed by Western blot analysis using antibodies targeting early viral lytic genes, anti-EA-D (BMRF1, Millipore Corp, Billerica, MA, USA), anti-EBV transcription factor R (BRLF1, Argene, Verniolle, France) and anti-EBV-ZEBRA (Argene, Verniolle, France).

Plasmid Constructs

The v-snoRNA1 expression vector were made by inserting the PCR fragments that contains v-snoRNA1 flanking sequence (AJ507799.2, 152761:153032) into pcDNA3.1 expression vector through *Hind III* and *Xba I* sites. The PCR products were amplified using genomic DNA from B95.8 (variant 1), C666-1 (variant 2) and one of the NPC biopsy (variant 3) as template. The sequences of the primes used for the PCR are F: GCC AAG CTT GCC CTT GCG TGT C and R: GGC TCT AGA AGG CTG GCA AAG ATC.

In Vitro Processing Assay

In vitro Drosha/DGCR8 and Dicer digestion assay was described previously [23].

Table S1. Oligonucleotide sequences and the specific annealing temperature used for QRT-PCR.

miRNA	Specific forward primer	Chimeric miRNA mimic	Annealing temp.
BART1-3p	AGCACCGCTATCCACTATGT	TAGCACCGCTATCCACTATrGrUrC	55
BART1-5p	TCTTAGTGGAAGTGACGTGCT	TCTTAGTGGAAGTGACGTGCTrGrUrG	60
BART2-3p	AGGAGCGATTTGGAGAAAATAA	AAGGAGCGATTTGGAGAAAATrArArA	60
BART2-5p	TATTTTCTGCATTCGCCCTTGC	TATTTTCTGCATTCGCCCTrUrGrC	60
BART3	CACCACTAGTCACCAGGTGT	CGCACCACTAGTCACCAGGrUrGrU	60
BART4	ACCTGATGCTGCTGGTGTGC	GACCTGATGCTGCTGGTGTrGrCrU	64
BART5	AAGGTGAATATAGCTGCCCAT	CAAGGTGAATATAGCTGCCCArUrCrG	55
BART6-3p	GGGATCGGACTAGCCTTAGA	CGGGGATCGGACTAGCCTTrArGrA	55
BART6-5p	TAAGGTTGGTCCAATCCATAGG	TAAGGTTGGTCCAATCCATrArGrG	55
BART7	CATCATAGTCCAGTGTCCAGGG	CATCATAGTCCAGTGTCCArGrGrG	60
BART8	TACGGTTTCCTAGATTGTACAG	TACGGTTTCCTAGATTGTArCrArG	55
BART9	TAACACTTCATGGGTCCCGTAGT	TAACACTTCATGGGTCCCGTrArGrU	55
BART10	TACATAACCATGGAGTTGGCTGT	TACATAACCATGGAGTTGGCrUrGrU	60
BART11-3p	ACGCACACCAGGCTGACTG	ACGCACACCAGGCTGACTrGrCrC	62
BART11-5p	AGACAGTTTGGTGCGCTAGT	TCAGACAGTTTGGTGCGCTAGrUrUrG	55
BART12	CTGTGGTGTTTGGTGTGGTT	TCCTGTGGTGTTTGGTGTGrGrUrU	62
BART13	ACACTCCAGCTGGGTGTAACTTGC CAGGGA	TGTAACTTGCCAGGGACGGCrUrGrA	62
BART14	TAAATGCTGCAGTAGTAGGGA	TAAATGCTGCAGTAGTAGGrGrArU	55
BART15	TCAGTGGTTTTGTTTCCTTGA	GTCAGTGGTTTTGTTTCCTrUrGrA	60
BART16	TAGATAGAGTGGGTGTGTGCTC	TTAGATAGAGTGGGTGTGTGCrUrCrU	64
BART17-3p	GTATGCCTGGTGTTCCCCTTA	TGTATGCCTGGTGTCCCCTTrArGrU	60
BART17-5p	TAAGAGGACGCAGGCATACA	TAAGAGGACGCAGCATACrArArG	55
BART18-3p	TATCGGAAGTTTGGGCTTCGT	TATCGGAAGTTTGGGCTTCrGrUrC	60
BART18-5p	TCAAGTTCGCACTTCCTATAC	TCAAGTTCGCACTTCCTATrArCrA	55
BART19-3p	ACACTCCAGCTGGGUUUUGUUUG CUUGGGA	TTTTGTTTGCTTGGGAATrGrCrU	55
BART19-5p	ACATTCCCCGCAAACATGACAT	ACATTCCCCGCAAACATGACrArUrG	55
BART20-3p	CATGAAGGCACAGCCTGTTAC	CATGAAGGCACAGCCTGTTrArCrC	60
BART20-5p	TAGCAGGCATGTCTTCATTCC	TAGCAGGCATGTCTTCATrUrCrC	60
BART21-5p	CACTAGTGAAGGCAACTAAC	TCACTAGTGAAGGCAACTrArArC	55
BART22	TTACAAAGTCATGGTCTAGTAGT	TTACAAAGTCATGGTCTAGTrArGrU	55

Figure S1. Lack of association between novel EBER end terminal fragments and AGO2: (**A**) Expression of small EBER fragments in NPCs was validated by Northern blot analysis. Locations of the EBER1-5p and -3p fragments are indicated with arrow; (**B**) Western blot analysis of hAGO2 protein from C666-1 whole cell lysate before (Input) and after immunoprecipitation (IP) with anti-AGO2 antibody (*left panel*). Monoclonal antibody against FLAG tag (lane 2) and against hAGO2 4 (lane 3) were used for IP from C666-1 cell lysate (lane 1). AGO-2 associated miRNAs were analyzed by Northern blot analysis with synthetic oligonucleotide probe for the specific small EBV fragments listed below the figure. Locations of the probed fragments are indicated with arrows. RNA loading was visualized by SYBR Gold stained PAGE. BART22 IP was included as positive control. The information of EBER1-5p and 3p were described in a previous publication [23].

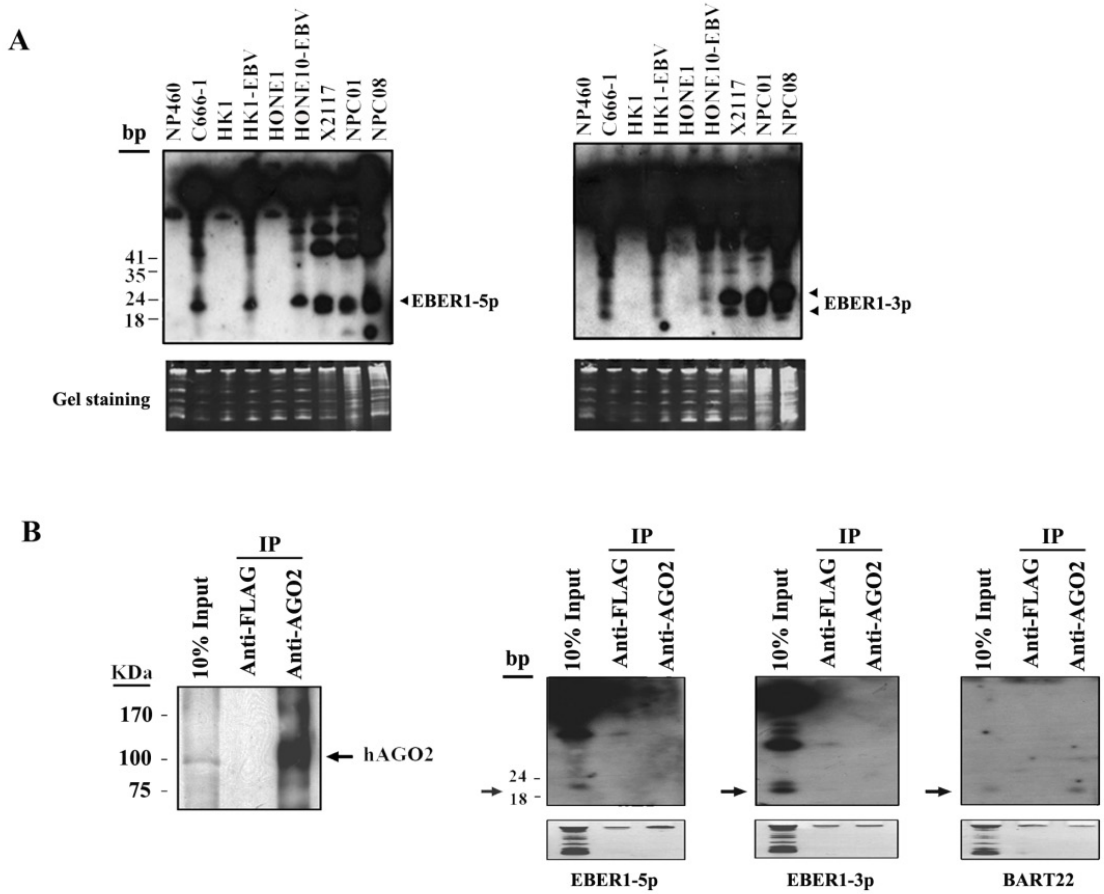


Figure S2. Absence of miR-BHRF1s expression in epithelial cancers: (**A**) Expression of BHRF1 mRNA transcripts in NPC cell lines (lanes 2–6), B-cell cell lines (lanes 7–10), NPC xenografts (lanes 11–12) and NPC biopsies (lanes 13–16) was analyzed by RT-PCR analysis. The RT-PCR products were separated on agarose gel (*left upper panel*) and transferred to a membrane for detection by Southern blot (*left middle panel*). Expression of actin in RT-PCR was used as a control (*left lower panel*). Representative Northern blot results showed the expression of ebv-miRNAs in different EBV positive samples (*right panel*); (**B**) Induction of EBV lytic cycles in C666-1 cells was confirmed by Western blot analysis (*left panel*). Expression of miR-BHRF1 and hsa-miR-16 was demonstrated by Northern blot analysis (*right panel*); (**C**) Representative Northern blot analysis on RNA samples from different EBV positive epithelial carcinomas was performed. The biopsies included NPC, Lymophoepithelial-like cholangiocarcinoma (LEL-CCA) and Lymphoepithelial-like carcinoma of the lung (LELC-Lung).

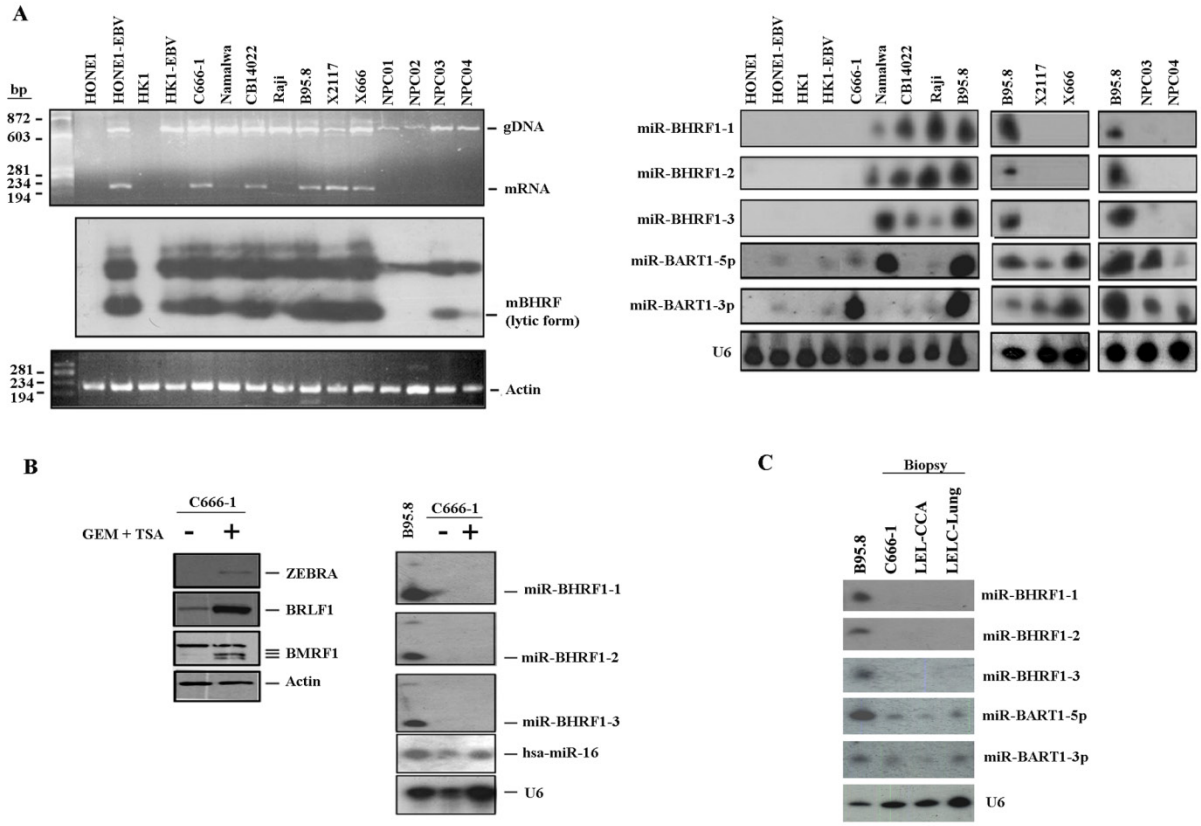
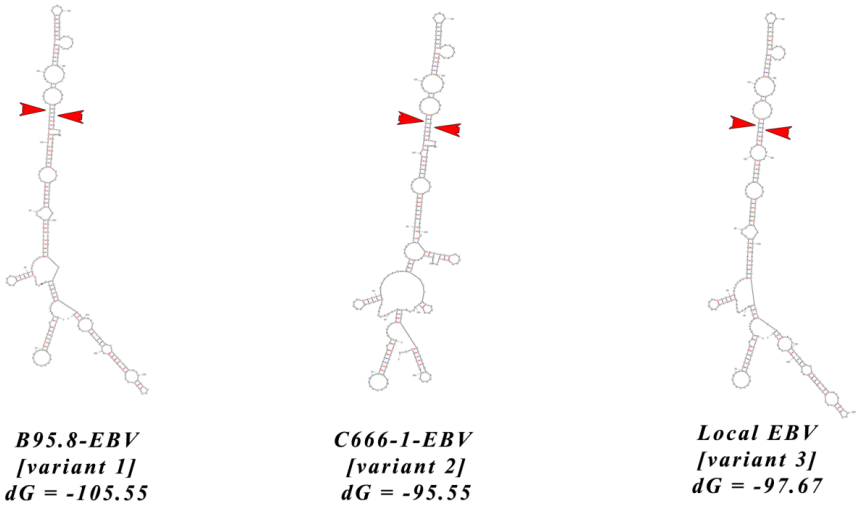


Figure S3. The predicted secondary structures of the long v-snoRNA1 containing fragment. The RNA folding of EBV fragment (AJ507799.2, 152761:153032) predicted by MFOLD are shown in B95.8-EBV (Variant 1) (*left panel*); C666-1-EBV (Variant 2) (*middle panel*) and the common EBV strain in our locality (Variant 3) (*right panel*). The folding energy (dG) with units (kcal/mol) is indicated. The location of v-snoRNA1 is indicated by sequences between red arrows.



References

1. Epstein, M.A.; Barr, Y.M.; Achong, B.G. A second virus-carrying tissue culture strain (Eb2) of lymphoblasts from burkitt's lymphoma. *Pathol. Biol.* **1964**, *12*, 1233–1234.
2. Zur Hausen, H.; Schulte-Holthausen, H. Presence of EB virus nucleic acid homology in a "virus-free" line of Burkitt tumour cells. *Nature* **1970**, *227*, 245–248.
3. Sitki-Green, D.L.; Edwards, R.H.; Covington, M.M.; Raab-Traub, N. Biology of Epstein-Barr virus during infectious mononucleosis. *J. Infect. Dis.* **2004**, *189*, 483–492.
4. Kutok, J.L.; Wang, F. Spectrum of Epstein-Barr virus-associated diseases. *Annu. Rev. Pathol.* **2006**, *1*, 375–404.
5. Babcock, G.J.; Decker, L.L.; Volk, M.; Thorley-Lawson, D.A. EBV persistence in memory B cells *in vivo*. *Immunity* **1998**, *9*, 395–404.
6. Young, L.S.; Rickinson, A.B. Epstein-Barr virus: 40 years on. *Nat. Rev. Cancer* **2004**, *4*, 757–768.
7. Gregory, C.D.; Rowe, M.; Rickinson, A.B. Different Epstein-Barr virus-B cell interactions in phenotypically distinct clones of a Burkitt's lymphoma cell line. *J. Gen. Virol.* **1990**, *71*, 1481–1495.
8. Timms, J.M.; Bell, A.; Flavell, J.R.; Murray, P.G.; Rickinson, A.B.; Traverse-Glehen, A.; Berger, F.; Delecluse, H.J. Target cells of Epstein-Barr-virus (EBV)-positive post-transplant lymphoproliferative disease: Similarities to EBV-positive Hodgkin's lymphoma. *Lancet* **2003**, *361*, 217–223.
9. Brooks, L.; Yao, Q.Y.; Rickinson, A.B.; Young, L.S. Epstein-Barr virus latent gene transcription in nasopharyngeal carcinoma cells: Coexpression of EBNA1, LMP1, and LMP2 transcripts. *J. Virol.* **1992**, *66*, 2689–2697.

10. Qu, L.; Rowe, D.T. Epstein-Barr virus latent gene expression in uncultured peripheral blood lymphocytes. *J. Virol.* **1992**, *66*, 3715–3724.
11. Shaknovich, R.; Basso, K.; Bhagat, G.; Mansukhani, M.; Hatzivassiliou, G.; Murty, V.V.; Buettner, M.; Niedobitek, G.; Alobeid, B.; Cattoretti, G. Identification of rare Epstein-Barr virus infected memory B cells and plasma cells in non-monomorphic post-transplant lymphoproliferative disorders and the signature of viral signaling. *Haematologica* **2006**, *91*, 1313–1320.
12. Kieff, E.; Johannsen, E.; Calderwood, M.A. *Epstein-Barr Virus: Latency and Transformation*; Robertson, E.S., Ed.;. Caister Academic Press: Norfolk, UK, 2010; pp. 1–24.
13. Lo, K.W.; To, K.F.; Huang, D.P. Focus on nasopharyngeal carcinoma. *Cancer Cell* **2004**, *5*, 423–428.
14. Hong Kong Cancer Registry. Available online: http://www3.ha.org.hk/cancereg/Statistics.html accessed on 30 June 2013, Hong Kong Cancer Registry 2010 (accessed on 20 August 2013).
15. Marks, J.E.; Phillips, J.L.; Menck, H.R. The national cancer data base report on the relationship of race and national origin to the histology of nasopharyngeal carcinoma. *Cancer* **1998**, *83*, 582–588.
16. Imai, S.; Nishikawa, J.; Takada, K., Cell-to-cell contact as an efficient mode of Epstein-Barr virus infection of diverse human epithelial cells. *J. Virol.* **1998**, *72*, 4371–4378.
17. Iizasa, H.; Nanbo, A.; Nishikawa, J.; Jinushi, M.; Yoshiyama, H. Epstein-Barr Virus (EBV)-associated gastric carcinoma. *Viruses* **2012**, *4*, 3420–3439.
18. Raab-Traub, N. Epstein-Barr virus in the pathogenesis of NPC. *Semin. Cancer Biol.* **2002**, *12*, 431–441.
19. Dawson, C.W.; Port, R.J.; Young, L.S. The role of the EBV-encoded latent membrane proteins LMP1 and LMP2 in the pathogenesis of nasopharyngeal carcinoma (NPC). *Semin. Cancer Biol.* **2012**, *22*, 144–153.
20. Lo, A.K.; Lo, K.W.; Ko, C.W.; Young, L.S.; Dawson, C.W. Inhibition of the LKB1-AMPK pathway by the Epstein-Barr virus-encoded LMP1 promotes proliferation and transformation of human nasopharyngeal epithelial cells. *J. Pathol.* **2013**, *230*, 336–346.
21. Lee, S.P. Nasopharyngeal carcinoma and the EBV-specific T cell response: Prospects for immunotherapy. *Semin. Cancer Biol.* **2002**, *12*, 463–471.
22. Gottschalk, S.; Heslop, H.E.; Rooney, C.M. Adoptive immunotherapy for EBV-associated malignancies. *Leuk. Lymphoma* **2005**, *46*, 1–10.
23. Lung, R.W.; Tong, J.H.; Sung, Y.M.; Leung, P.S.; Ng, D.C.; Chau, S.L.; Chan, A.W.; Ng, E.K.; Lo, K.W.; To, K.F. Modulation of LMP2A expression by a newly identified Epstein-Barr virus-encoded microRNA miR-BART22. *Neoplasia* **2009**, *11*, 1174–1184.
24. Chen, S.J.; Chen, G.H.; Chen, Y.H.; Liu, C.Y.; Chang, K.P.; Chang, Y.S.; Chen, H.C. Characterization of Epstein-Barr virus miRNAome in nasopharyngeal carcinoma by deep sequencing. *PLoS One* **2010**, *5*, e12745.
25. Burke, A.P.; Yen, T.S.; Shekitka, K.M.; Sobin, L.H. Lymphoepithelial carcinoma of the stomach with Epstein-Barr virus demonstrated by polymerase chain reaction. *Mod. Pathol.* **1990**, *3*, 377–380.
26. Shibata, D.; Tokunaga, M.; Uemura, Y.; Sato, E.; Tanaka, S.; Weiss, L.M. Association of Epstein-Barr virus with undifferentiated gastric carcinomas with intense lymphoid infiltration. Lymphoepithelioma-like carcinoma. *Am. J. Pathol.* **1991**, *139*, 469–474.

27. Osato, T.; Imai, S. Epstein-Barr virus and gastric carcinoma. *Semin. Cancer Biol.* **1996**, *7*, 175–182.
28. Shibata, D.; Weiss, L.M. Epstein-Barr virus-associated gastric adenocarcinoma. *Am. J. Pathol.* **1992**, *140*, 769–774.
29. Takada, K. Epstein-Barr virus and gastric carcinoma. *Mol. Pathol.* **2000**, *53*, 255–261.
30. Lee, J.H.; Kim, S.H.; Han, S.H.; An, J.S.; Lee, E.S.; Kim, Y.S. Clinicopathological and molecular characteristics of Epstein-Barr virus-associated gastric carcinoma: A meta-analysis. *J. Gastroenterol. Hepatol.* **2009**, *24*, 354–365.
31. Murphy, G.; Pfeiffer, R.; Camargo, M.C.; Rabkin, C.S. Meta-analysis shows that prevalence of Epstein-Barr virus-positive gastric cancer differs based on sex and anatomic location. *Gastroenterology* **2009**, *137*, 824–833.
32. Chen, J.N.; He, D.; Tang, F.; Shao, C.K. Epstein-Barr virus-associated gastric carcinoma: A newly defined entity. *J. Clin. Gastroenterol.* **2012**, *46*, 262–271.
33. Imai, S.; Koizumi, S.; Sugiura, M.; Tokunaga, M.; Uemura, Y.; Yamamoto, N.; Tanaka, S.; Sato, E.; Osato, T. Gastric carcinoma: Monoclonal epithelial malignant cells expressing Epstein-Barr virus latent infection protein. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 9131–9135.
34. Luo, B.; Wang, Y.; Wang, X.F.; Liang, H.; Yan, L.P.; Huang, B.H.; Zhao, P. Expression of Epstein-Barr virus genes in EBV-associated gastric carcinomas. *World J. Gastroenterol.* **2005**, *11*, 629–633.
35. Zhao, J.; Liang, Q.; Cheung, K.F.; Kang, W.; Lung, R.W.; Tong, J.H.; To, K.F.; Sung, J.J.; Yu, J. Genome-wide identification of Epstein-Barr virus-driven promoter methylation profiles of human genes in gastric cancer cells. *Cancer* **2013**, *119*, 304–312.
36. Zhao, J.; Jin, H.; Cheung, K.F.; Tong, J.H.; Zhang, S.; Go, M.Y.; Tian, L.; Kang, W.; Leung, P.P.; Zeng, Z.; *et al.* Zinc finger E-box binding factor 1 plays a central role in regulating Epstein-Barr virus (EBV) latent-lytic switch and acts as a therapeutic target in EBV-associated gastric cancer. *Cancer* **2012**, *118*, 924–936.
37. Zhao, J.; Liang, Q.; Cheung, K.F.; Kang, W.; Dong, Y.; Lung, R.W.; Tong, J.H.; To, K.F.; Sung, J.J.; Yu, J. Somatostatin receptor 1, a novel EBV-associated CpG hypermethylated gene, contributes to the pathogenesis of EBV-associated gastric cancer. *Br. J. Cancer* **2013**, *108*, 2557–2564.
38. Kim do, N.; Chae, H.S.; Oh, S.T.; Kang, J.H.; Park, C.H.; Park, W.S.; Takada, K.; Lee, J.M.; Lee, W.K.; Lee, S.K. Expression of viral microRNAs in Epstein-Barr virus-associated gastric carcinoma. *J. Virol.* **2007**, *81*, 1033–1036.
39. Marquitz, A.R.; Mathur, A.; Shair, K.H.; Raab-Traub, N. Infection of Epstein-Barr virus in a gastric carcinoma cell line induces anchorage independence and global changes in gene expression. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9593–9598.
40. Iezzoni, J.C.; Gaffey, M.J.; Weiss, L.M. The role of Epstein-Barr virus in lymphoepithelioma-like carcinomas. *Am. J. Clin. Pathol.* **1995**, *103*, 308–315.
41. Xiao, P.; Shi, H.; Zhang, H.; Meng, F.; Peng, J.; Ke, Z.; Wang, K.; Liu, Y.; Han, A. Epstein-Barr virus-associated intrahepatic cholangiocarcinoma bearing an intense lymphoplasmacytic infiltration. *J. Clin. Pathol.* **2012**, *65*, 570–573.
42. Hsu, J.L.; Glaser, S.L. Epstein-Barr virus-associated malignancies: Epidemiologic patterns and etiologic implications. *Crit. Rev. Oncol. Hematol.* **2000**, *34*, 27–53.
43. Hippocrate, A.; Oussaief, L.; Joab, I. Possible role of EBV in breast cancer and other unusually EBV-associated cancers. *Cancer Lett.* **2011**, *305*, 144–149.

44. Begin, L.R.; Eskandari, J.; Joncas, J.; Panasci, L. Epstein-Barr virus related lymphoepithelioma-like carcinoma of lung. *J. Surg. Oncol.* **1987**, *36*, 280–283.
45. Han, A.J.; Xiong, M.; Zong, Y.S. Association of Epstein-Barr virus with lymphoepithelioma-like carcinoma of the lung in southern China. *Am. J. Clin. Pathol.* **2000**, *114*, 220–226.
46. Lerner, M.R.; Andrews, N.C.; Miller, G.; Steitz, J.A. Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 805–809.
47. Takada, K. Role of EBER and BARF1 in nasopharyngeal carcinoma (NPC) tumorigenesis. *Semin. Cancer Biol.* **2012**, *22*, 162–165.
48. Sample, J.; Kieff, E. Transcription of the Epstein-Barr virus genome during latency in growth-transformed lymphocytes. *J. Virol.* **1990**, *64*, 1667–1674.
49. Howe, J.G.; Shu, M.D. Epstein-Barr virus small RNA (EBER) genes: Unique transcription units that combine RNA polymerase II and III promoter elements. *Cell* **1989**, *57*, 825–834.
50. Schwemmle, M.; Clemens, M.J.; Hilse, K.; Pfeifer, K.; Troster, H.; Muller, W.E.; Bachmann, M. Localization of Epstein-Barr virus-encoded RNAs EBER-1 and EBER-2 in interphase and mitotic Burkitt lymphoma cells. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10292–10296.
51. Gourzones, C.; Busson, P.; Raab-Traub, N. *Nasopharyngeal Carcinoma: Keys for Translational Medicine and Biology*; Busson, P., Ed.; Landes Bioscience and Springer Science: Villejuif, France, 2012; pp. 42–60.
52. Fok, V.; Mitton-Fry, R.M.; Grech, A.; Steitz, J.A. Multiple domains of EBER 1, an Epstein-Barr virus noncoding RNA, recruit human ribosomal protein L22. *RNA* **2006**, *12*, 872–882.
53. Toczyski, D.P.; Steitz, J.A. EAP, a highly conserved cellular protein associated with Epstein-Barr virus small RNAs (EBERs). *EMBO J.* **1991**, *10*, 459–466.
54. Toczyski, D.P.; Matera, A.G.; Ward, D.C.; Steitz, J.A. The Epstein-Barr virus (EBV) small RNA EBER1 binds and relocalizes ribosomal protein L22 in EBV-infected human B lymphocytes. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3463–3467.
55. Clarke, P.A.; Schwemmle, M.; Schickinger, J.; Hilse, K.; Clemens, M.J. Binding of Epstein-Barr virus small RNA EBER-1 to the double-stranded RNA-activated protein kinase DAI. *Nucleic Acids Res.* **1991**, *19*, 243–248.
56. Samanta, M.; Iwakiri, D.; Kanda, T.; Imaizumi, T.; Takada, K. EB virus-encoded RNAs are recognized by RIG-I and activate signaling to induce type I IFN. *EMBO J.* **2006**, *25*, 4207–4214.
57. Wong, H.L.; Wang, X.; Chang, R.C.; Jin, D.Y.; Feng, H.; Wang, Q.; Lo, K.W.; Huang, D.P.; Yuen, P.W.; Takada, K.; *et al.* Stable expression of EBERs in immortalized nasopharyngeal epithelial cells confers resistance to apoptotic stress. *Mol. Carcinog.* **2005**, *44*, 92–101.
58. Iwakiri, D.; Sheen, T.S.; Chen, J.Y.; Huang, D.P.; Takada, K. Epstein-Barr virus-encoded small RNA induces insulin-like growth factor 1 and supports growth of nasopharyngeal carcinoma-derived cell lines. *Oncogene* **2005**, *24*, 1767–1773.
59. Iwakiri, D.; Eizuru, Y.; Tokunaga, M.; Takada, K. Autocrine growth of Epstein-Barr virus-positive gastric carcinoma cells mediated by an Epstein-Barr virus-encoded small RNA. *Cancer Res.* **2003**, *63*, 7062–7067.

60. Rosa, M.D.; Gottlieb, E.; Lerner, M.R.; Steitz, J.A. Striking similarities are exhibited by two small Epstein-Barr virus-encoded ribonucleic acids and the adenovirus-associated ribonucleic acids VAI and VAII. *Mol. Cell Biol.* **1981**, *1*, 785–796.
61. Glickman, J.N.; Howe, J.G.; Steitz, J.A. Structural analyses of EBER1 and EBER2 ribonucleoprotein particles present in Epstein-Barr virus-infected cells. *J. Virol.* **1988**, *62*, 902–911.
62. Bhat, R.A.; Thimmappaya, B. Two small RNAs encoded by Epstein-Barr virus can functionally substitute for the virus-associated RNAs in the lytic growth of adenovirus 5. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 4789–4793.
63. Wang, Y.; Xue, S.A.; Hallden, G.; Francis, J.; Yuan, M.; Griffin, B.E.; Lemoine, N.R. Virus-associated RNA I-deleted adenovirus, a potential oncolytic agent targeting EBV-associated tumors. *Cancer Res.* **2005**, *65*, 1523–1531.
64. Aparicio, O.; Razquin, N.; Zaratiegui, M.; Narvaiza, I.; Fortes, P. Adenovirus virus-associated RNA is processed to functional interfering RNAs involved in virus production. *J. Virol.* **2006**, *80*, 1376–1384.
65. Sano, M.; Kato, Y.; Taira, K. Sequence-specific interference by small RNAs derived from adenovirus VAI RNA. *FEBS Lett.* **2006**, *580*, 1553–1564.
66. Han, J.; Lee, Y.; Yeom, K.H.; Nam, J.W.; Heo, I.; Rhee, J.K.; Sohn, S.Y.; Cho, Y.; Zhang, B.T.; Kim, V.N. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **2006**, *125*, 887–901.
67. Fok, V.; Friend, K.; Steitz, J.A. Epstein-Barr virus noncoding RNAs are confined to the nucleus, whereas their partner, the human La protein, undergoes nucleocytoplasmic shuttling. *J. Cell Biol.* **2006**, *173*, 319–325.
68. Gilligan, K.; Sato, H.; Rajadurai, P.; Busson, P.; Young, L.; Rickinson, A.; Tursz, T.; Raab-Traub, N. Novel transcription from the Epstein-Barr virus terminal EcoRI fragment, DIJhet, in a nasopharyngeal carcinoma. *J. Virol.* **1990**, *64*, 4948–4956.
69. Gilligan, K.J.; Rajadurai, P.; Lin, J.C.; Busson, P.; Abdel-Hamid, M.; Prasad, U.; Tursz, T.; Raab-Traub, N. Expression of the Epstein-Barr virus BamHI A fragment in nasopharyngeal carcinoma: Evidence for a viral protein expressed *in vivo*. *J. Virol.* **1991**, *65*, 6252–6259.
70. Robertson, E.S.; Tomkinson, B.; Kieff, E. An Epstein-Barr virus with a 58-kilobase-pair deletion that includes BARF0 transforms B lymphocytes *in vitro*. *J. Virol.* **1994**, *68*, 1449–1458.
71. Al-Mozaini, M.; Bodelon, G.; Karstegl, C.E.; Jin, B.; Al-Ahdal, M.; Farrell, P.J. Epstein-Barr virus BART gene expression. *J. Gen. Virol.* **2009**, *90*, 307–316.
72. Smith, P.R.; de Jesus, O.; Turner, D.; Hollyoake, M.; Karstegl, C.E.; Griffin, B.E.; Karran, L.; Wang, Y.; Hayward, S.D.; Farrell, P.J. Structure and coding content of CST (BART) family RNAs of Epstein-Barr virus. *J. Virol.* **2000**, *74*, 3082–3092.
73. Pfeffer, S.; Zavolan, M.; Grasser, F.A.; Chien, M.; Russo, J.J.; Ju, J.; John, B.; Enright, A.J.; Marks, D.; Sander, C.; *et al.* Identification of virus-encoded microRNAs. *Science* **2004**, *304*, 734–736.
74. Cai, X.; Schafer, A.; Lu, S.; Bilello, J.P.; Desrosiers, R.C.; Edwards, R.; Raab-Traub, N.; Cullen, B.R. Epstein-Barr virus microRNAs are evolutionarily conserved and differentially expressed. *PLoS Pathog.* **2006**, *2*, e23.
75. Grundhoff, A.; Sullivan, C.S.; Ganem, D. A combined computational and microarray-based approach identifies novel microRNAs encoded by human gamma-herpesviruses. *RNA* **2006**, *12*, 733–750.

76. Hutzinger, R.; Feederle, R.; Mrazek, J.; Schiefermeier, N.; Balwierz, P.J.; Zavolan, M.; Polacek, N.; Delecluse, H.J.; Huttenhofer, A. Expression and processing of a small nucleolar RNA from the Epstein-Barr virus genome. *PLoS Pathog.* **2009**, *5*, e1000547.
77. Hammond, S.M.; Boettcher, S.; Caudy, A.A.; Kobayashi, R.; Hannon, G.J. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **2001**, *293*, 1146–1150.
78. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297.
79. Carthew, R.W.; Sontheimer, E.J. Origins and mechanisms of miRNAs and siRNAs. *Cell* **2009**, *136*, 642–655.
80. Zhu, J.Y.; Pfuhl, T.; Motsch, N.; Barth, S.; Nicholls, J.; Grasser, F.; Meister, G. Identification of novel Epstein-Barr virus microRNA genes from nasopharyngeal carcinomas. *J. Virol.* **2009**, *83*, 3333–3341.
81. Matera, A.G.; Terns, R.M.; Terns, M.P. Non-coding RNAs: Lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 209–220.
82. Ender, C.; Krek, A.; Friedlander, M.R.; Beitzinger, M.; Weinmann, L.; Chen, W.; Pfeffer, S.; Rajewsky, N.; Meister, G. A human snoRNA with microRNA-like functions. *Mol. Cell* **2008**, *32*, 519–528.
83. Saraiya, A.A.; Wang, C.C. snoRNA, a novel precursor of microRNA in Giardia lamblia. *PLoS Pathog.* **2008**, *4*, e1000224.
84. Xing, L.; Kieff, E. Epstein-Barr virus BHRF1 micro- and stable RNAs during latency III and after induction of replication. *J. Virol.* **2007**, *81*, 9967–9975.
85. Imig, J.; Motsch, N.; Zhu, J.Y.; Barth, S.; Okoniewski, M.; Reineke, T.; Tinguely, M.; Faggioni, A.; Trivedi, P.; Meister, G.; *et al.* microRNA profiling in Epstein-Barr virus-associated B-cell lymphoma. *Nucleic Acids Res.* **2011**, *39*, 1880–1893.
86. Amoroso, R.; Fitzsimmons, L.; Thomas, W.A.; Kelly, G.L.; Rowe, M.; Bell, A.I. Quantitative studies of Epstein-Barr virus-encoded microRNAs provide novel insights into their regulation. *J. Virol.* **2011**, *85*, 996–1010.
87. Qiu, J.; Cosmopoulos, K.; Pegtel, M.; Hopmans, E.; Murray, P.; Middeldorp, J.; Shapiro, M.; Thorley-Lawson, D.A. A novel persistence associated EBV miRNA expression profile is disrupted in neoplasia. *PLoS Pathog.* **2011**, *7*, e1002193.
88. Cosmopoulos, K.; Pegtel, M.; Hawkins, J.; Moffett, H.; Novina, C.; Middeldorp, J.; Thorley-Lawson, D.A. Comprehensive profiling of Epstein-Barr virus microRNAs in nasopharyngeal carcinoma. *J. Virol.* **2009**, *83*, 2357–2367.
89. Chatterjee, S.; Fasler, M.; Bussing, I.; Grosshans, H. Target-mediated protection of endogenous microRNAs in C. elegans. *Dev. Cell* **2011**, *20*, 388–396.
90. Chen, H.; Huang, J.; Wu, F.Y.; Liao, G.; Hutt-Fletcher, L.; Hayward, S.D. Regulation of expression of the Epstein-Barr virus BamHI-A rightward transcripts. *J. Virol.* **2005**, *79*, 1724–1733.
91. Gourzones, C.; Jimenez, A.S.; Busson, P. Profiling of Epstein-Barr virus-encoded microRNAs in nasopharyngeal carcinoma reveals potential biomarkers and oncomirs. *Cancer* **2012**, 118, dio:10.1002/cncr.26514.
92. Wong, A.M.; Kong, K.L.; Tsang, J.W.; Kwong, D.L.; Guan, X.Y. Profiling of Epstein-Barr virus-encoded microRNAs in nasopharyngeal carcinoma reveals potential biomarkers and oncomirs. *Cancer* **2012**, *118*, 698–710.

93. Choy, E.Y.; Siu, K.L.; Kok, K.H.; Lung, R.W.; Tsang, C.M.; To, K.F.; Kwong, D.L.; Tsao, S.W.; Jin, D.Y. An Epstein-Barr virus-encoded microRNA targets PUMA to promote host cell survival. *J. Exp. Med.* **2008**, *205*, 2551–2560.
94. Lung, R.W.; Wang, X.; Tong, J.H.; Chau, S.L.; Lau, K.M.; Cheng, S.H.; Woo, J.K.; Woo, J.; Leung, P.C.; Ng, M.H.; *et al.* A single nucleotide polymorphism in microRNA-146a is associated with the risk for nasopharyngeal carcinoma. *Mol. Carcinog.* **2012**, doi:10.1002/mc.21937.
95. Cazalla, D.; Xie, M.; Steitz, J.A. A primate herpesvirus uses the integrator complex to generate viral microRNAs. *Mol. Cell* **2011**, *43*, 982–992.
96. Chi, S.W.; Zang, J.B.; Mele, A.; Darnell, R.B. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* **2009**, *460*, 479–486.
97. Ascano, M.; Hafner, M.; Cekan, P.; Gerstberger, S.; Tuschl, T. Identification of RNA-protein interaction networks using PAR-CLIP. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 159–177.
98. Hafner, M.; Lianoglou, S.; Tuschl, T.; Betel, D. Genome-wide identification of miRNA targets by PAR-CLIP. *Methods* **2012**, *58*, 94–105.
99. Gottwein, E.; Corcoran, D.L.; Mukherjee, N.; Skalsky, R.L.; Hafner, M.; Nusbaum, J.D.; Shamulailatpam, P.; Love, C.L.; Dave, S.S.; Tuschl, T.; *et al.* Viral microRNA targetome of KSHV-infected primary effusion lymphoma cell lines. *Cell Host Microbe* **2011**, *10*, 515–526.
100. Skalsky, R.L.; Corcoran, D.L.; Gottwein, E.; Frank, C.L.; Kang, D.; Hafner, M.; Nusbaum, J.D.; Feederle, R.; Delecluse, H.J.; Luftig, M.A.; *et al.* The viral and cellular microRNA targetome in lymphoblastoid cell lines. *PLoS Pathog.* **2012**, *8*, e1002484.
101. Riley, K.J.; Rabinowitz, G.S.; Yario, T.A.; Luna, J.M.; Darnell, R.B.; Steitz, J.A. EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency. *EMBO J.* **2012**, *31*, 2207–2221.
102. Barth, S.; Pfuhl, T.; Mamiani, A.; Ehses, C.; Roemer, K.; Kremmer, E.; Jaker, C.; Hock, J.; Meister, G.; Grasser, F.A. Epstein-Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5. *Nucleic Acids Res.* **2008**, *36*, 666–675.
103. Lo, A.K.; To, K.F.; Lo, K.W.; Lung, R.W.; Hui, J.W.; Liao, G.; Hayward, S.D. Modulation of LMP1 protein expression by EBV-encoded microRNAs. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16164–16169.
104. Hislop, A.D.; Taylor, G.S.; Sauce, D.; Rickinson, A.B. Cellular responses to viral infection in humans: Lessons from Epstein-Barr virus. *Annu. Rev. Immunol.* **2007**, *25*, 587–617.
105. Straathof, K.C.; Leen, A.M.; Buza, E.L.; Taylor, G.; Huls, M.H.; Heslop, H.E.; Rooney, C.M.; Bollard, C.M. Characterization of latent membrane protein 2 specificity in CTL lines from patients with EBV-positive nasopharyngeal carcinoma and lymphoma. *J. Immunol.* **2005**, *175*, 4137–4147.
106. Eliopoulos, A.G.; Dawson, C.W.; Mosialos, G.; Floettmann, J.E.; Rowe, M.; Armitage, R.J.; Dawson, J.; Zapata, J.M.; Kerr, D.J.; Wakelam, M.J.; *et al.* CD40-induced growth inhibition in epithelial cells is mimicked by Epstein-Barr virus-encoded LMP1: Involvement of TRAF3 as a common mediator. *Oncogene* **1996**, *13*, 2243–2254.
107. Dirmeier, U.; Hoffmann, R.; Kilger, E.; Schultheiss, U.; Briseno, C.; Gires, O.; Kieser, A.; Eick, D.; Sugden, B.; Hammerschmidt, W. Latent membrane protein 1 of Epstein-Barr virus coordinately regulates proliferation with control of apoptosis. *Oncogene* **2005**, *24*, 1711–1717.

108. Ramakrishnan, R.; Donahue, H.; Garcia, D.; Tan, J.; Shimizu, N.; Rice, A.P.; Ling, P.D. Epstein-Barr virus BART9 miRNA modulates LMP1 levels and affects growth rate of nasal NK T cell lymphomas. *PLoS One* **2011**, *6*, e27271.
109. Lomonosova, E.; Chinnadurai, G. BH3-only proteins in apoptosis and beyond: An overview. *Oncogene* **2008**, *27*, S2–S19.
110. Marquitz, A.R.; Mathur, A.; Nam, C.S.; Raab-Traub, N. The Epstein-Barr virus BART microRNAs target the pro-apoptotic protein Bim. *Virology* **2011**, *412*, 392–400.
111. Lin, T.C.; Liu, T.Y.; Hsu, S.M.; Lin, C.W. Epstein-Barr virus-encoded miR-BART20-5p inhibits T-bet translation with secondary suppression of p53 in invasive nasal NK/T-cell lymphoma. *Am. J. Pathol.* **2013**, *182*, 1865–1875.
112. Dolken, L.; Malterer, G.; Erhard, F.; Kothe, S.; Friedel, C.C.; Suffert, G.; Marcinowski, L.; Motsch, N.; Barth, S.; Beitzinger, M.; *et al.* Systematic analysis of viral and cellular microRNA targets in cells latently infected with human gamma-herpesviruses by RISC immunoprecipitation assay. *Cell Host Microbe* **2010**, *7*, 324–334.
113. Bellot, G.; Cartron, P.F.; Er, E.; Oliver, L.; Juin, P.; Armstrong, L.C.; Bornstein, P.; Mihara, K.; Manon, S.; Vallette, F.M. TOM22, a core component of the mitochondria outer membrane protein translocation pore, is a mitochondrial receptor for the proapoptotic protein Bax. *Cell Death Differ.* **2007**, *14*, 785–794.
114. Vereide, D.T.; Seto, E.; Chiu, Y.F.; Hayes, M.; Tagawa, T.; Grundhoff, A.; Hammerschmidt, W.; Sugden, B. Epstein-Barr virus maintains lymphomas via its miRNAs. *Oncogene* **2013**, doi:10.1038/onc.2013.71.
115. Ding, Y.; Xi, Y.; Chen, T.; Wang, J.Y.; Tao, D.L.; Wu, Z.L.; Li, Y.P.; Li, C.; Zeng, R.; Li, L. Caprin-2 enhances canonical Wnt signaling through regulating LRP5/6 phosphorylation. *J. Cell Biol.* **2008**, *182*, 865–872.
116. Aerbajinai, W.; Lee, Y.T.; Wojda, U.; Barr, V.A.; Miller, J.L. Cloning and characterization of a gene expressed during terminal differentiation that encodes a novel inhibitor of growth. *J. Biol. Chem.* **2004**, *279*, 1916–1921.
117. Choi, H.; Lee, H.; Kim, S.R.; Gho, Y.S.; Lee, S.K. EBV encoded miR-BART15-3p promotes cell apoptosis partially by targeting BRUCE. *J. Virol.* **2013**, doi:10.1128/JVI.03159-12.
118. Chen, Z.; Naito, M.; Hori, S.; Mashima, T.; Yamori, T.; Tsuruo, T. A human IAP-family gene, apollon, expressed in human brain cancer cells. *Biochem. Biophys. Res. Commun.* **1999**, *264*, 847–854.
119. Verhagen, A.M.; Coulson, E.J.; Vaux, D.L. Inhibitor of apoptosis proteins and their relatives: IAPs and other BIRPs. *Genome Biol.* **2001**, *2*, REVIEWS3009.1–REVIEWS3009.10.
120. Haneklaus, M.; Gerlic, M.; Kurowska-Stolarska, M.; Rainey, A.A.; Pich, D.; McInnes, I.B.; Hammerschmidt, W.; O'Neill, L.A.; Masters, S.L. Cutting edge: miR-223 and EBV miR-BART15 regulate the NLRP3 inflammasome and IL-1beta production. *J. Immunol.* **2012**, *189*, 3795–3799.
121. Kang, W.; Tong, J.H.; Chan, A.W.; Lee, T.L.; Lung, R.W.; Leung, P.P.; So, K.K.; Wu, K.; Fan, D.; Yu, J.; *et al.* Yes-associated protein 1 exhibits oncogenic property in gastric cancer and its nuclear accumulation associates with poor prognosis. *Clin. Cancer Res.* **2011**, *17*, 2130–2139.
122. Wong, Q.W.; Lung, R.W.; Law, P.T.; Lai, P.B.; Chan, K.Y.; To, K.F.; Wong, N. MicroRNA-223 is commonly repressed in hepatocellular carcinoma and potentiates expression of Stathmin1. *Gastroenterology* **2008**, *135*, 257–269.

123. Pegtel, D.M.; Cosmopoulos, K.; Thorley-Lawson, D.A.; van Eijndhoven, M.A.; Hopmans, E.S.; Lindenberg, J.L.; de Gruijl, T.D.; Wurdinger, T.; Middeldorp, J.M. Functional delivery of viral miRNAs via exosomes. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 6328–6333.
124. Meckes, D.G., Jr.; Shair, K.H.; Marquitz, A.R.; Kung, C.P.; Edwards, R.H.; Raab-Traub, N. Human tumor virus utilizes exosomes for intercellular communication. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 20370–20375.
125. Nachmani, D.; Stern-Ginossar, N.; Sarid, R.; Mandelboim, O. Diverse herpesvirus microRNAs target the stress-induced immune ligand MICB to escape recognition by natural killer cells. *Cell Host Microbe* **2009**, *5*, 376–385.
126. Yang, I.V.; Wade, C.M.; Kang, H.M.; Alper, S.; Rutledge, H.; Lackford, B.; Eskin, E.; Daly, M.J.; Schwartz, D.A. Identification of novel genes that mediate innate immunity using inbred mice. *Genetics* **2009**, *183*, 1535–1544.
127. Ullman, K.S.; Powers, M.A.; Forbes, D.J. Nuclear export receptors: From importin to exportin. *Cell* **1997**, *90*, 967–970.
128. Lei, T.; Yuen, K.S.; Xu, R.; Tsao, S.W.; Chen, H.; Li, M.; Kok, K.H.; Jin, D.Y. Targeting of DICE1 tumor suppressor by Epstein-Barr virus-encoded miR-BART3* microRNA in nasopharyngeal carcinoma. *Int. J. Cancer* **2013**, *133*, 79–87.
129. Ropke, A.; Buhtz, P.; Bohm, M.; Seger, J.; Wieland, I.; Allhoff, E.P.; Wieacker, P.F. Promoter CpG hypermethylation and downregulation of DICE1 expression in prostate cancer. *Oncogene* **2005**, *24*, 6667–6675.
130. Li, W.J.; Hu, N.; Su, H.; Wang, C.; Goldstein, A.M.; Wang, Y.; Emmert-Buck, M.R.; Roth, M.J.; Guo, W.J.; Taylor, P.R. Allelic loss on chromosome 13q14 and mutation in deleted in cancer 1 gene in esophageal squamous cell carcinoma. *Oncogene* **2003**, *22*, 314–318.
131. Wieland, I.; Arden, K.C.; Michels, D.; Klein-Hitpass, L.; Bohm, M.; Viars, C.S.; Weidle, U.H. Isolation of DICE1: A gene frequently affected by LOH and downregulated in lung carcinomas. *Oncogene* **1999**, *18*, 4530–4537.
132. Godshalk, S.E.; Bhaduri-McIntosh, S.; Slack, F.J. Epstein-Barr virus-mediated dysregulation of human microRNA expression. *Cell Cycle* **2008**, *7*, 3595–3600.
133. Iizasa, H.; Wulff, B.E.; Alla, N.R.; Maragkakis, M.; Megraw, M.; Hatzigeorgiou, A.; Iwakiri, D.; Takada, K.; Wiedmer, A.; Showe, L.; *et al.* Editing of Epstein-Barr virus-encoded BART6 microRNAs controls their dicer targeting and consequently affects viral latency. *J. Biol. Chem.* **2010**, *285*, 33358–33370.
134. ClinicalTrials.gov. Available online: http://clinicaltrials.gov/show/NCT01200420 (accessed on 20 August 2013).
135. Janssen, H.L.; Reesink, H.W.; Lawitz, E.J.; Zeuzem, S.; Rodriguez-Torres, M.; Patel, K.; van der Meer, A.J.; Patick, A.K.; Chen, A.; Zhou, Y.; *et al.* Treatment of HCV infection by targeting microRNA. *N. Engl. J. Med.* **2013**, *368*, 1685–1694.
136. Muthiah, M.; Park, I.K.; Cho, C.S. Nanoparticle-mediated delivery of therapeutic genes: Focus on miRNA therapeutics. *Expert. Opin. Drug. Deliv.* **2013**, doi:10.1517/17425247.2013.798640.
137. Garzon, R.; Marcucci, G.; Croce, C.M. Targeting microRNAs in cancer: Rationale, strategies and challenges. *Nat. Rev. Drug Discov.* **2010**, *9*, 775–789.

138. Lo, A.K.; Dawson, C.W.; Jin, D.Y.; Lo, K.W. The pathological roles of BART miRNAs in nasopharyngeal carcinoma. *J. Pathol.* **2012**, *227*, 392–403.
139. Babu, S.G.; Ponia, S.S.; Kumar, D.; Saxena, S. Cellular oncomiR orthologue in EBV oncogenesis. *Comput. Biol. Med.* **2011**, *41*, 891–898.
140. Keryer-Bibens, C.; Pioche-Durieu, C.; Villemant, C.; Souquere, S.; Nishi, N.; Hirashima, M.; Middeldorp, J.; Busson, P. Exosomes released by EBV-infected nasopharyngeal carcinoma cells convey the viral latent membrane protein 1 and the immunomodulatory protein galectin 9. *BMC Cancer* **2006**, *6*, 283.
141. Gourzones, C.; Gelin, A.; Bombik, I.; Klibi, J.; Verillaud, B.; Guigay, J.; Lang, P.; Temam, S.; Schneider, V.; Amiel, C.; *et al.* Extra-cellular release and blood diffusion of BART viral micro-RNAs produced by EBV-infected nasopharyngeal carcinoma cells. *Virol. J.* **2010**, *7*, 271.
142. Rechavi, O.; Erlich, Y.; Amram, H.; Flomenblit, L.; Karginov, F.V.; Goldstein, I.; Hannon, G.J.; Kloog, Y. Cell contact-dependent acquisition of cellular and viral nonautonomously encoded small RNAs. *Genes Dev.* **2009**, *23*, 1971–1979.
143. Verweij, F.J.; van Eijndhoven, M.A.; Hopmans, E.S.; Vendrig, T.; Wurdinger, T.; Cahir-McFarland, E.; Kieff, E.; Geerts, D.; van der Kant, R.; Neefjes, J.; *et al.* LMP1 association with CD63 in endosomes and secretion via exosomes limits constitutive NF-kappaB activation. *EMBO J.* **2011**, *30*, 2115–2129.
144. Pegtel, D.M.; van de Garde, M.D.; Middeldorp, J.M. Viral miRNAs exploiting the endosomal-exosomal pathway for intercellular cross-talk and immune evasion. *Biochim. Biophys. Acta* **2011**, *1809*, 715–721.
145. Gourzones, C.; Ferrand, F.R.; Amiel, C.; Verillaud, B.; Barat, A.; Guerin, M.; Gattolliat, C.H.; Gelin, A.; Klibi, J.; Chaaben, A.B.; *et al.* Consistent high concentration of the viral microRNA BART17 in plasma samples from nasopharyngeal carcinoma patients-evidence of non-exosomal transport. *Virol. J.* **2013**, *10*, 119.
146. Bala, S.; Petrasek, J.; Mundkur, S.; Catalano, D.; Levin, I.; Ward, J.; Alao, H.; Kodys, K.; Szabo, G. Circulating microRNAs in exosomes indicate hepatocyte injury and inflammation in alcoholic, drug-induced, and inflammatory liver diseases. *Hepatology* **2012**, *56*, 1946–1957.
147. Arroyo, J.D.; Chevillet, J.R.; Kroh, E.M.; Ruf, I.K.; Pritchard, C.C.; Gibson, D.F.; Mitchell, P.S.; Bennett, C.F.; Pogosova-Agadjanyan, E.L.; Stirewalt, D.L.; *et al.* Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 5003–5008.
148. Chan, J.Y.; Gao, W.; Ho, W.K.; Wei, W.I.; Wong, T.S. Overexpression of Epstein-Barr virus-encoded microRNA-BART7 in undifferentiated nasopharyngeal carcinoma. *Anticancer Res.* **2012**, *32*, 3201–3210.
149. Bustin, S.A. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J. Mol. Endocrinol.* **2000**, *25*, 169–193.
150. Oudejans, J.J.; van den Brule, A.J.; Jiwa, N.M.; de Bruin, P.C.; Ossenkoppele, G.J.; van der Valk, P.; Walboomers, J.M.; Meijer, C.J. BHRF1, the Epstein-Barr virus (EBV) homologue of the BCL-2 protooncogene, is transcribed in EBV-associated B-cell lymphomas and in reactive lymphocytes. *Blood* **1995**, *86*, 1893–1902.

Reprinted from *IJMS*. Cite as: Calore, F.; Lovat, F.; Garofalo, M. Non-Coding RNAs and Cancer. *Int. J. Mol. Sci.* **2013**, *14*, 17085-17110.

Review

Non-Coding RNAs and Cancer

Federica Calore †, Francesca Lovat † and Michela Garofalo *

Department of Molecular Virology, Immunology and Medical Genetics,
Comprehensive Cancer Center, Ohio State University, Columbus, OH 43210, USA;
E-Mails: federica.calore@osumc.edu (F.C.); francesca.lovat@osumc.edu (F.L.)

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: michela.garofalo@osumc.edu; Tel.: +1-614-688-8056; Fax: +1-614-292-4097.

Received: 5 July 2013; in revised form: 1 August 2013 / Accepted: 8 August 2013 / Published: 19 August 2013

Abstract: The discovery of the biological relevance of non-coding RNA (ncRNAs) molecules represents one of the most significant advances in contemporary molecular biology. Expression profiling of human tumors, based on the expression of miRNAs and other short or long ncRNAs, has identified signatures associated with diagnosis, staging, progression, prognosis, and response to treatment. In this review we will discuss the recent remarkable advancement in the understanding the biological functions of human ncRNAs in cancer, the mechanisms of expression and the therapeutic potential.

Keywords: small non-coding RNAs; long non-coding RNAs; cancer

---

1. Introduction

The human genome sequencing performed by the International Human Genome Sequencing Consortium revealed that the number of protein-coding genes corresponded only to 20–25,000 [1]. While, at first, it was common belief that the remaining, bigger portion of the human genome was not functional and considered as "junk DNA", several studies based on advanced technologies such as tiling arrays and RNA deep sequencing have recently pointed out the identification of thousands of RNA transcripts not derived from known genes and not encoding for a protein [2,3]. These molecules have been classified as non-coding RNAs (ncRNAs).

NcRNAs could be grouped into two major classes based on the transcript size: small ncRNAs less than 200 bp, such as piRNAs (Piwi-associated RNAs), miRNAs (microRNAs), and snoRNAs (small nucleolar RNAs), and long ncRNAs (lncRNAs), greater than 200 bp. Each of these classes can be further divided, whereas novel subclasses are still being discovered and characterized. All these ncRNAs form huge molecular networks and play a central role in regulating cellular activities in Eukaryotes. The alteration and dysregulation of several ncRNA has been reported in various human diseases, including cancer, providing evidence for targeting these molecules as anticancer agents. Here, we will summarize the current knowledge about regulatory functions of ncRNAs, with special emphasis on their effects in cancer formation and progression.

2. piRNAs

PIWI-family proteins and their associated small RNAs (piRNAs) provide an essential protection for germ-cell genomes against the activity of transponsable elements (TE). They help to maintain genome integrity, silencing TE [4] and this role is highly conserved across animal species. Unlike the other classes of small noncoding RNAs, which are 24–32 nt in length, they are generated from single-stranded RNA precursors through a Dicer-independent mechanism [5–7]. PiRNAs associate with PIWI proteins, which are germline-specific members of the Argonaute protein family, while siRNAs and miRNAs associate with ubiquitously expressed AGO subfamily members. The PIWI protein family is highly conserved across a variety of species and organisms. MIWI, MILI, and MIWI2 (Piwil4) are the three mouse PIWI proteins [8–10], whereas PIWIL1/HIWI, PIWIL2/HILI, PIWIL3, and PIWIL4/HIWI2 are the four PIWI proteins expressed in humans [11]. PIWI mutations in mice, Drosophila, and zebrafish, result in consistent defects in spermatogenic cells, demonstrating the essential role of PIWI proteins in germline development [12–15]. PiRNAs are more expressed in testes than other small noncoding RNAs [16–19] and are involved in spermatogenesis by regulating meiosis and/or suppressing TE. Hundreds of thousands of different piRNA species have been found in mammals [20], with no clear secondary structure motifs but with a common bias 5' uridine. At 3' termini piRNAs present a 2'-*O*-methylation, a process mediated by methyltransferase HEN1, which is associated with PIWI proteins [21,22]. This modification protects piRNA from 3'→5' exonucleases activity, suggesting an increase of their stability [23]. PiRNAs are not distributed across the whole genome but they are clustered in few hundred genomic loci called piRNA clusters [6]. The biogenesis of piRNA could be divided into two main pathways: primary processing and ping-pong amplification cycle (Figure 1). First, piRNA clusters are transcribed in both directions and provide a pool of fragmented primary piRNAs. Primary piRNA transcripts are exported to the cytoplasm where numerous factors (*i.e.*, Zucchini, Armitage and YB) participate in piRNA processing and loading onto PIWI proteins [24,25]. Piwi-piRNA complexes are transported into the nucleus, where they inhibit transcription of TE [26,27]. This first process (primary processing) is similar between germline and somatic cells. The next phase, the ping-pong amplification, is restricted to germline cells and requires slicer activity of PIWI proteins [25,26]. Recently, few works have demonstrated, by deep sequencing, that piRNA population is present in many more cell types than germline cells. For example, Lee *et al.* indentified the presence of a limited set of piRNAs in the mouse hippocampus. The most up-regulated, *DQ541777,* controls spine shape [28]. Moreover, another study described piRNA expression in more than 130 fruit fly, mouse, and rhesus macaque samples. Further, in mouse pancreas and macaque

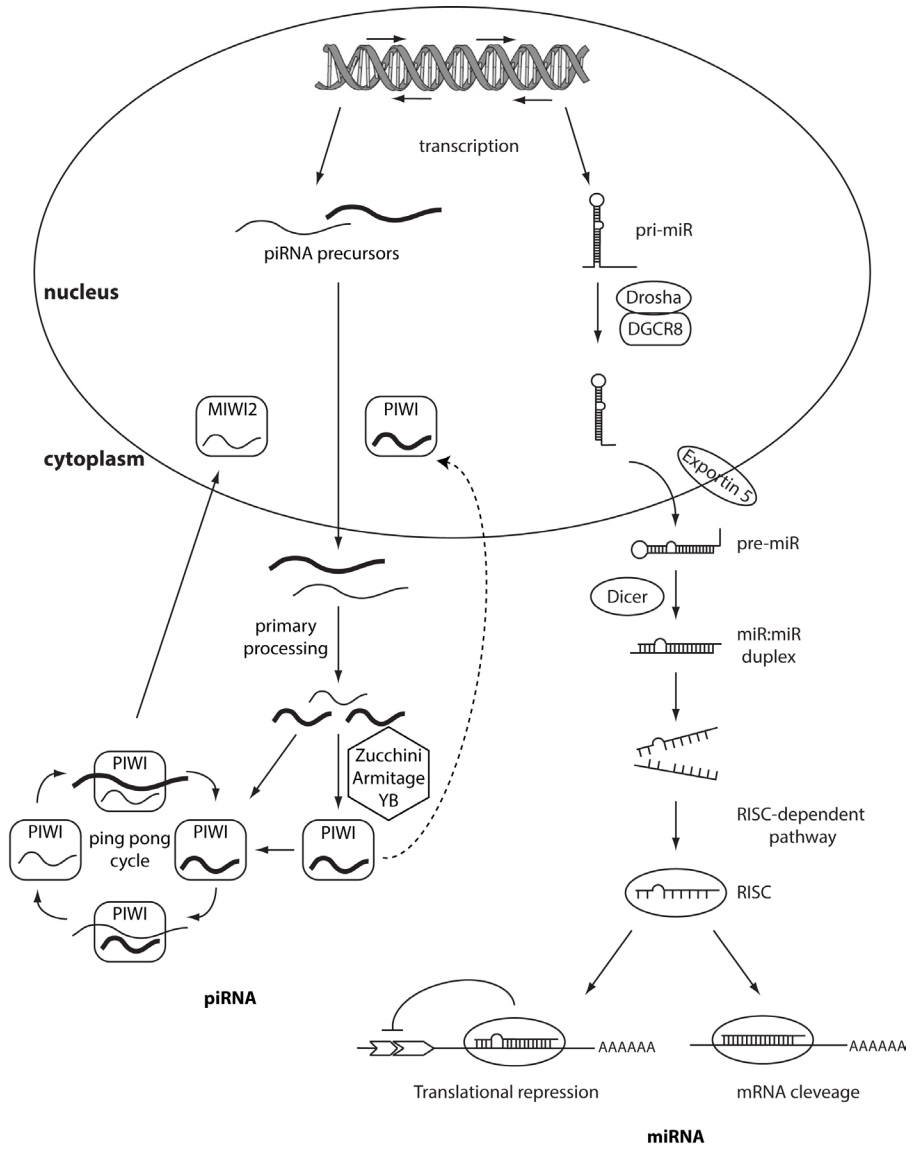
epididymis, piRNA are abundant as much as piRNA abundance in the germline [29]. An emerging number of studies highlights the role of piRNAs or PIWI proteins in the regulation of tumorigenesis. Indeed, piRNAs have been described in HeLa cells [30] and gastric, colon, lung, and breast cancer tissues [31]. These discoveries should not be surprising considering that cancer cells and germ cells share common features such as rapid proliferation and potentially infinite self-renewal. The first evidence of the role of piRNAs in cancer is described by Qiao *et al.* Hiwi, a Piwi family member, is over-expressed in seminomas but not in nonseminomas or in somatic tumors of the adult testis [32]. Moreover, HIWI over-expression has been also shown in cervical, pancreatic, colorectal, endometrial, esophageal, liver cancer, and gliomas [33–39]. Recently, Cheng and colleagues demonstrated that the expression of *piR-651* in gastric, colon, lung, and breast cancer tissues was higher compared to normal adjacent tissues. The levels of *piR-651* were associated with tumor-node-metastasis (TNM) stages. Inhibition of *piR-651* caused the arrest of gastric cancer cells at the G_2/M phase [31]; therefore this pi-RNA shows an oncogenic role and plays a crucial function in carcinogenesis. Another study demonstrated the down-regulation of *piR-823* in gastric cancer tissues compared to normal tissues suggesting its potential tumor suppressive role [40]. *In vivo* studies showed that the over-expression of *piR-823* significantly inhibited tumor growth in a dose-dependent manner. Moreover, *piR-823* was significantly lower in peripheral blood of gastric cancer patients compared to healthy controls. The levels of *piR-823* were positively associated with TNM stages and distant metastasis, suggesting that *piR-823* should be tested as a biomarker for detecting circulating gastric cancer cells in the blood [41]. All these data may suggest an important role of the axis PIWI and PIWI-associated RNAs going beyond the regulation of the genome in germline tissues and more studies are needed in order to investigate their specific role in tumorigenesis.

3. MicroRNAs

In 1993, Victor Ambros and colleagues discovered a gene, *lin-4*, that affected the development of *Caenorhabditis elegans* and found that its product was a small nonprotein-coding RNA [42]. The number of known small RNAs in different organisms such as *Caenorhabditis elegans*, *Drosophila melanogaster*, plants, and mammals, including humans, has since expanded substantially. MicroRNAs (miRNAs) are 19- to 24- nucleotide non-coding RNA molecules that regulate the expression of target mRNAs both at the transcriptional and translational level [43,44]. While in plants such regulation occurs through perfect base-pairing, usually in the 3' untranslated region (UTR) of the targeted mRNA, in mammals the base-pairing is only partial [45,46].

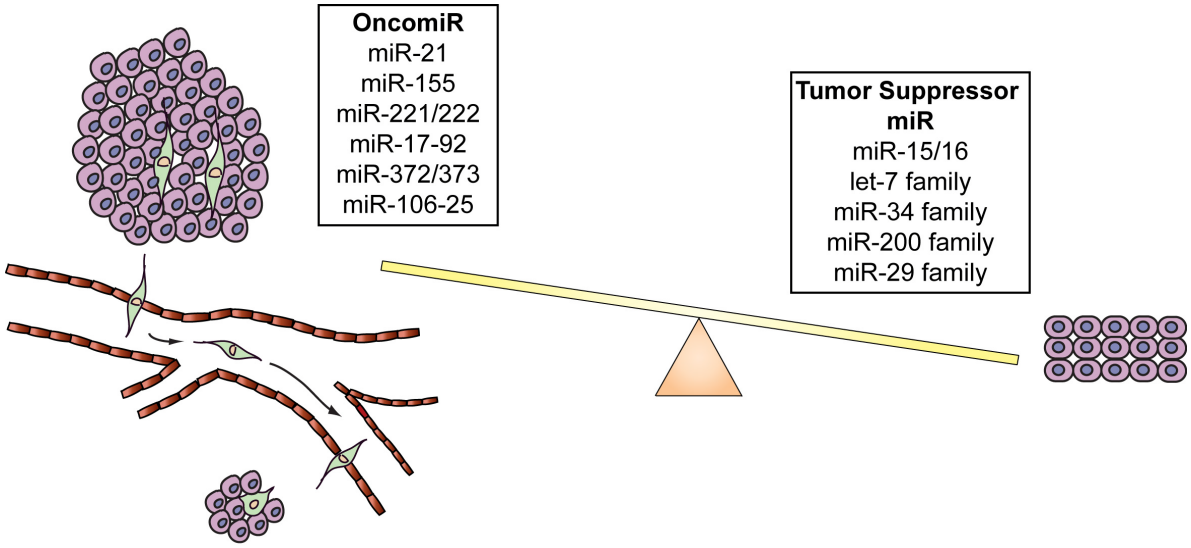
Each member of this large family of non-coding RNAs can have hundreds of different targets and nearly 30% of the genes are regulated by, at least, one miRNA [43]. Several studies have demonstrated the involvement of miRNAs in different biological processes such as proliferation, cell cycle regulation, proliferation, apoptosis, differentiation, development, metabolism, neuronal patterning, and aging [43–53].

Figure 1. PiRNAs and microRNAs biogenesis. On the left, piRNAs biogenesis. PiRNAs are processed from single-stranded RNA precursors. The biogenesis of piRNAs could be divided in two main pathways: primary processing and ping-pong amplification cycle. MIWI2, a PIWI protein, translocates processed piRNAs into the nucleus, where they block the transcription of the TE (trasposon elements). On the right, miRNA biogenesis. Primary transcripts (pri-miRs) are transcribed by the RNA polymerase II. In the nucleus pri-miRs are processed by Drosha-DGCR8 into pre-miRs of ~60–70 nt. The produced pre-miRNAs are exported by the Exportin 5 to the cytoplasm where they are processed in ~18–22-nucleotide miRNA duplexes by the cytoplasmic RNase III Dicer. Normally, one strand of this duplex is degraded, whereas the other strand accumulates as a mature miRNA. From the miRNA-miRNA duplex, only the miRNA enters preferentially in the protein effector complex, formed by the RNA-induced silencing complex (RISC) and miRgonaute. Perfect or nearly perfect complementarities between miRNA and its target 3' UTR induce RISC to cleave the target mRNA, whereas imperfect base matching induces mainly translational silencing of the target.



MiRNAs are transcribed by an RNA polymerase II as long, capped and polyadenylated precursors called pri-miRNAs [54,55], which are further cleaved into hairpin-shaped ~70–100 nucleotides precursors (pre-miRNAs) by a ribonuclease III (Drosha) and the double-stranded DNA binding protein DGCR8/Pasha [56] (Figure 1). Exportin 5 then translocates the pre-miRNAs to the cytoplasm [57], where another RNAse III, Dicer [43,58] further processes the precursor in a double strand RNA of about 24 nt. The double-stranded RNA is incorporated into the RISC (RNA-induced silencing complex) but only one strand, the mature microRNA, remains stably associated with the RISC and will drive the complex to the target mRNA. If the base-pairing between miRNA and the 3' UTR of the target mRNA is perfect the messenger is cleaved and degraded, whereas imperfect complementarity will result in translational silencing without mRNA degradation [59,60]. Several studies have demonstrated that miRNAs have a crucial role in cancer formation and spread. These small non-coding RNAs are, in fact, usually located in minimal regions of amplification, loss of heterozygosity, fragile sites, and common breakpoint regions in or in proximity of oncogenes or tumor suppressor genes. Moreover, profiling studies have demonstrated that miRNAs are differentially expressed in tumors *vs.* normal human tissues. These data have allowed the classification of microRNAs into two groups: oncomiRs (which act as oncogenes and are usually overexpressed in cancer, promoting tumor formation and spread) and tumor-suppressor miRs (which impair tumor growth and are usually silenced because of mutations, promoter methylation, or chromosomal rearrangements) [61–64], although some microRNAs can act as both oncogene or tumor-suppressor gene depending on the cellular context [65] (Figure 2 and Table1).

Figure 2. OncomiRs and tumor suppressor miRs. Correct cellular homeostasis is driven by a proper balance between oncomiRs and tumor suppressor miRs. OncomiRs are usually located in the amplified regions of the genome and are frequently over-expressed in cancer, promoting tumor growth and metastasis. Tumor suppressor miRs are often down-regulated in cancer and inhibit tumor growth inducing apoptosis and blocking cell migration.



3.1. OncomiRs

One of the most well-known oncomiRs is *miR-21*, overexpressed in different types of cancer such as chronic lymphocytic leukemia (CLL) [62], acute myelogenous leukemia (AML) [66], glioblastoma [67], pancreatic, prostate, colon, gastric, breast, and lung cancer [68]. In 2008, Asangani and coworkers [69] demonstrated that *miR-21* downregulated the tumor suppressor PCDC4 (programmed cell death 4) promoting tumor invasion and metastasis in colorectal cancer. Zhang and colleagues [70] showed that *miR-21* induced growth and invasion in non-small cell lung cancer by repressing PTEN (phosphatase and tensin homolog); moreover, *miR-21* modulate TRAIL sensitivity in glioma cells mainly by modulating caspase-3 and TAp63 expression and TRAIL-induced caspase machinery [71], confirming that *miR-21* acts like an oncogene by blocking the expression of critical apoptosis-related genes.

Another example of oncomiR is represented by *miR-155*. Similarly to *miR-21*, *miR-155* is highly expressed in CLL [72], AML [73], lung, breast and pancreatic cancer [68], Hodgkin disease [72], and primary mediastinal non-Hodgkin's lymphoma [62]. In 2010, Jiang and coworkers demonstrated that *miR-155* targeted the tumor suppressor gene *Socs1* (*suppressor of cytokine signaling 1* gene) in human breast cancer cells, promoting cell proliferation, colony formation, and xenograft tumor growth [74]. *MiR-155* has also been found to be one of the most potent miRNAs suppressing apoptosis in human T cell leukemia (Jurkat cells) and in MDA-MB-453 breast cancer cells [75]. Moreover, in a transgenic mouse model, selective overexpression of *miR-155* in B cells led to early B cells polyclonal proliferation with a high-grade lymphoma-pre-B leukemia, suggesting that *miR-155* promotes the initiation and progression of the disease [76].

MiR-221 and *-222* are also up-regulated in several solid tumors, such as hepatocarcinoma [77], breast estrogen negative cells [78], melanoma cells [79], thyroid cancer [80]. Both these miRNAs induce tumor growth and spread of several cancer cell lines [81–83]. In 2009, our group demonstrated that hepatocyte growth factor receptor (MET) oncogene, through Jun transcriptional activation, upregulated *miR-221* and *-222* expression, which in turn, by targeting *PTEN* and *TIMP3*, conferred resistance to tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)-induced cell death and enhanced tumorigenicity of lung and liver cancer cells. Therefore, the use of microRNAs in therapeutic intervention could sensitize tumor cells to drug-inducing apoptosis and also inhibit their survival, proliferation, and invasive abilities [84].

The oncomiR group is wide, and comprises other microRNAs such as the *miR-17-92* cluster, which is crucial for B-cell proliferation and its absence induces an increase of the proapoptotic protein Bim and inhibits the pro-B to pre-B cell development [85]; *miR-372/373*, which are involved in the development of human testicular germ cell tumors by neutralizing the TP53 pathway [86]; *miR-10b*, which promotes cell migration and invasion in breast cancer [87]; the polycistron *miR-106-25*, which acts as an oncogene by interfering with the synthesis of p21 and Bim [88].

3.2. Tumor Suppressor MicroRNAs

The group of miRNAs able to inhibit cell growth, induce apoptosis, and block cell cycle, are called tumor suppressor miRs. Normally, oncomiRs are located mainly in the amplified regions in human

cancers and are frequently over-expressed in neoplastic tissues. Conversely, tumor suppressor miRs are located in the deleted regions and are often down-regulated in cancerous tissues.

The first evidence that miRNAs are involved in cancer comes from the finding that *miR-15* and *miR-16* are down-regulated or deleted in most patients with chronic lymphocytic leukemia [61].

Their expression is inversely related to several oncogenes, such as *Bcl-2* [89], *CCND1*, *WNT3A* [90], *Ccne1*, *Bmi-1* [91], and *VEGF-A* [92], which induce cell proliferation, survival, invasion and angiogenesis. Recently it has been shown that *miR-15* and *-16* are involved in drug resistance. Pouliot *et al.* demonstrated that *miR-15* and *-16* sensitized cisplatin-resistant epidermoid carcinoma cells to apoptosis by targeting *WEE1* and *CHK1* [93].

Another example of tumor suppressor miR is represented by the *let-7* family. Several studies described the down-regulation of *let-7* family in numerous tumors, including lung [94], gastric [95], colon cancer [96], and Burkitt's lymphoma [97]. *Let-7* family targets and inhibits the expression of several oncogenes such as *c-Myc* [97], *Ras* [98], *high-mobility group A* (*HMGA*) [99], *Janus protein tyrosine kinase* (*JAK*) and signal transducer and activator of transcription 3 (*STAT3*) pathway [100]. A recent study also reported that let-7 directly targets *PAK1*, *DIAPH2*, *RDX*, and *ITGB8*, multiple genes involved in the actin cytoskeleton pathway, inhibiting breast cancer cell migration [101].

The tumor suppressor activity of miR-34 family has been demonstrated in cancer cell types of lung [102], liver [103], breast [104], colon [105], brain [106], ovary [107], esophagus [108], and the lymphoid system [109]. In mammals, *miR-34* family comprises three processed miRNAs that are encoded by two different genes: *miR-34a* is encoded by its own transcript, whereas *miR-34b* and *-34c* share a common primary transcript. Their expression is directly induced by p53 in response to DNA damage or oncogenic stress [110]. *MiR-34* family inhibits many different oncogenic pathways involved in the control of cellular proliferation, cell cycle, and senescence by targeting oncogenes such as mitogen-activated protein kinase kinase 1 (*MEK1, MAP2K1*), R-Ras (*RRAS*), platelet-derived growth factor receptors (*PDGFRA and PDGFRB*) [111], and hepatocyte growth factor receptor (*MET*), *BCL2* and survivin.

MiR-200 family is commonly lost in aggressive tumors such as lung, prostate and pancreatic cancer. It has been shown that *miR-200* family members directly target *ZEB1* and *ZEB2*, transcriptional repressors of E-cadherin [112], and *BMI1*, reducing epithelial mesenchimal transition [113].

MiR-29s are also downregulated in multiple cancer types such as CLL [62], breast [114] and cervical cancer [115], and hepatocellular carcinoma [116]. *MiR-29* family targets and inhibits *DNMT3A* and *-3B* (DNA methyltransferases 3A and 3B) [117], *Tcl1* in chronic lymphocytic leukaemia [118] and *Mcl1* in cholangio-carcinoma [119]. Moreover, it has been demonstrated that the down-regulation of *miR-29* by MYC, HDAC, and EZH2 promotes cell survival and growth in MYC-associated lymphomas [120]. In conclusion, the correct cell homeostasis and survival are driven by a proper balance between oncomiRs and tumor suppressor miRs. Up-regulation of oncomiRs or down-regulation of tumor suppressor miRs leads to cancer formation and progression (Figure 2 and Table 1).

Table 1. OncomiRs and tumor suppressor miRs.

miRNA	Tumor type	Status	References
miR-21	CLL, AML, glioblastoma, pancreatic, prostate, colon, gastric, breast and lung cancer	Up-regulated	[62,66–71]
miR-155	CLL, AML, lung, breast and pancreatic cancer, Hodgkin disease, primary mediastinal non-Hodgkin's lymphoma	Up-regulated	[62,66,68,72,74–76]
miR-221/222	hepatocarcinoma, breast cancer, melanoma, thyroid cancer and glioma	Up-regulated	[77–84]
miR-17-92	AML	Up-regulated	[85]
miR-372/373	testicular germ cell tumor	Up-regulated	[86]
miR-10b	breast cancer	Up-regulated	[87]
miR-106-25	gastric cancer	Up-regulated	[88]
miR-15-16	CLL, prostate and ovarian cancer and multiple myeloma	Down-regulated	[61,89–93]
let-7 family	lung, gastric, colon, breast cancer and Burkitt's lymphoma	Down-regulated	[94–101]
miR-34	lung, liver, breast, colon, brain, ovary, esophageal cancer and non-small cell lung cancer (NSCLC)	Down-regulated	[102–111]
miR-200	lung, prostate and pancreatic cancer	Down-regulated	[112,113]
miR-29	CLL, breast and cervical cancer hepatocellular and cholangio-carcinoma	Down-regulated	[62,114–120]

4. snoRNAs

Small nucleolar RNAs (snoRNAs) are small non-coding RNAs whose length ranges from 60 to 300 nucleotides. SnoRNAs are normally located within introns of protein-coding genes and are transcribed by RNA polymerase II, but in some cases they can be found within introns of lncRNAs [121,122]. Within the cell, snoRNAs specifically accumulate in the nucleolar compartment, where they are responsible of the 2'-*O*-ribose methylation and pseudouridylation of specific ribosomal RNA nucleotides, essential modifications for the efficient and accurate production of the ribosome [123].

SnoRNAs can be classified into two groups: H/ACA box and C/D box. In both cases, snoRNAs hybridize specifically to the complementary sequence in the rRNAs and the associated protein complexes (C/D or H/ACA snoRNP) carry out the appropriate modification on the nucleotide that is identified by snoRNAs [124–126].

The H/ACA box snoRNAs family is involved in pseudouridylation of rRNAs. These ncRNAs have two major hairpin elements, connected by a hinge, and followed by a short tail region containing the conserved H and ACA box motifs that are located at the bases of the 5' and 3', respectively. The sequence specificity for the pseudouridylation is carried by two short antisense elements located in an internal loop of the 5' and/or 3' hairpins [127].

The C/D box snoRNAs, instead, are mainly involved in the 2'-*O*-methylation of rRNAs. This group of ncRNAs contains two short sequence motifs, box C (5'-PuUGAUGA-3') and box D (5'-CUGA-3')

located near the 5' and the 3' ends, respectively. These elements form a terminal stem-box structure, involving not only elements required for snoRNAs nuclear localization, but also another copy of the box C, named box C', in their central portion, and another box D, named box D'. 2'-*O*-methylation is carried out through one or two antisense elements located upstream box D and/or box D' and complementary to a site of rRNA 2'-*O*-ribose methylation [128]. The process of snoRNAs maturation has not been entirely unveiled, however it has been demonstrated that the maturation of box C/D snoRNAs in yeast can occur through two pathways (Figure 3) [129]. In the first pathway, splicing of a pre-mRNA leads to a snoRNA-containing lariat, which is then linearized by the enzyme Dbr1p. Thanks to the activity of endonucleases and exonucleases the mature snoRNA is finally released. The second pathway, instead, is splicing-independent: the snoRNA is excided from the intron of the pre-mRNA directly, leading to the destruction of the mRNA. However, this latter pathway is still not well defined and the enzymes involved in this process have not been totally identified.

Although the main function of snoRNAs seems to be related to rRNA folding and stabilization, recent discoveries have pointed out a wider regulatory function for these small ncRNAs. For example, snoRNAs seem to be involved in miRNA synthesis. In 2010, Breimer and coworkers identified several box C/D sno-miRNAs, originating from relatively short snoRNAs (such as *U27* and *HBII-336*) displaying miRNA features in mRNAs silencing in different cell types, therefore controlling several biological processes normally regulated by miRNAs [130].

SnoRNAs are also involved in the onset of the Prader-Willy syndrome (PWS), induced by the genetic loss of the 15q11–q13 locus, normally active only on the paternal allele. This site is characterized by several copies of the *HBII-85* snoRNA, whose loss seems to be correlated with the PWS phenotype, both in human and in mice [131,132]. Moreover, recent studies reported the involvement of snoRNAs in cancer formation and progression, although the exact molecular mechanisms by which snoRNAs regulate cancer are still unknown.

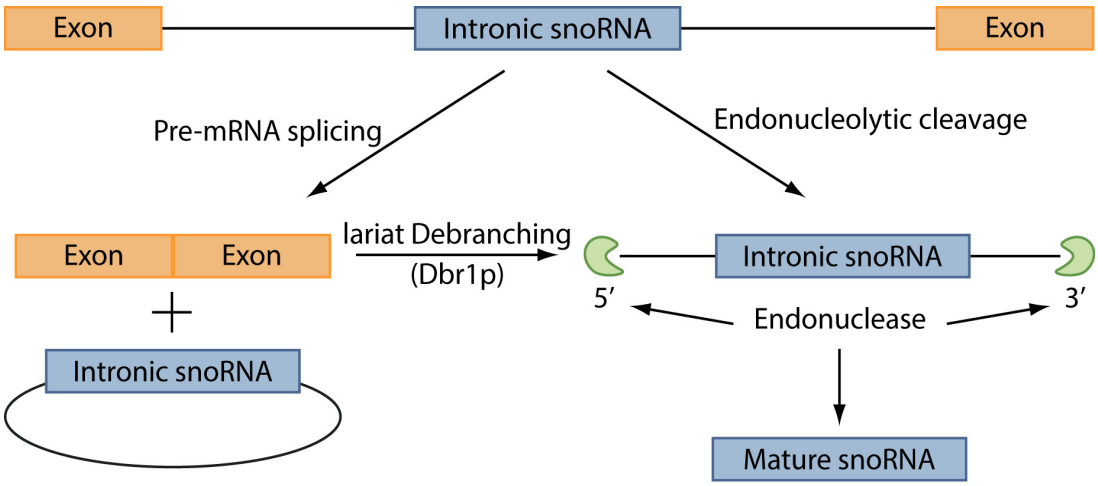
Similarly to miRNAs, snoRNA expression has been found deregulated in cancer patient samples. In fact, the expression of *GAS5* (growth arrest specific 5), a gene which encodes an lncRNA but also harbors ten intronic snoRNAs, is downregulated in breast cancer compared to normal adjacent epithelial breast tissue. *GAS5* transcript sensitizes mammalian cells to apoptosis inducers, therefore displaying a tumor-suppressor role [133]. Moreover, Nakamura and coworkers demonstrated that *GAS5* was a partner of BCL6 in a patient with diffuse large B-cell lymphoma, carrying the chromosomal translocation t (1; 3) (q25; q27) [134], while Gee showed that *GAS5* low expression correlated with poor prognosis in breast cancer and head and neck squamous carcinoma [135]. The same authors also reported that snoRNA U50 is frequently transcriptionally downregulated in breast and prostate cancer [136] and that its 2-nucleotides somatic and germline deletion led to increased incidence of homozygosity for the deletion in cancer cells.

Other snoRNAs, such as *snoRNA42*, overexpressed in NSCLC cells, are located at frequently amplified genomic regions in tumors, therefore acting like oncogenes and promoting tumor growth. In 2011 Mei and coworkers found that *snoRNA42* knockdown in NSCLC cells impaired tumorigenicity *in vitro* and *in vivo* promoting apoptosis in a p53-dependent manner; conversely its enforced expression in bronchial epitheliums promoted cell growth [137].

Moreover, Liao *et al.* performed a profiling study on 22 NSCLC tissues. They found an overexpression of six snoRNAs compared to normal specimens. In addition to *snoRNA42*, they

identified *SNORD33*, *SNORD66*, *SNORD73B*, *SNORD76* and *SNORD78*. Of these, *SNORD33*, *SNORD66*, and *SNORD76* expression in the plasma of NSCLC patients was higher compared to cancer-free individuals [138]. It has been demonstrated that, in addition to deregulated snoRNAs, also mutations of genes encoding for snoRNPs (snoRNA-associated proteins) can promote tumorigenesis. One of these genes is the human dyskerin, a putative pseudouridine synthase involved in the rRNA pseudouridylation and in the stabilization of the telomerase RNA elements. Mutations of its gene, *DKC1*, cause the X-linked genetic disease dyskeratosis congenita and promote tumor formation in mice [139]. The same effects have been described when point mutations occur in the genes encoding NOP10 and NHP2, both components of the H/ACA snoRNPs.

Figure 3. Intronic snoRNA processing. SnoRNA maturation occurs through two distinct pathways: splicing-dependent and splicing-independent. In the first pathway, the splicing of a pre-mRNA leads to a snoRNA-containing lariat, which is linearized by the enzyme Dbr1p and then endonucleases and exonucleases release the mature snoRNA. In the splicing-independent pathway the snoRNA is directly excided from the intron of the pre-mRNA by endonucleolytic cleavage.



5. Long Noncoding RNAs

Several studies based on RNA deep sequencing and genome-wide analysis have recently pointed out that the genome of mammals, as well of other organisms, contains thousands of long transcripts whose length ranges from 200 nt to 100 kilobases, called long non-coding RNAs (lncRNAs or lincRNA, for long intergenic ncRNA) [3,140–144]. LncRNAs are located within nuclear or cytosolic fractions [145]. They are usually transcribed by RNA polymerase II but have no open reading frame [146], and map to intronic and intergenic regions [147]. Moreover, they display epigenetic features common to protein-coding genes, such as trimethylation of histone 3 lysine 4 (H3K4me3) at the transcriptional start site (TSS), and trimethylation of histone 3 lysine 36 (H3K36me3) throughout the gene body [148,149]. It has been estimated that nearly 15,000 lncRNAs are present in the human genome, but only a small fraction is expressed in a given cell type. All the information regarding identified lncRNAs has been catalogued and is available at the website http://www.lncrnadb.org [150–152].

Although they were initially thought to be the product of a "noisy" inconsequential transcription resulting from low RNA polymerase fidelity [142], recent studies have demonstrated that lncRNAs regulate several biological processes such as transcription [153–156], translation [157], cellular differentiation [158], regulation of gene expression [159], cell cycle regulation [160,161], chromatin modification [143,162,163], and nuclear-cytoplasmic trafficking [159,164–167].

LncRNAs have been also found to guide protein complexes which regulate chromatin modification or transcription to their targets [143,162,168,169]. Finally, it has been demonstrated that lncRNAs are dysregulated in several human diseases, including cancer.

Dysregulated expression of lncRNAs in cancer marks the spectrum of disease progression [170] and may serve as an independent predictor for patient outcomes [171].

Long non-coding RNAs can mediate epigenetic changes by recruiting chromatin-remodeling complexes to specific genomic loci. A recent study found that 20% of 3300 human long non coding RNAs are bound by Polycomb Repressive Complex 2 (PRC2) [162]. Although the specific molecular mechanisms are not defined, there are several examples that illustrate the silencing potential of lncRNAs. The first, most known example is represented by *Xist* (X-inactive-specific transcript) gene, which encodes an lncRNA crucial for the inactivation of the X-chromosome in mammals [172]. Basically, *Xist* physically coats one of the two X-chromosomes and recruits the chromatin regulator PRC2 (Polycomb chromatin remodeling complex) to this chromosome, promoting the formation of heterochromatin through histone modifications [173]. Another important example is represented by the hundreds of long ncRNAs which are sequentially expressed in the human homeobox (Hox) loci, where they define chromatin domains of differential histone methylation and RNA polymerase accessibility [174]. One of these ncRNAs, Hox transcript antisense RNA (*HOTAIR*) regulates *in trans* human *HOXD* genes expression through the induction of a repressive chromatin state. This occurs through the association of *HOTAIR* with the chromatin-modifying complexes PRC2, LSD1, and coREST/REST [143,162,171]. As modulator of epigenetic landmark, it has been shown that *HOTAIR* has a profound effect on tumorigenesis. Indeed, it is upregulated in breast and colon cancers and it is associated with metastasis and poor prognosis [171]. Another important effect of lncRNAs on chromatin modification with important consequences in cancer is represented by the lncRNA *ANRIL*, which controls the epigenetic status of the locus *INK4b/ARF/INK4a* by interacting with subunits of PRC1 and PRC2. High expression of *ANRIL* has been found in some cancer tissues such as melanoma and prostate cancers ([175,176]. The long noncoding RNA *MALAT1* (metastasis-associated lung adenocarcinoma transcript 1), also known as *NEAT2* (nuclear-enriched abundant transcript 2), is a highly conserved nuclear noncoding RNA (ncRNA) which acts as molecular decoy serving as a structural link in ribonucleoprotein (RNPs) complexes. Gutschner and colleagues developed a *MALAT1* knockout model in human lung tumor cells. In lung cancer, *MALAT1* does not alter alternative splicing but actively regulates gene expression including a set of metastasis-associated genes. Consequently, *MALAT1*-deficient cells are impaired in migration and form fewer tumor nodules in a mouse xenograft. Antisense oligonucleotides (ASO), blocking *MALAT1*, prevent metastasis formation after tumor implantation [177]. In addition to these active lncRNAs acting as oncogenes, there are also lncRNAs with tumor suppressor function. One very famous example is the ncRNA *GAS5* (Growth Arrest-Specific 5). It was originally identified based on its increased levels in growth-arrested mouse NIH3T3 fibroblasts [132]. *GAS5* binds to the DNA-binding domain of the

glucocorticoid receptor (GR) by acting as a decoy glucocorticoid response element (GRE), thus competing with DNA GREs for binding to the GR [156]. *GAS5* negatively regulates the survival of lymphoid and breast cells, and is aberrantly expressed in several cancers. Pickard *et al.* showed that *GAS5* promotes apoptosis of prostate cells after irradiation with UV-C light and low levels of *GAS5* expression may therefore reduce the effectiveness of chemotherapeutic agents [178].

Recently, lncRNAs have also shown their tumorigenic potential by modulating the transcriptional program of p53 [179].

A 3kb lncRNA, *linc-RNA-p21*, transcriptionally activated by p53, has been shown to collaborate with p53 in order to control gene expression in response to DNA damage. Silencing of *lincRNA-p21* derepresses the expression of hundreds of genes through the interaction with hnRNP-K (Heterogeneous nuclear ribonucleoprotein K), thus, promoting apoptosis of abnormal cells or restraining tumors [179].

LncRNA *PANDA* is induced in response to external stimuli in a p53-dependent manner. After DNA damage p53 directly binds to the *CDKN1A* locus and activates *PANDA*, which enables cell-cycle arrest and impairs the expression of pro-apoptotic genes thanks to its interaction with the transcription factor NF-YA [LNC8] [180].

In addition to the features described above, recent studies have unveiled other properties of lncRNAs. For instance, it has been demonstrated that pseudogene transcripts are biologically active as they can regulate mRNA stability. One example is given by the tumor suppressor pseudogene *PTENP1*, whose 3' UTR region is very similar to the untraslated region of *PTEN* transcript. Both these regions bind the same set of miRNAs, and *PTENP1* pseudogene may act as “decoy” by protecting *PTEN* mRNA from common miRNA binding and allowing the expression of the tumor suppressor protein. *PTENP1* pseudogene therefore belongs to the group of competing endogenous RNAs (ceRNAs). Similarly, *KRAS* and *KRAS1P* transcript levels have been found positively correlated, corroborating that pseudogene functions mirror the role of their cognate genes as explained by a miRNA decoy mechanism. In cancer, specific mutations at the binding site of these pseudogenes impair their activity, therefore promoting tumor progression [181].

Enhancer-like lncRNAs (eRNAs) were discovered by Ørom and colleagues in 2010 [141]. The authors used a GENCODE annotation of the human genome to characterize over a thousand lncRNAs in several cell lines, finding that some of these RNAs displayed an enhancer-like function. Depletion of these ncRNAs led to a decreased expression of their neighboring protein-codon genes, such as the regulator of hematopoiesis *SCL*, *Snai1*, and *Snai2*, indicating that eRNAs play a pivotal role in development and differentiation. Moreover, Melo *et al.* showed through genome-wide chromatin-binding profiles that p53 protein binds also to regions located at distant sites from any known p53 target gene. Such regions were characterized not only by conserved p53-binding sites but displayed also enhancer activity and interacted with multiple neighboring genes allowing long-distance p53-dependent transcription regulation [182].

Finally, Natural Antisense Transcripts (NATs) are a large class of lncRNA transcribed from the opposite DNA strand to other transcripts and overlap in part with sense RNA. NATs play an important role in antisense regulation in gene expression. NATs have been implicated in several processes such as RNA translation [183] and transcriptional interference [184], and they have a pivotal role also in cancer. *aHIF*, a NAT derived from the 3' UTR of *HIF1*, represents the first case of overexpression of a

NAT associated with a specific human malignant disease: non-papillary clear-cell renal carcinomas, but not in papillary renal carcinomas [185]. Moreover, it has been demonstrated that *aHIF* expression is a poor prognosis marker in breast cancer [186].

The already mentioned *ANRIL* is an antisense lncRNA originates from the *INK4B-ARF-INK4A* locus, which contains three tumour suppressor genes, and it is overexpressed in prostate cancer tissues. Repression of *ANRIL* expression was associated with a reduction in cellular proliferation and increased the expression of both $p16^{Ink4A}$ and $p15^{INK4B}$, which are encoded by *CDKN2A* and *CDKN2B*, respectively [176]. *BOKAS* is a natural antisense transcript of Bok, a proapoptotic member in the Bcl-2 family. The expression of *BOKAS* was only detected in testis and different cancer tissues but not in other normal adult tissues. Overexpression of *BOKAS* was able to inhibit Bok-induced apoptosis in HeLa cells [187].

Another example of NAT is represented by *Zeb2/Sip1* NAT. This NAT regulates E-cadherin expression by increasing the levels of Zeb2 protein, a transcriptional repressor of E-cadherin, suggesting a role for noncoding RNAs in the control of epithelial morphology [188].

6. Conclusions

The recent discoveries regarding the biogenesis and function of ncRNAs have definitely improved our undestanding of the complexity of the human genome and the regulation of several processes. In particular, the involvement of microRNAs in the regulation of cell cycle, proliferation, differentiation and, most of all, in cancer formation and progression has certainly opened new fields of research aimed to better elucidate their mechanisms of action. Recently, our group reported that miRNAs secreted through exosomes bind to Toll-like receptor 8 (TLR8) in human and TLR7 in mouse inducing a pro-inflammatory response [189]. Therefore, in addition to their post-transcriptional regulatory function, miRNAs act like hormones and are involved into cell-to-cell communication. Other studies have shown the presence of tumor-derived microRNAs in serum or plasma as an approach for blood-based detection of human cancers, indicating that microRNAs could be used as circulating biomarkers [190]. Several groups are currently investigating the possibility to use microRNAs as therapeutic tools alone or in combination with chemotherapy.

Successfull systemic delivery of miRNAs as anti-cancer approaches in preclinical models using liposomes [191], viral vectors [192], and nanoparticles [193] has been reported. There is no doubt that miRNAs and other ncRNAs play a very important role in the regulation of pathways involved in tumor development and progression. Although there are still several obstacles to overcome before clinical testing of miRNA therapeutics, such as delivery and chemical modification of miRNA modulators, the fact that ncRNAs are natural antisense interactors and regulate many genes involved in survival and proliferation makes them excellent candidates to become powerful therapeutic tools in the near future.

Acknowledgements

We thank Justin Middleton for editorial assistance. Michela Garofalo is recipient of the Kimmel Award.

Conflicts of Interest

The authors declare no conflict of interests.

References

1. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* **2004**, *431*, 931–945.
2. Kapranov, P.; Cheng, J.; Dike, S.; Nix, D.A.; Duttagupta, R.; Willingham, A.T.; Stadler, P.F.; Hertel, J.; Hackermuller, J.; Hofacker, I.L.; *et al.* RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **2007**, *316*, 1484–1488.
3. Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M.C.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C.; *et al.* The transcriptional landscape of the mammalian genome. *Science* **2005**, *309*, 1559–1563.
4. Senti, K.A.; Brennecke, J. The piRNA pathway: A fly's perspective on the guardian of the genome. *Trends Genet.* **2010**, *26*, 499–509.
5. Vagin, V.V.; Sigova, A.; Li, C.; Seitz, H.; Gvozdev, V.; Zamore, P.D. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **2006**, *313*, 320–324.
6. Brennecke, J.; Aravin, A.A.; Stark, A.; Dus, M.; Kellis, M.; Sachidanandam, R.; Hannon, G.J. Discrete small RNA-generating loci as master regulators of transposon activity in drosophila. *Cell* **2007**, *128*, 1089–1103.
7. Houwing, S.; Kamminga, L.M.; Berezikov, E.; Cronembold, D.; Girard, A.; van den Elst, H.; Filippov, D.V.; Blaser, H.; Raz, E.; Moens, C.B.; *et al.* A role for piwi and piRNAs in germ cell maintenance and transposon silencing in zebrafish. *Cell* **2007**, *129*, 69–82.
8. Kuramochi-Miyagawa, S.; Kimura, T.; Ijiri, T.W.; Isobe, T.; Asada, N.; Fujita, Y.; Ikawa, M.; Iwai, N.; Okabe, M.; Deng, W.; *et al.* Mili, a mammalian member of *piwi* family gene, is essential for spermatogenesis. *Development* **2004**, *131*, 839–849.
9. Deng, W.; Lin, H. *Miwi*, a murine homolog of *piwi*, encodes a cytoplasmic protein essential for spermatogenesis. *Dev. Cell* **2002**, *2*, 819–830.
10. Carmell, M.A.; Girard, A.; van de Kant, H.J.; Bourc'his, D.; Bestor, T.H.; de Rooij, D.G.; Hannon, G.J. Miwi2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev. Cell* **2007**, *12*, 503–514.
11. Sasaki, T.; Shiohama, A.; Minoshima, S.; Shimizu, N. Identification of eight members of the argonaute family in the human genome small star, filled. *Genomics* **2003**, *82*, 323–330.
12. Chen, Y.; Pane, A.; Schupbach, T. Cutoff and aubergine mutations result in retrotransposon upregulation and checkpoint activation in drosophila. *Curr. Biol.* **2007**, *17*, 637–642.
13. Cox, D.N.; Chao, A.; Baker, J.; Chang, L.; Qiao, D.; Lin, H. A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev.* **1998**, *12*, 3715–3727.
14. Houwing, S.; Berezikov, E.; Ketting, R.F. Zili is required for germ cell differentiation and meiosis in zebrafish. *EMBO* **2008**, *27*, 2702–2711.

15. Schupbach, T.; Wieschaus, E. Female sterile mutations on the second chromosome of drosophila melanogaster. II. mutations blocking oogenesis or altering egg morphology. *Genetics* **1991**, *129*, 1119–1136.
16. Aravin, A.; Gaidatzis, D.; Pfeffer, S.; Lagos-Quintana, M.; Landgraf, P.; Iovino, N.; Morris, P.; Brownstein, M.J.; Kuramochi-Miyagawa, S.; Nakano, T.; *et al.* A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **2006**, *442*, 203–207.
17. Girard, A.; Sachidanandam, R.; Hannon, G.J.; Carmell, M.A. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **2006**, *442*, 199–202.
18. Grivna, S.T.; Beyret, E.; Wang, Z.; Lin, H. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev.* **2006**, *20*, 1709–1714.
19. Watanabe, T.; Takeda, A.; Tsukiyama, T.; Mise, K.; Okuno, T.; Sasaki, H.; Minami, N.; Imai, H. Identification and characterization of two novel classes of small RNAs in the mouse germline: Retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes Dev.* **2006**, *20*, 1732–1743.
20. Das, P.P.; Bagijn, M.P.; Goldstein, L.D.; Woolford, J.R.; Lehrbach, N.J.; Sapetschnig, A.; Buhecha, H.R.; Gilchrist, M.J.; Howe, K.L.; Stark, R.; *et al.* Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Mol. Cell* **2008**, *31*, 79–90.
21. Horwich, M.D.; Li, C.; Matranga, C.; Vagin, V.; Farley, G.; Wang, P.; Zamore, P.D. The drosophila RNA methyltransferase, dmhen1, modifies germline piRNAs and single-stranded siRNAs in risc. *Curr. Biol.* **2007**, *17*, 1265–1272.
22. Saito, K.; Sakaguchi, Y.; Suzuki, T.; Suzuki, T.; Siomi, H.; Siomi, M.C. Pimet, the drosophila homolog of hen1, mediates 2'-*O*-methylation of piwi-interacting RNAs at their 3' ends. *Genes Dev.* **2007**, *21*, 1603–1608.
23. Kirino, Y.; Mourelatos, Z. Mouse piwi-interacting RNAs are 2'-*O*-methylated at their 3' termini. *Nat. Struct. Mol. Biol.* **2007**, *14*, 347–348.
24. Nishimasu, H.; Ishizu, H.; Saito, K.; Fukuhara, S.; Kamatani, M.K.; Bonnefond, L.; Matsumoto, N.; Nishizawa, T.; Nakanaga, K.; Aoki, J.; *et al.* Structure and function of zucchini endoribonuclease in piRNA biogenesis. *Nature* **2012**, *491*, 284–287.
25. Saito, K.; Ishizu, H.; Komai, M.; Kotani, H.; Kawamura, Y.; Nishida, K.M.; Siomi, H.; Siomi, M.C. Roles for the Yb body components armitage and Yb in primary piRNA biogenesis in drosophila. *Genes Dev.* **2010**, *24*, 2493–2498.
26. Le Thomas, A.; Rogers, A.K.; Webster, A.; Marinov, G.K.; Liao, S.E.; Perkins, E.M.; Hur, J.K.; Aravin, A.A.; Toth, K.F. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev.* **2013**, *27*, 390–399.
27. Rozhkov, N.V.; Hammell, M.; Hannon, G.J. Multiple roles for piwi in silencing drosophila transposons. *Genes Dev.* **2013**, *27*, 400–412.
28. Lee, E.J.; Banerjee, S.; Zhou, H.; Jammalamadaka, A.; Arcila, M.; Manjunath, B.S.; Kosik, K.S. Identification of piRNAs in the central nervous system. *RNA* (*New York, N.Y.*) **2011**, *17*, 1090–1099.
29. Yan, Z.; Hu, H.Y.; Jiang, X.; Maierhofer, V.; Neb, E.; He, L.; Hu, Y.; Hu, H.; Li, N.; Chen, W.; *et al.* Widespread expression of piRNA-like molecules in somatic tissues. *Nucleic Acids Res.* **2011**, *39*, 6596–6607.

30. Lu, Y.; Li, C.; Zhang, K.; Sun, H.; Tao, D.; Liu, Y.; Zhang, S.; Ma, Y. Identification of piRNAs in hela cells by massive parallel sequencing. *BMB Rep.***2010**, *43*, 635–641.
31. Cheng, J.; Guo, J.M.; Xiao, B.X.; Miao, Y.; Jiang, Z.; Zhou, H.; Li, Q.N. PiRNA, the new non-coding RNA, is aberrantly expressed in human cancer cells. *Clin. Chim. Acta* **2011**, *412*, 1621–1625.
32. Qiao, D.; Zeeman, A.M.; Deng, W.; Looijenga, L.H.; Lin, H. Molecular characterization of hiwi, a human member of the *piwi* gene family whose overexpression is correlated to seminomas. *Oncogene* **2002**, *21*, 3988–3999.
33. Grochola, L.F.; Greither, T.; Taubert, H.; Moller, P.; Knippschild, U.; Udelnow, A.; Henne-Bruns, D.; Wurl, P. The stem cell-associated *hiwi* gene in human adenocarcinoma of the pancreas: Expression and risk of tumour-related death. *Br. J. Cancer* **2008**, *99*, 1083–1088.
34. He, W.; Wang, Z.; Wang, Q.; Fan, Q.; Shou, C.; Wang, J.; Giercksky, K.E.; Nesland, J.M.; Suo, Z. Expression of *hiwi* in human esophageal squamous cell carcinoma is significantly associated with poorer prognosis. *BMC Cancer* **2009**, *9*, 426.
35. Liu, W.K.; Jiang, X.Y.; Zhang, Z.X. Expression of *psca*, *piwil1*, and *tbx2* in endometrial adenocarcinoma. *Onkologie* **2010**, *33*, 241–245.
36. Liu, W.K.; Jiang, X.Y.; Zhang, Z.X. Expression of *psca*, *piwil1* and *tbx2* and its correlation with *hpv16* infection in formalin-fixed, paraffin-embedded cervical squamous cell carcinoma specimens. *Arch. Virol.* **2010**, *155*, 657–663.
37. Sun, G.; Wang, Y.; Sun, L.; Luo, H.; Liu, N.; Fu, Z.; You, Y. Clinical significance of *hiwi* gene expression in gliomas. *Brain Res.* **2011**, *1373*, 183–188.
38. Zeng, Y.; Qu, L.K.; Meng, L.; Liu, C.Y.; Dong, B.; Xing, X.F.; Wu, J.; Shou, C.C. *Hiwi* expression profile in cancer cells and its prognostic value for patients with colorectal cancer. *Chin. Med. J.* **2011**, *124*, 2144–2149.
39. Zhao, Y.M.; Zhou, J.M.; Wang, L.R.; He, H.W.; Wang, X.L.; Tao, Z.H.; Sun, H.C.; Wu, W.Z.; Fan, J.; Tang, Z.Y.; *et al. Hiwi* is associated with prognosis in patients with hepatocellular carcinoma after curative resection. *Cancer* **2012**, *118*, 2708–2717.
40. Cheng, J.; Deng, H.; Xiao, B.; Zhou, H.; Zhou, F.; Shen, Z.; Guo, J. *PiR-823*, a novel non-coding small RNA, demonstrates *in vitro* and *in vivo* tumor suppressive activity in human gastric cancer cells. *Cancer Lett.* **2012**, *315*, 12–17.
41. Cui, L.; Lou, Y.; Zhang, X.; Zhou, H.; Deng, H.; Song, H.; Yu, X.; Xiao, B.; Wang, W.; Guo, J. Detection of circulating tumor cells in peripheral blood from patients with gastric cancer using piRNAs as markers. *Clin. Biochem.* **2011**, *44*, 1050–1057.
42. Lee, R.C.; Feinbaum, R.L.; Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **1993**, *75*, 843–854.
43. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297.
44. Denli, A.M.; Tops, B.B.; Plasterk, R.H.; Ketting, R.F.; Hannon, G.J. Processing of primary microRNAs by the microprocessor complex. *Nature* **2004**, *432*, 231–235.
45. Lagos-Quintana, M.; Rauhut, R.; Lendeckel, W.; Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **2001**, *294*, 853–858.

46. Lee, R.C.; Ambros, V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **2001**, *294*, 862–864.
47. Hu, W.; Chan, C.S.; Wu, R.; Zhang, C.; Sun, Y.; Song, J.S.; Tang, L.H.; Levine, A.J.; Feng, Z. Negative regulation of tumor suppressor p53 by microRNA *miR-504*. *Mol. Cell* **2010**, *38*, 689–699.
48. Bagga, S.; Bracht, J.; Hunter, S.; Massirer, K.; Holtz, J.; Eachus, R.; Pasquinelli, A.E. Regulation by *let-7* and *lin-4* miRNAs results in target mRNA degradation. *Cell* **2005**, *122*, 553–563.
49. Harfe, B.D. MicroRNAs in vertebrate development. *Curr. Opin. Genet. Dev.* **2005**, *15*, 410–415.
50. Boehm, M.; Slack, F.J. MicroRNA control of lifespan and metabolism. *Cell Cycle* **2006**, *5*, 837–840.
51. Calin, G.A.; Garzon, R.; Cimmino, A.; Fabbri, M.; Croce, C.M. MicroRNAs and leukemias: How strong is the connection? *Leuk. Res.* **2006**, *30*, 653–655.
52. Arisawa, T.; Tahara, T.; Shibata, T.; Nagasaka, M.; Nakamura, M.; Kamiya, Y.; Fujita, H.; Hasegawa, S.; Takagi, T.; Wang, F.Y.; *et al.* A polymorphism of *microRNA 27a* genome region is associated with the development of gastric mucosal atrophy in japanese male subjects. *Dig. Dis. Sci.* **2007**, *52*, 1691–1697.
53. Carleton, M.; Cleary, M.A.; Linsley, P.S. MicroRNAs and cell cycle regulation. *Cell Cycle* **2007**, *6*, 2127–2132.
54. Lee, Y.; Kim, M.; Han, J.; Yeom, K.H.; Lee, S.; Baek, S.H.; Kim, V.N. MicroRNA genes are transcribed by RNA polymerase ii. *EMBO* **2004**, *23*, 4051–4060.
55. Cai, X.; Hagedorn, C.H.; Cullen, B.R. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* (*New York*, *N.Y.*) **2004**, *10*, 1957–1966.
56. Ambros, V. The functions of animal microRNAs. *Nature* **2004**, *431*, 350–355.
57. Bohnsack, M.T.; Czaplinski, K.; Gorlich, D. Exportin 5 is a rangtp-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* (*New York*, *N.Y.*) **2004**, *10*, 185–191.
58. Esquela-Kerscher, A.; Johnson, S.M.; Bai, L.; Saito, K.; Partridge, J.; Reinert, K.L.; Slack, F.J. Post-embryonic expression of *C. elegans* microRNAs belonging to the *lin-4* and *let-7* families in the hypodermis and the reproductive system. *Dev. Dyn.* **2005**, *234*, 868–877.
59. Achard, P.; Herr, A.; Baulcombe, D.C.; Harberd, N.P. Modulation of floral development by a gibberellin-regulated microRNA. *Development* **2004**, *131*, 3357–3365.
60. Gregory, R.I.; Chendrimada, T.P.; Shiekhattar, R. MicroRNA biogenesis: Isolation and characterization of the microprocessor complex. *Methods Mol. Biol.* **2006**, *342*, 33–47.
61. Calin, G.A.; Dumitru, C.D.; Shimizu, M.; Bichi, R.; Zupo, S.; Noch, E.; Aldler, H.; Rattan, S.; Keating, M.; Rai, K.; *et al.* Frequent deletions and down-regulation of micro-RNA genes *miR15* and *miR16* at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15524–15529.
62. Calin, G.A.; Ferracin, M.; Cimmino, A.; Di Leva, G.; Shimizu, M.; Wojcik, S.E.; Iorio, M.V.; Visone, R.; Sever, N.I.; Fabbri, M.; *et al.* A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N. Engl. J. Med.* **2005**, *353*, 1793–1801.
63. Nakamura, T.; Canaani, E.; Croce, C.M. Oncogenic all1 fusion proteins target drosha-mediated microRNA processing. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10980–10985.

64. Saito, Y.; Liang, G.; Egger, G.; Friedman, J.M.; Chuang, J.C.; Coetzee, G.A.; Jones, P.A. Specific activation of *microRNA-127* with downregulation of the proto-oncogene *bcl6* by chromatin-modifying drugs in human cancer cells. *Cancer Cell* **2006**, *9*, 435–443.
65. Fabbri, M.; Croce, C.M.; Calin, G.A. MicroRNAs. *Cancer J.* **2008**, *14*, 1–6.
66. Garzon, R.; Volinia, S.; Liu, C.G.; Fernandez-Cymering, C.; Palumbo, T.; Pichiorri, F.; Fabbri, M.; Coombes, K.; Alder, H.; Nakamura, T.; *et al.* MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia. *Blood* **2008**, *111*, 3183–3189.
67. Ciafre, S.A.; Galardi, S.; Mangiola, A.; Ferracin, M.; Liu, C.G.; Sabatino, G.; Negrini, M.; Maira, G.; Croce, C.M.; Farace, M.G. Extensive modulation of a set of microRNAs in primary glioblastoma. *Biochem. Biophys. Res. Commun.* **2005**, *334*, 1351–1358.
68. Volinia, S.; Calin, G.A.; Liu, C.G.; Ambs, S.; Cimmino, A.; Petrocca, F.; Visone, R.; Iorio, M.; Roldo, C.; Ferracin, M.; *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 2257–2261.
69. Asangani, I.A.; Rasheed, S.A.; Nikolova, D.A.; Leupold, J.H.; Colburn, N.H.; Post, S.; Allgayer, H. *MicroRNA-21* (*miR-21*) post-transcriptionally downregulates tumor suppressor pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene* **2008**, *27*, 2128–2136.
70. Zhang, J.G.; Wang, J.J.; Zhao, F.; Liu, Q.; Jiang, K.; Yang, G.H. *MicroRNA-21* (*miR-21*) represses tumor suppressor pten and promotes growth and invasion in non-small cell lung cancer (nsclc). *Clin. Chim. Acta* **2010**, *411*, 846–852.
71. Quintavalle, C.; Donnarumma, E.; Iaboni, M.; Roscigno, G.; Garofalo, M.; Romano, G.; Fiore, D.; De Marinis, P.; Croce, C.M.; Condorelli, G. Effect of *miR-21* and *miR-30b/c* on trail-induced apoptosis in glioma cells. *Oncogene* **2012**, doi:10.1038/onc.2012.410.
72. Kluiver, J.; Poppema, S.; de Jong, D.; Blokzijl, T.; Harms, G.; Jacobs, S.; Kroesen, B.J.; van den Berg, A. *Bic* and *miR-155* are highly expressed in hodgkin, primary mediastinal and diffuse large B cell lymphomas. *J. Pathol.* **2005**, *207*, 243–249.
73. Calin, G.A.; Cimmino, A.; Fabbri, M.; Ferracin, M.; Wojcik, S.E.; Shimizu, M.; Taccioli, C.; Zanesi, N.; Garzon, R.; Aqeilan, R.I.; *et al. Mir-15a* and *miR-16-1* cluster functions in human leukemia. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5166–5171.
74. Jiang, S.; Zhang, H.W.; Lu, M.H.; He, X.H.; Li, Y.; Gu, H.; Liu, M.F.; Wang, E.D. *MicroRNA-155* functions as an oncomir in breast cancer by targeting the suppressor of cytokine signaling 1 gene. *Cancer Res.* **2010**, *70*, 3119–3127.
75. Ovcharenko, D.; Kelnar, K.; Johnson, C.; Leng, N.; Brown, D. Genome-scale microRNA and small interfering RNA screens identify small RNA modulators of trail-induced apoptosis pathway. *Cancer Res.* **2007**, *67*, 10782–10788.
76. Costinean, S.; Zanesi, N.; Pekarsky, Y.; Tili, E.; Volinia, S.; Heerema, N.; Croce, C.M. Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in *e(mu)-miR155* transgenic mice. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 7024–7029.
77. Fornari, F.; Gramantieri, L.; Ferracin, M.; Veronese, A.; Sabbioni, S.; Calin, G.A.; Grazi, G.L.; Giovannini, C.; Croce, C.M.; Bolondi, L.; *et al. Mir-221* controls *cdkn1c/p57* and *cdkn1b/p27* expression in human hepatocellular carcinoma. *Oncogene* **2008**, *27*, 5651–5661.

78. Di Leva, G.; Gasparini, P.; Piovan, C.; Ngankeu, A.; Garofalo, M.; Taccioli, C.; Iorio, M.V.; Li, M.; Volinia, S.; Alder, H.; *et al.* MicroRNA cluster *221–222* and estrogen receptor alpha interactions in breast cancer. *J. Natl. Cancer Inst.* **2010**, *102*, 706–721.
79. Felicetti, F.; Errico, M.C.; Bottero, L.; Segnalini, P.; Stoppacciaro, A.; Biffoni, M.; Felli, N.; Mattia, G.; Petrini, M.; Colombo, M.P.; *et al.* The promyelocytic leukemia zinc finger-*microRNA-221/-222* pathway controls melanoma progression through multiple oncogenic mechanisms. *Cancer Res.* **2008**, *68*, 2745–2754.
80. Pallante, P.; Visone, R.; Ferracin, M.; Ferraro, A.; Berlingieri, M.T.; Troncone, G.; Chiappetta, G.; Liu, C.G.; Santoro, M.; Negrini, M.; *et al.* MicroRNA deregulation in human thyroid papillary carcinomas. *Endovr.-Relat. Cancer* **2006**, *13*, 497–508.
81. le Sage, C.; Nagel, R.; Egan, D.A.; Schrier, M.; Mesman, E.; Mangiola, A.; Anile, C.; Maira, G.; Mercatelli, N.; Ciafre, S.A.; *et al.* Regulation of the p27(kip1) tumor suppressor by *miR-221* and *miR-222* promotes cancer cell proliferation. *EMBO* **2007**, *26*, 3699–3708.
82. Garofalo, M.; Quintavalle, C.; Di Leva, G.; Zanca, C.; Romano, G.; Taccioli, C.; Liu, C.G.; Croce, C.M.; Condorelli, G. MicroRNA signatures of trail resistance in human non-small cell lung cancer. *Oncogene* **2008**, *27*, 3845–3855.
83. Quintavalle, C.; Garofalo, M.; Zanca, C.; Romano, G.; Iaboni, M.; del Basso De Caro, M.; Martinez-Montero, J.C.; Incoronato, M.; Nuovo, G.; Croce, C.M.; *et al. MiR-221/222* overexpession in human glioblastoma increases invasiveness by targeting the protein phosphate ptpmu. *Oncogene* **2012**, *31*, 858–868.
84. Garofalo, M.; Di Leva, G.; Romano, G.; Nuovo, G.; Suh, S.S.; Ngankeu, A.; Taccioli, C.; Pichiorri, F.; Alder, H.; Secchiero, P.; *et al. Mir-221 & 222* regulate trail resistance and enhance tumorigenicity through pten and timp3 downregulation. *Cancer Cell* **2009**, *16*, 498–509.
85. Manni, I.; Artuso, S.; Careccia, S.; Rizzo, M.G.; Baserga, R.; Piaggio, G.; Sacchi, A. The microRNA *miR-92* increases proliferation of myeloid cells and by targeting p63 modulates the abundance of its isoforms. *FASEB* **2009**, *23*, 3957–3966.
86. Voorhoeve, P.M.; le Sage, C.; Schrier, M.; Gillis, A.J.; Stoop, H.; Nagel, R.; Liu, Y.P.; van Duijse, J.; Drost, J.; Griekspoor, A.; *et al.* A genetic screen implicates *miRNA-372* and *miRNA-373* as oncogenes in testicular germ cell tumors. *Cell* **2006**, *124*, 1169–1181.
87. Ma, L.; Teruya-Feldstein, J.; Weinberg, R.A. Tumour invasion and metastasis initiated by *microRNA-10b* in breast cancer. *Nature* **2007**, *449*, 682–688.
88. Petrocca, F.; Visone, R.; Onelli, M.R.; Shah, M.H.; Nicoloso, M.S.; de Martino, I.; Iliopoulos, D.; Pilozzi, E.; Liu, C.G.; Negrini, M.; *et al.* E2f1-regulated microRNAs impair tgfbeta-dependent cell-cycle arrest and apoptosis in gastric cancer. *Cancer Cell* **2008**, *13*, 272–286.
89. Cimmino, A.; Calin, G.A.; Fabbri, M.; Iorio, M.V.; Ferracin, M.; Shimizu, M.; Wojcik, S.E.; Aqeilan, R.I.; Zupo, S.; Dono, M.; *et al. MiR-15* and *miR-16* induce apoptosis by targeting *bcl2*. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13944–13949.
90. Bonci, D.; Coppola, V.; Musumeci, M.; Addario, A.; Giuffrida, R.; Memeo, L.; D'Urso, L.; Pagliuca, A.; Biffoni, M.; Labbaye, C.; *et al. The miR-15a-miR-16-1* cluster controls prostate cancer by targeting multiple oncogenic activities. *Nat. Med.* **2008**, *14*, 1271–1277.

91. Bhattacharya, R.; Nicoloso, M.; Arvizo, R.; Wang, E.; Cortez, A.; Rossi, S.; Calin, G.A.; Mukherjee, P. *MiR-15a* and *miR-16* control *bmi-1* expression in ovarian cancer. *Cancer Res.* **2009**, *69*, 9090–9095.
92. Sun, C.Y.; She, X.M.; Qin, Y.; Chu, Z.B.; Chen, L.; Ai, L.S.; Zhang, L.; Hu, Y. *MiR-15a* and *miR-16* affect the angiogenesis of multiple myeloma by targeting *vegf*. *Carcinogenesis* **2013**, *34*, 426–435.
93. Pouliot, L.M.; Chen, Y.C.; Bai, J.; Guha, R.; Martin, S.E.; Gottesman, M.M.; Hall, M.D. Cisplatin sensitivity mediated by wee1 and chk1 is mediated by *miR-155* and the *miR-15* family. *Cancer Res.* **2012**, *72*, 5945–5955.
94. Takamizawa, J.; Konishi, H.; Yanagisawa, K.; Tomida, S.; Osada, H.; Endoh, H.; Harano, T.; Yatabe, Y.; Nagino, M.; Nimura, Y.; *et al.* Reduced expression of the *let-7* microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res.* **2004**, *64*, 3753–3756.
95. Zhang, H.H.; Wang, X.J.; Li, G.X.; Yang, E.; Yang, N.M. Detection of *let-7a* microRNA by real-time PCR in gastric carcinoma. *World J. Gastroenterol.* **2007**, *13*, 2883–2888.
96. Akao, Y.; Nakagawa, Y.; Naoe, T. *Let-7* microRNA functions as a potential growth suppressor in human colon cancer cells. *Biol. Pharm. Bull.* **2006**, *29*, 903–906.
97. Sampson, V.B.; Rong, N.H.; Han, J.; Yang, Q.; Aris, V.; Soteropoulos, P.; Petrelli, N.J.; Dunn, S.P.; Krueger, L.J. MicroRNA *let-7a* down-regulates *myc* and reverts *myc*-induced growth in burkitt lymphoma cells. *Cancer Res.* **2007**, *67*, 9762–9770.
98. Johnson, S.M.; Grosshans, H.; Shingara, J.; Byrom, M.; Jarvis, R.; Cheng, A.; Labourier, E.; Reinert, K.L.; Brown, D.; Slack, F.J. Ras is regulated by the *let-7* microRNA family. *Cell* **2005**, *120*, 635–647.
99. Mayr, C.; Hemann, M.T.; Bartel, D.P. Disrupting the pairing between *let-7* and *hmga2* enhances oncogenic transformation. *Science* **2007**, *315*, 1576–1579.
100. Wang, Y.; Lu, Y.; Toh, S.T.; Sung, W.K.; Tan, P.; Chow, P.; Chung, A.Y.; Jooi, L.L.; Lee, C.G. *Lethal-7* is down-regulated by the hepatitis B virus x protein and targets signal transducer and activator of transcription 3. *J. Hepatol.* **2010**, *53*, 57–66.
101. Hu, X.; Guo, J.; Zheng, L.; Li, C.; Zheng, T.M.; Tanyi, J.L.; Liang, S.; Benedetto, C.; Mitidieri, M.; Katsaros, D.; *et al.* The heterochronic microRNA *let-7* inhibits cell motility by regulating the genes in the actin cytoskeleton pathway in breast cancer. *Mol. Cancer Res.* **2013**, *11*, 240–250.
102. Wiggins, J.F.; Ruffino, L.; Kelnar, K.; Omotola, M.; Patrawala, L.; Brown, D.; Bader, A.G. Development of a lung cancer therapeutic based on the tumor suppressor *microRNA-34*. *Cancer Res.* **2010**, *70*, 5923–5930.
103. Li, N.; Fu, H.; Tie, Y.; Hu, Z.; Kong, W.; Wu, Y.; Zheng, X. *Mir-34a* inhibits migration and invasion by down-regulation of c-met expression in human hepatocellular carcinoma cells. *Cancer Lett.* **2009**, *275*, 44–53.
104. Lodygin, D.; Tarasov, V.; Epanchintsev, A.; Berking, C.; Knyazeva, T.; Korner, H.; Knyazev, P.; Diebold, J.; Hermeking, H. Inactivation of *miR-34a* by aberrant CpG methylation in multiple types of cancer. *Cell Cycle* **2008**, *7*, 2591–2600.

105. Tazawa, H.; Tsuchiya, N.; Izumiya, M.; Nakagama, H. Tumor-suppressive *miR-34a* induces senescence-like growth arrest through modulation of the e2f pathway in human colon cancer cells. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15472–15477.
106. Welch, C.; Chen, Y.; Stallings, R.L. *MicroRNA-34a* functions as a potential tumor suppressor by inducing apoptosis in neuroblastoma cells. *Oncogene* **2007**, *26*, 5017–5022.
107. Corney, D.C.; Hwang, C.I.; Matoso, A.; Vogt, M.; Flesken-Nikitin, A.; Godwin, A.K.; Kamat, A.A.; Sood, A.K.; Ellenson, L.H.; Hermeking, H.; *et al.* Frequent downregulation of *miR-34* family in human ovarian cancers. *Clin. Cancer Res.* **2010**, *16*, 1119–1128.
108. Chen, X.; Hu, H.; Guan, X.; Xiong, G.; Wang, Y.; Wang, K.; Li, J.; Xu, X.; Yang, K.; Bai, Y. CpG island methylation status of miRNAs in esophageal squamous cell carcinoma. *Int. J. Cancer* **2012**, *130*, 1607–1613.
109. Mraz, M.; Malinova, K.; Kotaskova, J.; Pavlova, S.; Tichy, B.; Malcikova, J.; Stano Kozubik, K.; Smardova, J.; Brychtova, Y.; Doubek, M.; *et al. Mir-34a*, *miR-29c* and *miR-17–5p* are downregulated in cll patients with tp53 abnormalities. *Leukemia* **2009**, *23*, 1159–1163.
110. He, L.; He, X.; Lim, L.P.; de Stanchina, E.; Xuan, Z.; Liang, Y.; Xue, W.; Zender, L.; Magnus, J.; Ridzon, D.; *et al.* A microRNA component of the p53 tumour suppressor network. *Nature* **2007**, *447*, 1130–1134.
111. Garofalo, M.; Jeon, Y.J.; Nuovo, G.J.; Middleton, J.; Secchiero, P.; Joshi, P.; Alder, H.; Nazaryan, N.; Di Leva, G.; Romano, G.; *et al. Mir-34a/c*-dependent pdgfr-alpha/beta downregulation inhibits tumorigenesis and enhances trail-induced apoptosis in lung cancer. *PLoS One* **2013**, *8*, e67581.
112. Park, S.M.; Gaur, A.B.; Lengyel, E.; Peter, M.E. The *miR-200* family determines the epithelial phenotype of cancer cells by targeting the e-cadherin repressors zeb1 and zeb2. *Genes Dev.* **2008**, *22*, 894–907.
113. Liu, S.; Tetzlaff, M.T.; Cui, R.; Xu, X. *Mir-200c* inhibits melanoma progression and drug resistance through down-regulation of *bmi-1*. *Am. J. Pathol.* **2012**, *181*, 1823–1835.
114. Iorio, M.V.; Ferracin, M.; Liu, C.G.; Veronese, A.; Spizzo, R.; Sabbioni, S.; Magri, E.; Pedriali, M.; Fabbri, M.; Campiglio, M.; *et al.* MicroRNA gene expression deregulation in human breast cancer. *Cancer Res.* **2005**, *65*, 7065–7070.
115. Li, Y.; Wang, F.; Xu, J.; Ye, F.; Shen, Y.; Zhou, J.; Lu, W.; Wan, X.; Ma, D.; Xie, X. Progressive miRNA expression profiles in cervical carcinogenesis and identification of *hpv*-related target genes for *miR-29*. *J. Pathol.* **2011**, *224*, 484–495.
116. Xiong, Y.; Fang, J.H.; Yun, J.P.; Yang, J.; Zhang, Y.; Jia, W.H.; Zhuang, S.M. Effects of *microRNA-29* on apoptosis, tumorigenicity, and prognosis of hepatocellular carcinoma. *Hepatology* **2010**, *51*, 836–845.
117. Fabbri, M.; Garzon, R.; Cimmino, A.; Liu, Z.; Zanesi, N.; Callegari, E.; Liu, S.; Alder, H.; Costinean, S.; Fernandez-Cymering, C.; *et al. MicroRNA-29* family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3a and 3b. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15805–15810.
118. Pekarsky, Y.; Santanam, U.; Cimmino, A.; Palamarchuk, A.; Efanov, A.; Maximov, V.; Volinia, S.; Alder, H.; Liu, C.G.; Rassenti, L.; *et al. Tcl1* expression in chronic lymphocytic leukemia is regulated by *miR-29* and *miR-181*. *Cancer Res.* **2006**, *66*, 11590–11593.

119. Mott, J.L.; Kobayashi, S.; Bronk, S.F.; Gores, G.J. *Mir-29* regulates mcl-1 protein expression and apoptosis. *Oncogene* **2007**, *26*, 6133–6140.
120. Zhang, X.; Zhao, X.; Fiskus, W.; Lin, J.; Lwin, T.; Rao, R.; Zhang, Y.; Chan, J.C.; Fu, K.; Marquez, V.E.; *et al.* Coordinated silencing of *myc*-mediated *miR-29* by hdac3 and ezh2 as a therapeutic target of histone modification in aggressive B-cell lymphomas. *Cancer Cell* **2012**, *22*, 506–523.
121. Smith, C.M.; Steitz, J.A. Sno storm in the nucleolus: New roles for myriad small rnps. *Cell* **1997**, *89*, 669–672.
122. Bortolin, M.L.; Kiss, T. Human u19 intron-encoded snoRNA is processed from a long primary transcript that possesses little potential for protein coding. *RNA* (*New York*, *N.Y.*) **1998**, *4*, 445–454.
123. Kiss-Laszlo, Z.; Henry, Y.; Bachellerie, J.P.; Caizergues-Ferrer, M.; Kiss, T. Site-specific ribose methylation of preribosomal RNA: A novel function for small nucleolar RNAs. *Cell* **1996**, *85*, 1077–1088.
124. Lafontaine, D.L.; Tollervey, D. Birth of the snornps: The evolution of the modification-guide snoRNAs. *Trends Biochem. Sci.* **1998**, *23*, 383–388.
125. Terns, M.P.; Terns, R.M. Small nucleolar RNAs: Versatile trans-acting molecules of ancient evolutionary origin. *Gene Expr.* **2002**, *10*, 17–39.
126. Weinstein, L.B.; Steitz, J.A. Guided tours: From precursor snoRNA to functional snornp. *Curr. Opin. Cell Biol.* **1999**, *11*, 378–384.
127. Kiss, A.M.; Jady, B.E.; Bertrand, E.; Kiss, T. Human box h/aca pseudouridylation guide RNA machinery. *Mol. Cell. Biol.* **2004**, *24*, 5797–5807.
128. Bachellerie, J.P.; Cavaille, J.; Huttenhofer, A. The expanding snoRNA world. *Biochimie* **2002**, *84*, 775–790.
129. Coughlin, D.J.; Pleiss, J.A.; Walker, S.C.; Whitworth, G.B.; Engelke, D.R. Genome-wide search for yeast RNase p substrates reveals role in maturation of intron-encoded box c/d small nucleolar RNAs. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 12218–12223.
130. Brameier, M.; Herwig, A.; Reinhardt, R.; Walter, L.; Gruber, J. Human box c/d snoRNAs with miRNA like functions: Expanding the range of regulatory RNAs. *Nucleic Acids Res.* **2011**, *39*, 675–686.
131. Li, R.; Wang, H.; Bekele, B.N.; Yin, Z.; Caraway, N.P.; Katz, R.L.; Stass, S.A.; Jiang, F. Identification of putative oncogenes in lung adenocarcinoma by a comprehensive functional genomic approach. *Oncogene* **2006**, *25*, 2628–2635.
132. Schneider, C.; King, R.M.; Philipson, L. Genes specifically expressed at growth arrest of mammalian cells. *Cell* **1988**, *54*, 787–793.
133. Mourtada-Maarabouni, M.; Pickard, M.R.; Hedge, V.L.; Farzaneh, F.; Williams, G.T. Gas5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene* **2009**, *28*, 195–208.
134. Nakamura, Y.; Takahashi, N.; Kakegawa, E.; Yoshida, K.; Ito, Y.; Kayano, H.; Niitsu, N.; Jinnai, I.; Bessho, M. The *gas5* (growth arrest-specific transcript 5) gene fuses to bcl6 as a result of t (1; 3) (q25; q27) in a patient with B-cell lymphoma. *Cancer Genet. Cytogenet.* **2008**, *182*, 144–149.

135. Gee, H.E.; Buffa, F.M.; Camps, C.; Ramachandran, A.; Leek, R.; Taylor, M.; Patil, M.; Sheldon, H.; Betts, G.; Homer, J.; *et al.* The small-nucleolar RNAs commonly used for microRNA normalisation correlate with tumour pathology and prognosis. *Br. J. Cancer* **2011**, *104*, 1168–1177.
136. Tanaka, R.; Satoh, H.; Moriyama, M.; Satoh, K.; Morishita, Y.; Yoshida, S.; Watanabe, T.; Nakamura, Y.; Mori, S. Intronic u50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t (3; 6) (q27; q15) of human B-cell lymphoma. *Genes Cells* **2000**, *5*, 277–287.
137. Mei, Y.P.; Liao, J.P.; Shen, J.; Yu, L.; Liu, B.L.; Liu, L.; Li, R.Y.; Ji, L.; Dorsey, S.G.; Jiang, Z.R.; *et al. Small nucleolar RNA 42* acts as an oncogene in lung tumorigenesis. *Oncogene* **2012**, *31*, 2794–2804.
138. Liao, J.; Yu, L.; Mei, Y.; Guarnera, M.; Shen, J.; Li, R.; Liu, Z.; Jiang, F. Small nucleolar RNA signatures as biomarkers for non-small-cell lung cancer. *Mol. Cancer* **2010**, *9*, 198.
139. Ruggero, D.; Grisendi, S.; Piazza, F.; Rego, E.; Mari, F.; Rao, P.H.; Cordon-Cardo, C.; Pandolfi, P.P. Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. *Science* **2003**, *299*, 259–262.
140. Cabili, M.N.; Trapnell, C.; Goff, L.; Koziol, M.; Tazon-Vega, B.; Regev, A.; Rinn, J.L. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **2011**, *25*, 1915–1927.
141. Orom, U.A.; Derrien, T.; Beringer, M.; Gumireddy, K.; Gardini, A.; Bussotti, G.; Lai, F.; Zytnicki, M.; Notredame, C.; Huang, Q.; *et al.* Long noncoding RNAs with enhancer-like function in human cells. *Cell* **2010**, *143*, 46–58.
142. Ponting, C.P.; Oliver, P.L.; Reik, W. Evolution and functions of long noncoding RNAs. *Cell* **2009**, *136*, 629–641.
143. Khalil, A.M.; Guttman, M.; Huarte, M.; Garber, M.; Raj, A.; Rivea Morales, D.; Thomas, K.; Presser, A.; Bernstein, B.E.; van Oudenaarden, A.; *et al.* Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11667–11672.
144. Guttman, M.; Amit, I.; Garber, M.; French, C.; Lin, M.F.; Feldser, D.; Huarte, M.; Zuk, O.; Carey, B.W.; Cassady, J.P.; *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **2009**, *458*, 223–227.
145. Birney, E.; Stamatoyannopoulos, J.A.; Dutta, A.; Guigo, R.; Gingeras, T.R.; Margulies, E.H.; Weng, Z.; Snyder, M.; Dermitzakis, E.T.; Thurman, R.E.; *et al.* Identification and analysis of functional elements in 1% of the human genome by the encode pilot project. *Nature* **2007**, *447*, 799–816.
146. Brosnan, C.A.; Voinnet, O. The long and the short of noncoding RNAs. *Curr. Opin. Cell Biol.* **2009**, *21*, 416–425.
147. Amaral, P.P.; Mattick, J.S. Noncoding RNA in development. *Mamm. Genome* **2008**, *19*, 454–492.
148. Pang, K.C.; Frith, M.C.; Mattick, J.S. Rapid evolution of noncoding RNAs: Lack of conservation does not mean lack of function. *Trends Genet.* **2006**, *22*, 1–5.

149. Chen, L.L.; Carmichael, G.G. Decoding the function of nuclear long non-coding RNAs. *Curr. Opin. Cell Biol.* **2010**, *22*, 357–364.
150. Amaral, P.P.; Clark, M.B.; Gascoigne, D.K.; Dinger, M.E.; Mattick, J.S. Lncrnadb: A reference database for long noncoding RNAs. *Nucleic Acids Res.* **2011**, *39*, D146–D151.
151. Dinger, M.E.; Pang, K.C.; Mercer, T.R.; Crowe, M.L.; Grimmond, S.M.; Mattick, J.S. NRED: A database of long noncoding RNA expression. *Nucleic Acids Res.* **2009**, *37*, D122–D126.
152. Derrien, T.; Johnson, R.; Bussotti, G.; Tanzer, A.; Djebali, S.; Tilgner, H.; Guernec, G.; Martin, D.; Merkel, A.; Knowles, D.G.; *et al.* The gencode v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.* **2012**, *22*, 1775–1789.
153. Wang, X.; Arai, S.; Song, X.; Reichart, D.; Du, K.; Pascual, G.; Tempst, P.; Rosenfeld, M.G.; Glass, C.K.; Kurokawa, R. Induced ncRNAs allosterically modify RNA-binding proteins in *cis* to inhibit transcription. *Nature* **2008**, *454*, 126–130.
154. Schwartz, J.C.; Younger, S.T.; Nguyen, N.B.; Hardy, D.B.; Monia, B.P.; Corey, D.R.; Janowski, B.A. Antisense transcripts are targets for activating small RNAs. *Nat. Struct. Mol. Biol.* **2008**, *15*, 842–848.
155. Tasheva, E.S.; Roufa, D.J. Regulation of human *rps14* transcription by intronic antisense RNAs and ribosomal protein s14. *Genes Dev.* **1995**, *9*, 304–316.
156. Kino, T.; Hurt, D.E.; Ichijo, T.; Nader, N.; Chrousos, G.P. Noncoding RNA *gas5* is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Science Signal.* **2010**, *3*, doi:10.1126/scisignal.2000568.
157. Lin, D.; Pestova, T.V.; Hellen, C.U.; Tiedge, H. Translational control by a small RNA: Dendritic bc1 RNA targets the eukaryotic initiation factor 4a helicase mechanism. *Mol. Cell. Biol.* **2008**, *28*, 3008–3019.
158. Young, T.L.; Matsuda, T.; Cepko, C.L. The noncoding RNA taurine upregulated gene 1 is required for differentiation of the murine retina. *Curr. Biol.* **2005**, *15*, 501–512.
159. Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914.
160. Yochum, G.S.; Cleland, R.; McWeeney, S.; Goodman, R.H. An antisense transcript induced by wnt/beta-catenin signaling decreases e2f4. *J. Biol. Chem.* **2007**, *282*, 871–878.
161. Mourtada-Maarabouni, M.; Hedge, V.L.; Kirkham, L.; Farzaneh, F.; Williams, G.T. Growth arrest in human t-cells is controlled by the non-coding RNA growth-arrest-specific transcript 5 (*gas5*). *J. Cell Sci.* **2008**, *121*, 939–946.
162. Rinn, J.L.; Kertesz, M.; Wang, J.K.; Squazzo, S.L.; Xu, X.; Brugmann, S.A.; Goodnough, L.H.; Helms, J.A.; Farnham, P.J.; Segal, E.; *et al.* Functional demarcation of active and silent chromatin domains in human hox loci by noncoding RNAs. *Cell* **2007**, *129*, 1311–1323.
163. Yu, W.; Gius, D.; Onyango, P.; Muldoon-Jacobs, K.; Karp, J.; Feinberg, A.P.; Cui, H. Epigenetic silencing of tumour suppressor gene *p15* by its antisense RNA. *Nature* **2008**, *451*, 202–206.
164. Mercer, T.R.; Dinger, M.E.; Mattick, J.S. Long non-coding RNAs: Insights into functions. *Nat. Rev. Genet.* **2009**, *10*, 155–159.
165. Nagano, T.; Fraser, P. No-nonsense functions for long noncoding RNAs. *Cell* **2011**, *145*, 178–181.

166. Clark, M.B.; Mattick, J.S. Long noncoding RNAs in cell biology. *Semin. Cell Dev. Biol.* **2011**, *22*, 366–376.
167. Mattick, J.S.; Amaral, P.P.; Dinger, M.E.; Mercer, T.R.; Mehler, M.F. RNA regulation of epigenetic processes. *BioEssays* **2009**, *31*, 51–59.
168. Nagano, T.; Mitchell, J.A.; Sanz, L.A.; Pauler, F.M.; Ferguson-Smith, A.C.; Feil, R.; Fraser, P. The air noncoding RNA epigenetically silences transcription by targeting g9a to chromatin. *Science* **2008**, *322*, 1717–1720.
169. Pandey, R.R.; Mondal, T.; Mohammad, F.; Enroth, S.; Redrup, L.; Komorowski, J.; Nagano, T.; Mancini-Dinardo, D.; Kanduri, C. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell* **2008**, *32*, 232–246.
170. Prensner, J.R.; Iyer, M.K.; Balbin, O.A.; Dhanasekaran, S.M.; Cao, Q.; Brenner, J.C.; Laxman, B.; Asangani, I.A.; Grasso, C.S.; Kominsky, H.D.; *et al.* Transcriptome sequencing across a prostate cancer cohort identifies *pcat-1*, an unannotated lincRNA implicated in disease progression. *Nat. Biotechnol.* **2011**, *29*, 742–749.
171. Gupta, R.A.; Shah, N.; Wang, K.C.; Kim, J.; Horlings, H.M.; Wong, D.J.; Tsai, M.C.; Hung, T.; Argani, P.; Rinn, J.L.; *et al.* Long non-coding RNA hotair reprograms chromatin state to promote cancer metastasis. *Nature* **2010**, *464*, 1071–1076.
172. Lee, J.T. Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 815–826.
173. Zhao, J.; Sun, B.K.; Erwin, J.A.; Song, J.J.; Lee, J.T. Polycomb proteins targeted by a short repeat RNA to the mouse x chromosome. *Science* **2008**, *322*, 750–756.
174. Morris, K.V.; Santoso, S.; Turner, A.M.; Pastori, C.; Hawkins, P.G. Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. *PLoS Genet.* **2008**, *4*, e1000258.
175. Vanneste, R.; Smith, E.; Graham, G. Multiple neurofibromas as the presenting feature of familial atypical multiple malignant melanoma (fammm) syndrome. *Am. J. Med. Genet.* **2013**, *161*, 1425–1431.
176. Yap, K.L.; Li, S.; Munoz-Cabello, A.M.; Raguz, S.; Zeng, L.; Mujtaba, S.; Gil, J.; Walsh, M.J.; Zhou, M.M. Molecular interplay of the noncoding RNA anril and methylated histone h3 lysine 27 by polycomb cbx7 in transcriptional silencing of ink4a. *Mol. Cell* **2010**, *38*, 662–674.
177. Gutschner, T.; Hammerle, M.; Eissmann, M.; Hsu, J.; Kim, Y.; Hung, G.; Revenko, A.; Arun, G.; Stentrup, M.; Gross, M.; *et al.* The noncoding RNA *malat1* is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res.* **2013**, *73*, 1180–1189.
178. Pickard, M.R.; Mourtada-Maarabouni, M.; Williams, G.T. Long non-coding RNA *gas5* regulates apoptosis in prostate cancer cell lines. *Biochim. Biophys. Acta* **2013**, *1832*, 1613–1623.
179. Huarte, M.; Guttman, M.; Feldser, D.; Garber, M.; Koziol, M.J.; Kenzelmann-Broz, D.; Khalil, A.M.; Zuk, O.; Amit, I.; Rabani, M.; *et al.* A large intergenic noncoding RNA induced by *p53* mediates global gene repression in the *p53* response. *Cell* **2010**, *142*, 409–419.
180. Hung, T.; Wang, Y.; Lin, M.F.; Koegel, A.K.; Kotake, Y.; Grant, G.D.; Horlings, H.M.; Shah, N.; Umbricht, C.; Wang, P.; *et al.* Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat. Genet.* **2011**, *43*, 621–629.

181. Poliseno, L.; Salmena, L.; Zhang, J.; Carver, B.; Haveman, W.J.; Pandolfi, P.P. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **2010**, *465*, 1033–1038.
182. Melo, C.A.; Drost, J.; Wijchers, P.J.; van de Werken, H.; de Wit, E.; Oude Vrielink, J.A.; Elkon, R.; Melo, S.A.; Leveille, N.; Kalluri, R.; *et al.* ERNAs are required for *p53*-dependent enhancer activity and gene transcription. *Mol. Cell* **2013**, *49*, 524–535.
183. Brantl, S. Antisense-RNA regulation and RNA interference. *Biochim. Biophys. Acta* **2002**, *1575*, 15–25.
184. Prescott, E.M.; Proudfoot, N.J. Transcriptional collision between convergent genes in budding yeast. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8796–8801.
185. Thrash-Bingham, C.A.; Tartof, K.D. *Ahif*: A natural antisense transcript overexpressed in human renal cancer and during hypoxia. *J. Natl. Cancer Inst.* **1999**, *91*, 143–151.
186. Cayre, A.; Rossignol, F.; Clottes, E.; Penault-Llorca, F. *Ahif* but not *hif-1alpha* transcript is a poor prognostic marker in human breast cancer. *Breast Cancer Res.* **2003**, *5*, R223–230.
187. Zhang, H.; Gao, S.; De Geyter, C. A natural antisense transcript, bokas, regulates the pro-apoptotic activity of human bok. *Int. J. Oncol.* **2009**, *34*, 1135–1138.
188. Beltran, M.; Puig, I.; Pena, C.; Garcia, J.M.; Alvarez, A.B.; Pena, R.; Bonilla, F.; de Herreros, A.G. A natural antisense transcript regulates *zeb2/sip1* gene expression during snail1-induced epithelial-mesenchymal transition. *Genes Dev.* **2008**, *22*, 756–769.
189. Fabbri, M.; Paone, A.; Calore, F.; Galli, R.; Gaudio, E.; Santhanam, R.; Lovat, F.; Fadda, P.; Mao, C.; Nuovo, G.J.; *et al.* MicroRNAs bind to toll-like receptors to induce prometastatic inflammatory response. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E2110–2116.
190. Mitchell, P.S.; Parkin, R.K.; Kroh, E.M.; Fritz, B.R.; Wyman, S.K.; Pogosova-Agadjanyan, E.L.; Peterson, A.; Noteboom, J.; O'Briant, K.C.; Allen, A.; *et al.* Circulating microRNAs as stable blood-based markers for cancer detection. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10513–10518.
191. Rai, K.; Takigawa, N.; Ito, S.; Kashihara, H.; Ichihara, E.; Yasuda, T.; Shimizu, K.; Tanimoto, M.; Kiura, K. Liposomal delivery of microRNA-7-expressing plasmid overcomes epidermal growth factor receptor tyrosine kinase inhibitor-resistance in lung cancer cells. *Mol. Cancer Ther.* **2011**, *10*, 1720–1727.
192. Kota, J.; Chivukula, R.R.; O'Donnell, K.A.; Wentzel, E.A.; Montgomery, C.L.; Hwang, H.W.; Chang, T.C.; Vivekanandan, P.; Torbenson, M.; Clark, K.R.; *et al.* Therapeutic microRNA delivery suppresses tumorigenesis in a murine liver cancer model. *Cell* **2009**, *137*, 1005–1017.
193. Su, J.; Baigude, H.; McCarroll, J.; Rana, T.M. Silencing microRNA by interfering nanoparticles in mice. *Nucleic Acids Res.* **2011**, *39*, e38.

Reprinted from *IJMS*. Cite as: Hauptman, N.; Glavač, D. Long Non-Coding RNA in Cancer *Int. J. Mol. Sci.* **2013**, *14*, 4655-4669.

Review

Long Non-Coding RNA in Cancer

Nina Hauptman † and Damjan Glavač †,*

Department of Molecular Genetics, Institute of Pathology, University of Ljubljana, SI-1000 Ljubljana, Slovenia; E-Mail: nina.hauptman@mf.uni-lj.si

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: damjan.glavac@mf.uni-lj.si; Tel.: +386-1-543-7180; Fax: +386-1-543-7181.

Received: 1 November 2012; in revised form: 3 January 2013 / Accepted: 31 January 2013 / Published: 26 February 2013

Abstract: Long non-coding RNAs (lncRNAs) are pervasively transcribed in the genome and are emerging as new players in tumorigenesis due to their various functions in transcriptional, posttranscriptional and epigenetic mechanisms of gene regulation. LncRNAs are deregulated in a number of cancers, demonstrating both oncogenic and tumor suppressive roles, thus suggesting their aberrant expression may be a substantial contributor in cancer development. In this review, we will summarize their emerging role in human cancer and discuss their perspectives in diagnostics as potential biomarkers.

Keywords: long non-coding RNA; cancer; oncogenic lncRNA; tumor suppressor lncRNA

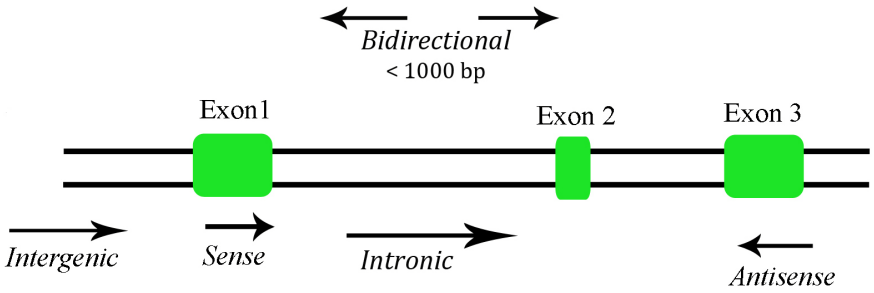
1. Introduction

The central dogma of molecular biology postulates gene-coding through storage of genetic information and proteins as the main molecules of cellular functions, while RNA has the role of an intermediary between DNA sequence and encoded protein. The findings of the human genome project thus came as a surprise, since only 1.5% of the human genome encodes protein-coding genes [1–5]. Development of new techniques revolutionized the molecular world with evidence that at least 90% of the human genome is actively transcribed [6,7]. The human transcriptome has shown more complexity than previously assumed since the protein-coding transcripts are being a minority, compared to a more complex group of non-coding RNAs (ncRNAs), such as microRNAs (miRNAs), long non-coding RNAs (lncRNAs), small nucleolar RNAs (snoRNAs), small interfering RNAS (siRNAs), small nuclear

(snRNAs), and piwi-interacting RNAs (piRNAs) [8–15]. Although initially thought to be transcriptional noise, ncRNA may play a crucial role in cellular development, physiology and pathologies.

Depending on their size, ncRNAs are divided into two major groups. Transcripts shorter than 200 nucleotides are referred to as small ncRNAs, which include miRNA, siRNA, piRNA, *etc.* The other group is composed of lncRNA, where the transcripts lack a significant open reading frame, and have length of 200 nt up to 100 kilobases. A lncRNA can be placed into one or more of five broad categories: (1) sense, or (2) antisense, when overlapping one or more exons of another transcript on the same, or opposite, strand, respectively; (3) bidirectional, when the sequence is located on the opposite strand from a neighboring coding transcript whose transcription is initiated less than 1000 base pairs away, (4) intronic, when it is derived wholly from within an intron of a second transcript, or (5) intergenic, when it lies within the genomic interval between two genes [16] (Figure 1). There are some lncRNAs that are transcribed by RNA polymerase III while the majority of lncRNAs are transcribed by RNA polymerase II, spliced and polyadenylated [17]. Most of the lncRNAs are located in the cytoplasm, although there are some found in both cytoplasm and nucleus [18].

Figure 1. Categories of long non-coding RNA (lncRNA).



2. Long Non-Coding RNA Functions

LncRNAs are involved in almost every step of a life cycle of genes and regulate diverse functions. Several lncRNAs can regulate gene expression at various levels, including chromatin modification, transcription, and posttranscriptional processing [19].

So far, their role was extensively studied in epigenetic regulation, such as imprinting. Diploid organisms carry two alleles of each of the parents' autosomal genes. In most cases, both of the alleles are expressed equally, except when a subset of genes shows imprinting in which expression is restricted by epigenetic mechanism to either maternal or paternal allele [17]. X-inactivation (XCI) is a process that equalizes gene expression between males and females by inactivating one X in female cells [17]. Some lncRNAs participate in global cellular behavior by controlling apoptosis, cell death and cell growth [15,20]. LncRNA can also mediate epigenetic modification by recruiting chromatin remodeling complex to specific chromatin loci, e.g., HOTAIR by polycomb repression complex 2 (PCR2) and/or lysine-specific demethylase 1 (LSD1), CCND1 by protein termed translocated in liposarcoma (TLS), and ANRIL by polycomb repression complex 1 and 2 (PCR1 and PCR2) [5,21–25]. The mode of action of some lncRNAs is interaction with their intracellular steroid receptors. Other lncRNAs function by regulating transcription through a variety of mechanisms that include interacting with RNA-binding proteins, acting as a coactivator of transcription factors, or repressing a major promoter of their target gene [22]. In addition to chromatin modification and transcriptional regulation,

lncRNAs can regulate gene expression at the posttranscriptional level.

3. Oncogenic lncRNA

SRA—Steroid Receptor RNA Activator is a coactivator for steroid receptors and acts as an ncRNA found in the nucleus and cytoplasm. SRA regulates gene expression mediated by steroid receptors through complexing with proteins also containing steroid receptor coactivator 1 (SRC-1) [26]. The SRA1 gene can also encode a protein that acts as a coactivator and corepressor [27]. SRA levels have been found to be upregulated in breast tumors where it is assumed that increased SRA levels change the steroid receptors' actions, contributing to breast tumorigenesis. While the expression of SRA in normal tissues is low, it is highly up-regulated in various tumors of the human breast, uterus and ovary. This evidence supports that SRA is a potential biomarker of steroid-dependent tumors [26].

HOTAIR—HOX Antisense Intergenic RNA with a length of 2.2 kb was found in the *HOXC* locus and is transcribed in antisense manner [28]. It is the first lncRNA discovered to be involved in tumorigenesis. In breast cancer, both primary and metastatic, the expression is up regulated; in the latter case up to 2000-fold increase was shown [23]. The high expression level of HOTAIR in primary breast cancer is also correlated to metastasis, and poor survival rate [23]. The level of HOTAIR expression is higher in patients with lymph node metastasis in hepatocellular cancer [29].

Polycomb group proteins mediate repression of transcription of thousands of genes that control differentiation pathways during development, and have roles in stem cell pluripotency and human cancer [23,30–34]. The target of PRC2 is the *HOXD* locus on chromosome 2 where the PRC2 in association with HOTAIR causes the transcriptional silencing of several metastasis suppressor genes resulting in breast epithelial cells having the expression of embryonic fibroblast. Alternating the level of HOTAIR results in enhanced PRC2 repressive activity [23]. HOTAIR acts as a molecular scaffold having two known chromatin modification complexes. The 5′ region of lncRNA binds to the PRC2 complex responsible for H3K27 methylation and the 3′ region binds to LSD1, which mediates enzymatic demethylation of H3K4 [24,30,35]. This result suggests the possible function of HOTAIR as a scaffold binding to selected histone modification enzymes and therefore causing histone modification on target genes [30]. Although the precise mechanism is still not known, it is clear that HOTAIR remodels chromatin to promote cancer invasiveness.

HOTAIR as an epigenetic regulator in gene expression is deregulated in different cancers [23,36–38]. In hepatocellular carcinoma (HCC) and HCC patients with liver transplantation, the levels of HOTAIR compared with normal liver tissue are elevated. Expression levels of HOTAIR can also be used as an independent prognostic marker for HCC recurrence and lower survival rate [31]. HOTAIR can be a potential biomarker for the existence of lymph node metastasis in HCC [29].

ANRIL—Antisense ncRNA in the INK4 locus

Many transcripts coding for proteins have anti-sense partners, whose perturbation can alter the expression of the sense transcripts [39]. Some of these genes are tumor suppressors, which can be epigenetically silenced by antisense ncRNA [40].

ANRIL activates two polycomb repressor complexes, PRC1 and PRC2 [21,25], resulting in chromatin reorganization which silences the INK4b-ARF-INK4a locus encoding tumor suppressors $p15^{INK4b}$, $p14^{ARF}$, $p16^{INK4a}$, which are active in cell cycle inhibition, senescence and stress-induced apoptosis. Overexpression of ANRIL in prostate cancer has shown silencing of INK4b-ARF-INK4a and p15/CDKN2B by heterochromatin reformation [25,41]. The repression is mediated by direct binding to combox 7 (CBX 7) and SUZ12, members of PRC1 and PRC2, respectively [21,25].

MALAT 1—Metastasis-Associated Lung Adenocarcinoma Transcript 1

This lncRNA was first associated with high metastatic potential and poor patient prognosis during a comparative screen of non-small cell lung cancer patients with or without metastatic tumors [42]. MALAT1 is widely expressed in normal human tissues [42,43] and is found to be upregulated in a variety of human cancers of the breast, prostate, colon, liver and uterus [44–47]. The MALAT1 locus at 11q13.1 has been identified to harbor chromosomal translocation break points associated with cancer [48–50]. MALAT1 is localized in nuclear speckles and widely expressed in normal tissues [42,43], but was found to be upregulated in hepatocellular carcinoma, breast, pancreas, osteosarcoma, colon and prostate cancers [44–47,51]. It has been shown that increased expression of MALAT1 can be used as a prognostic marker for HCC patients following liver transplantation [52].

A number of studies have implicated MALAT1 in the regulation of cell mobility, due to its high levels of expression in cancers. For example, RNA interference-mediated silencing of MALAT1 reduced the *in vitro* migration of lung adenocarcinoma cells by influencing the expression of motility-related genes [53]. Recent studies on knockout MALAT1 mice have not displayed any cellular phenotype. Future studies will be needed where mice will be exposed to different stresses, such as induction of cancer, which will potentially unveil its function. It is known that MALAT1 as well as HOTAIR play vital roles in human cells but it is possible that they have no significant role in living animals under normal physiological conditions [54–56].

4. Oncogenic and Tumor Suppressor lncRNA

H19 is expressed from the maternal allele and has a pivotal role in genomic imprinting during cell growth and development [57]. The locus contains *H19* and insulin-like growing factor 2 (*IGF2*), which are imprinted. This leads to differential expression of both genes, *H19* from maternal and *IGF2* from paternal allele [57,58]. The loss of imprinting results in misexpression of *H19* and was observed in many tumors including hepatocellular and bladder cancer [59,60]. This lncRNA has been linked to oncogenic and tumor suppressor properties [57]. cMYC induces the expression of *H19* in different cell types where *H19* potentiates tumorigenesis [58]. In addition c-MYC also down-regulates expression of *IGF2* imprinted gene. *H19* transcripts are precursors for miR-675 which functionally down-regulates the tumor suppressor gene for retinoblastoma in human colorectal cancer [61]. Data support *H19*

deregulation causing either oncogenic or tumor suppressor properties, although the exact mechanism is still elusive.

5. Tumor Suppressor lncRNA

MEG3—Maternally Expressed Gene 3

LncRNA MEG3 is a transcript of the maternally imprinted gene. In normal pituitary cells MEG3 is expressed, the loss of expression is observed in pituitary adenomas and the majority of meningiomas and meningioma cell lines [62,63]. MEG3 activates regulation of tumor suppressor protein p53. Normally, p53 protein levels are extremely low due to its rapid degradation via the ubiquitin-proteasome pathway. The ubiquitination of p53 is mainly mediated by MDM2, an E3 ubiquitin ligase. MEG3 down-regulates MDM2 expression, which suggests that MDM2 down-regulation is one of the mechanisms whereby MEG3 activates p53 [64]. MEG3 significantly increases p53 protein level and stimulates p53-dependent transcription [65]. MEG3 enhances p53 binding to target promoters such as GDF15 but not p21 and is also able to inhibit cell proliferation in the absence p53, suggesting that MEG3 is a p53 dependent and independent tumor suppressor [62–65].

GAS5—Growth Arrest-Specific 5 is widely expressed in embryonic and adult tissues. Expression is almost undetectable in growing leukemia cells and abundant in saturation density-arrested cells [66,67]. GAS5 functions as a starvation or growth arrest-linked riborepressor for the glucocorticoid receptors by binding to their DNA binding domain inhibiting the association of these receptors with their DNA recognition sequence. This suppresses the induction of several responsive genes including the gene encoding cellular inhibitor of apoptosis 2 (cIAP2), reducing cell metabolism and synthesizes cells to apoptosis [4,67]. GAS5 can induce apoptosis directly or indirectly in the prostate and breast cancer cell lines, where it was shown that GAS5 has a significantly lower expression in breast cancers compared to normal breast epithelial tissues [68].

CCND1/Cyclin D1 is a heterogenous lncRNA transcribed from the promoter region of the Cyclin D1 gene. Cyclin D1 is a cell cycle regulator that is frequently mutated, amplified and over expressed in a variety of cancers [69]. LncRNA recruits the RNA-binding proteinTLS, which is a key transcriptional regulatory sensor of DNA damage signals. Upon binding TLS undergoes allosteric modification, modulating activities of CREB-binding protein (CBP) and p300, resulting in inhibition of the cyclin D1 gene expression [22].

LincRNA-p21 expression is directly induced by the p53 signaling pathway. It is required for global repression of genes that interfere with p53 function to regulate cellular apoptosis. Lincrna-p21 mediated gene repression occurs through physical interaction with RNA-binding protein hnRNP K leading to the promoters of genes being repressed in a p53 dependent manner [70].

6. Diagnostic Benefits of lncRNA

So far, the majority of cancer biomarkers are protein-coding genes, their transcripts or the proteins. The non-coding regions are evolving as a biomarker hotspots only recently. By the advent of high-throughput sequencing, we are now able to identify deregulated expression of transcriptome at much higher resolution, what allow us to decipher smaller changes in the expression level. In the case

of lncRNAs, where their main function is regulation of other genes expression, the importance of lncRNAs maintained expression is evident. Since cancer is a complicated disease, which involves many factors, molecular biomarkers are valuable diagnostic and prognostic tools that could ease the disease management. Compared to protein-coding RNAs, using lncRNA as markers is of advantage since their own expression is a better indicator of the tumor status. Many lncRNAs are now connected to cancer due to new technologies and are emerging into the field of molecular biology as new regulatory players. Several lncRNA were found to be deregulated in a wide variety of cancers (Table 1).

Table 1. Cancer associated lncRNAs (adapted from [15,71]).

Name	Cytoband (Size)	Cancer Types	References
AK023948	8q24.22 (2807 nt)	Papillary thyroid carcinoma (down regulated)	[72]
ANRIL	9p21.3 (~3.9kb)	Prostate, leukemia	[36,41]
BC200	2p21 (200 nt)	Breast, cervix, esophagus, lung, ovary, parotid, tongue	[73,74]
PRNCR1	8q24.2 (13 kb)	Prostate	[75]
H19	11p15.5 (2.3 kb)	Bladder, lung, liver, breast, esophagus, choriocarcinoma, colon	[57,58,76–80]
HOTAIR	12q13.13 (2.2 kb)	Breast, hepatocellular	[23,29,30,36]
HULC	6p24.3 (~500 nt)	Hepatocellular	[4,81,82]
LincRNA-p21	~3.1 kb	Represses p53 pathway; induces apoptosis	[70]
Loc285194	3q13.31 (2105 nt)	Osteosarcoma	[83]
Malat1	11q13.1 (7.5 kb)	breast, prostate, colon, liver, uterus	[44–47]
MEG3	14q32.2 (1.6 kb)	Brain (down-regulated)	[62,65]
PTNEP1	9p13.3 (3.9 kb)	Prostate	[84]
Spry4-it1	5q31.3 (~700 nt)	Melanoma	[85]
SRA	5q31.3 (1965 nt)	Breast, uterus, ovary (down-regulated)	[26,27]
UCA1/CUDR	19p13.12 (1.4, 2.2, 2.7 kb)	Bladder, colon, cervix, lung, thyroid, liver, breast, esophagus, stomach	[86,87]
Wt1-as	11p13 (isoforms)	acute myeloid leukemia	[88]
PCA3	9q21.22 (0.6–4 kb)	Prostate	[89]
GAS5	1q25.1 (isoforms)	Breast (down-regulated)	[68]

In breast cancer research higher expressions of SRA and SRAP, compared to normal tissue were observed. Possibly SRAP expression contributes to higher survival for patient undergoing Tamoxifen treatment [90].

The expression of MALAT 1 is elevated in osteosarcoma patients with poor response to chemotherapy, which suggests that this transcript plays a crucial role in the pathology of tumors [53]. Additionally MALAT 1 serves as an independent prognostic marker for patient survival in early stage non-small cell lung cancer [42].

In hepatocellular carcinoma, (HCC) definitive diagnosis of lymph node metastasis is difficult without histological evidence. It has been demonstrated that a significant correlation between *HOTAIR* gene expression and lymph node metastasis exists, suggesting that measuring HOTAIR lncRNA is a potential biomarker for predicting lymph node metastasis [29]. Upregulation of HOTAIR is closely associated with gastrointestinal stromal tumor (GIST) aggressiveness and metastasis and it can be used as a potential biomarker [38,91].

MALAT1 is a powerful biomarker for HCC recurrence prediction following liver transplantation. Moreover, silencing MALAT1 activity in HCC would be a potential anticancer therapy to prevent tumor recurrence after orthotopic liver transplantation [52].

SPRY4-IT1 expression is substantially increased in patient melanoma cell samples compared to melanocytes. The elevated expression of SPRY4-IT1 in melanoma cells, its accumulation in the cell cytoplasm, and its effects on cell dynamics suggest that the misexpression of SPRY4-IT1 may have an important role in melanoma development, and could be an early biomarker and a key regulator for melanoma pathogenesis in humans [85].

The novel potential biomarkers can be discovered through certain types of highly expressed cancer-associated lncRNAs [92]. Therapeutic benefit can be obtained through pathways mediating transcriptional gene silencing, especially those of tumor suppressors and oncogenes [93]. For patients' comfort, biomarkers should be detected in samples obtained in a non-invasive way. Desirable samples are body fluids, such as serum or urine, where circulating nuclear acids (CNAs), both DNA and RNA species, are found. CNAs are found in plasma, cell-free serum, sputum and urine [29,94–97].

PRNCR1 (prostate cancer non-coding RNA1) expression was upregulated in some of the prostate cancer cells as well as precursor lesion prostatic intraepithelial neoplasia and considered as a tumor marker [75].

Suggestions that lncRNA can be used as biomarkers and/or drug targets have arisen from numerous studies observing the expression patterns of tumor tissues comparing to normal ones [14]. The possible therapies arising from this knowledge would be beneficial in cases where protein target drugs have not been effective. A recent study has shown that reduced expression of ncRAN enhanced the chemotherapeutic drug *in vitro* [98]. This opens another possibility of cancer treatment, where a combination of drugs would have much higher effect.

Often lncRNAs exhibit tissue specific patterns that distinguish them from miRNAs and protein-coding mRNAs that are expressed from multiple tissue types. Their specificity makes them precise biomarkers for cancer diagnostics [99]. PCA3 is a prostate-specific lncRNA overexpressed in prostate cancer. Although its functions are not understood, it was still utilized as a biomarker in a clinical test. Expression of the PCA3 transcript is determined from prostate cells in urine samples of patients [100,101]. Another lncRNA detected in body fluids is HULC, expression of which is disrupted in hepatocellular carcinomas and can be monitored in patients' blood sera [102].

To understand the biology of cancer it will be essential to identify, annotate lncRNAs and study their expression profiles in human tissues and diseases [103,104]. With this, the potential of lncRNAs on biology and medicine will be revealed. Long non-coding RNAs have recently arisen as new discoveries in the field of molecular biology. Since only a few individual lncRNAs have been functionally studied, still a lot of questions remain to be addressed [4]. At the moment the full potential of cancer therapy is not yet developed. The future of it lies in specific targeting of cancer cells and specific delivery of the drugs. LncRNAs are a possible resource for developing diagnostics and therapies, although a better understanding of their function and precise mechanism through which they function are needed first [4]. Another possibility for cancer treatment lies in combination of drugs, where one would change the expression of lncRNA in a way for chemotherapeutic drug to have a higher effect. Since probably the lncRNA function through their secondary structure special molecules could be developed to disrupt their secondary structure or bind to them to form complexes through

which an inactivation of lncRNA would occur. These molecules should be highly specific in order not to disrupt other molecules and mechanisms. To discover the right molecules more studies of the complex mechanisms involving lncRNA are needed.

7. Conclusions

RNA used to be just a messenger between coding genes and proteins encoded by them. However, "transcriptional noise" is turning out to be a very important part of regulation processes. With the discovery of LncRNA and their functions, the new world of molecular biology is emerging. There is much research still on the way towards a deeper understanding of regulation processes in which lncRNA is one of the important players. LncRNA deregulation in human disease is unveiling the complexity of cellular processes. Studying the mechanisms of lncRNA involvement in oncogenic and tumor suppressive pathways will lead to new cancer diagnostic markers and will pave the way to novel therapeutic targets.

Acknowledgments

This work was supported by program P3-0054 of the Slovenian Research Agency.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Stein, L.D. Human genome: End of the beginning. *Nature* **2004**, *431*, 915–916.
2. Ponting, C.P.; Belgard, T.G. Transcribed dark matter: meaning or myth? *Hum. Mol. Genet.* **2010**, *19*, R162–R168.
3. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; *et al.* Initial sequencing and analysis of the human genome. *Nature* **2001**, *49*, 860–921.
4. Gutschner, T.; Diederichs, S. The Hallmarks of Cancer: A long non-coding RNA point of view. *RNA Biol.* **2012**, *9*, 703–719.
5. Nie, L.; Wu, H.-J.; Hsu, J.-M.; Chang, S.-S.; LaBaff, A.; Li, C.-W.; Wang, Y.; Hsu, J.L.; Hung, M.-C. Long non-coding RNAs: Versatile master regulators of gene expression and crucial players in cancer. *Am. J. Transl. Res.* **2012**, *4*, 127–150.
6. Birney, E.; Stamatoyannopoulos, J.A.; Dutta, A.; Guigó, R.; Gingeras, T.R.; Margulies, E.H.; Weng, Z.; Snyder, M.; Dermitzakis, E.T.; Thurman, R.E.; *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Bioessays* **2007**, *32*, 599–608.
7. Costa, F.F. Non-coding RNAs: Meet thy masters. *Bioessays* **2010**, *32*, 599–608.
8. Kapranov, P.; Willingham, A.T.; Gingeras, T.R. Genome-wide transcription and the implications for genomic organization. *Nat. Rev. Genet.* **2007**, *8*, 413–423.
9. Frith, M.C.; Pheasant, M.; Mattick, J.S. The amazing complexity of the human transcriptome. *Eur. J. Hum. Genet.* **2005**, *13*, 894–897.

10. Khachane, A.N.; Harrison, P.M. Mining mammalian transcript data for functional long non-coding RNAs. *PLoS One* **2010**, *5*, doi:10.1371/journal.pone.0010316.
11. Mattick, J.S.; Makunin, I.V. Non-coding RNA. *Hum. Mol. Genet.* **2006**, *15*, R17–R29.
12. Guttman, M.; Amit, I.; Garber, M.; French, C.; Lin, M.F.; Feldser, D.; Huarte, M.; Zuk, O.; Carey, B.W.; Cassady, J.P.; *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **2009**, *458*, 223–227.
13. Washietl, S.; Hofacker, I.L.; Lukasser, M.; Huttenhofer, A.; Stadler, P.F. Mapping of conserved RNA secondary structures predicts thousands of functional noncoding RNAs in the human genome. *Nat. Biotechnol.* **2005**, *23*, 1383–1390.
14. Taft, R.J.; Pang, K.C.; Mercer, T.R.; Dinger, M.; Mattick, J.S. Non-coding RNAs: Regulators of disease. *J. Pathol.* **2010**, *220*, 126–139.
15. Sana, J.; Faltejskova, P.; Svoboda, M.; Slaby, O. Novel classes of non-coding RNAs and cancer. *J. Transl. Med.* **2012**, *10*, doi:10.1186/1479-5876-10-103.
16. Ponting, C.P.; Oliver, P.L.; Reik, W. Evolution and functions of long noncoding RNAs. *Cell* **2009**, *136*, 629–641.
17. Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914.
18. Banfai, B.; Jia, H.; Khatun, J.; Wood, E.; Risk, B.; Gundling, W.E.; Kundaje, A.; Gunawardena, H.P.; Yu, Y.; Xie, L.; *et al.* Long noncoding RNAs are rarely translated in two human cell lines. *Genome Res.* **2012**, *22*, 1646–1657.
19. Wilusz, J.E.; Sunwoo, H.; Spector, D.L. Long noncoding RNAs: Functional surprises from the RNA world. *Genes Dev.* **2009**, *23*, 1494–1504.
20. Wapinski, O.; Chang, H.Y. Long noncoding RNAs and human disease. *Trends. Cell Biol.* **2011**, *21*, 354–361.
21. Kotake, Y.; Nakagawa, T.; Kitagawa, K.; Suzuki, S.; Liu, N.; Kitagawa, M.; Xiong, Y. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. *Oncogene* **2011**, *30*, 1956–1962.
22. Wang, X.; Arai, S.; Song, X.; Reichart, D.; Du, K.; Pascual, G.; Tempst, P.; Rosenfeld, M.G.; Glass, C.K.; Kurokawa, R. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* **2008**, *454*, 126–130.
23. Gupta, R.A.; Shah, N.; Wang, K.C.; Kim, J.; Horlings, H.M.; Wong, D.J.; Tsai, M.C.; Hung, T.; Argani, P.; Rinn, J.L.; *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **2010**, *464*, 1071–1076.
24. Hayami, S.; Kelly, J.D.; Cho, H.S.; Yoshimatsu, M.; Unoki, M.; Tsunoda, T.; Field, H.I.; Neal, D.E.; Yamaue, H.; Ponder, B.A.; *et al.* Overexpression of LSD1 contributes to human carcinogenesis through chromatin regulation in various cancers. *Int. J. Cancer* **2011**, *128*, 574–586.
25. Yap, K.L.; Li, S.; Munoz-Cabello, A.M.; Raguz, S.; Zeng, L.; Mujtaba, S.; Gil, J.; Walsh, M.J.; Zhou, M.M. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* **2010**, *38*, 662–674.
26. Lanz, R.B.; Chua, S.S.; Barron, N.; Soder, B.M.; DeMayo, F.; O'Malley, B.W. Steroid receptor RNA activator stimulates proliferation as well as apoptosis *in vivo*. *Mol. Cell. Biol.* **2003**, *23*, 7163–7176.

27. Chooniedass-Kothari, S.; Vincett, D.; Yan, Y.; Cooper, C.; Hamedani, M.K.; Myal, Y.; Leygue, E. The protein encoded by the functional steroid receptor RNA activator is a new modulator of ER alpha transcriptional activity. *FEBS Lett.* **2010**, *584*, 1174–1180.
28. Rinn, J.L.; Kertesz, M.; Wang, J.K.; Squazzo, S.L.; Xu, X.; Brugmann, S.A.; Goodnough, L.H.; Helms, J.A.; Farnham, P.J.; Segal, E.; *et al.* Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell* **2007**, *129*, 1311–1323.
29. Geng, Y.J.; Xie, S.L.; Li, Q.; Ma, J.; Wang, G.Y. Large intervening non-coding RNA HOTAIR is associated with hepatocellular carcinoma progression. *J. Int. Med. Res.* **2011**, *39*, 2119–2128.
30. Tsai, M.C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J.K.; Lan, F.; Shi, Y.; Segal, E.; Chang, H.Y. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **2010**, *329*, 689–693.
31. Morey, L.; Helin, K. Polycomb group protein-mediated repression of transcription. *Trends. Biochem. Sci.* **2010**, *35*, 323–332.
32. Zhao, J.; Ohsumi, T.K.; Kung, J.T.; Ogawa, Y.; Grau, D.J.; Sarma, K.; Song, J.J.; Kingston, R.E.; Borowsky, M.; Lee, J.T. Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol. Cell* **2010**, *40*, 939–953.
33. Zhang, Z.; Jones, A.; Sun, C.W.; Li, C.; Chang, C.W.; Joo, H.Y.; Dai, Q.; Mysliwiec, M.R.; Wu, L.C.; Guo, Y.; *et al.* PRC2 complexes with JARID2, MTF2, and esPRC2p48 in ES cells to modulate ES cell pluripotency and somatic cell reprogramming. *Stem Cells* **2011**, *29*, 229–240.
34. Simon, J.A.; Lange, C.A. Roles of the EZH2 histone methyltransferase in cancer epigenetics. *Mutat. Res.* **2008**, *647*, 21–29.
35. Sirchia, S.M.; Tabano, S.; Monti, L.; Recalcati, M.P.; Gariboldi, M.; Grati, F.R.; Porta, G.; Finelli, P.; Radice, P.; Miozzo, M. Misbehaviour of XIST RNA in breast cancer cells. *PLoS One* **2009**, *4*, doi:10.1371/journal.pone.0005559.
36. Yang, Z.; Zhou, L.; Wu, L.M.; Lai, M.C.; Xie, H.Y.; Zhang, F.; Zheng, S.S. Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. *Ann. Surg. Oncol.* **2011**, *18*, 1243–1250.
37. Kogo, R.; Shimamura, T.; Mimori, K.; Kawahara, K.; Imoto, S.; Sudo, T.; Tanaka, F.; Shibata, K.; Suzuki, A.; Komune, S.; *et al.* Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res.* **2011**, *71*, 6320–6326.
38. Niinuma, T.; Suzuki, H.; Nojima, M.; Nosho, K.; Yamamoto, H.; Takamaru, H.; Yamamoto, E.; Maruyama, R.; Nobuoka, T.; Miyazaki, Y.; *et al.* Upregulation of miR-196a and HOTAIR drive malignant character in gastrointestinal stromal tumors. *Cancer Res.* **2012**, *72*, 1126–1136.
39. Katayama, S.; Tomaru, Y.; Kasukawa, T.; Waki, K.; Nakanishi, M.; Nakamura, M.; Nishida, H.; Yap, C.C.; Suzuki, M.; Kawai, J.; *et al.* Antisense transcription in the mammalian transcriptome. *Science* **2005**, *309*, 1564–1566.
40. Kim, W.Y.; Sharpless, N.E. The regulation of INK4/ARF in cancer and aging. *Cell* **2006**, *127*, 265–275.
41. Yu, W.; Gius, D.; Onyango, P.; Muldoon-Jacobs, K.; Karp, J.; Feinberg, A.P.; Cui, H. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature* **2008**, *451*, 202–206.

42. Ji, P.; Diederichs, S.; Wang, W.; Boing, S.; Metzger, R.; Schneider, P.M.; Tidow, N.; Brandt, B.; Buerger, H.; Bulk, E.; *et al.* MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* **2003**, *22*, 8031–8041.
43. Hutchinson, J.N.; Ensminger, A.W.; Clemson, C.M.; Lynch, C.R.; Lawrence, J.B.; Chess, A. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **2007**, *8*, doi:10.1186/1471-2164-8-39.
44. Guffanti, A.; Iacono, M.; Pelucchi, P.; Kim, N.; Solda, G.; Croft, L.J.; Taft, R.J.; Rizzi, E.; Askarian-Amiri, M.; Bonnal, R.J.; *et al.* A transcriptional sketch of a primary human breast cancer by 454 deep sequencing. *BMC Genomics* **2009**, *10*, doi:10.1186/1471-2164-10-163.
45. Yamada, K.; Kano, J.; Tsunoda, H.; Yoshikawa, H.; Okubo, C.; Ishiyama, T.; Noguchi, M. Phenotypic characterization of endometrial stromal sarcoma of the uterus. *Cancer Sci.* **2006**, *97*, 106–112.
46. Lin, R.; Maeda, S.; Liu, C.; Karin, M.; Edgington, T.S. A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. *Oncogene* **2007**, *26*, 851–858.
47. Luo, J.H.; Ren, B.; Keryanov, S.; Tsang, G.C.; Reo, U.N.M.; Monga, S.P.; Storm, A.; Demetris, A.J.; Nalesnik, M.; Yu, Y.P.; *et al.* Transcriptomic and genomic analysis of human hepatocellular carcinomas and hepatoblastomas. *Hepatology* **2006**, *44*, 1012–1024.
48. Davis, I.J.; Hsi, B.L.; Arroyo, J.D.; Vargas, S.O.; Yeh, Y.A.; Motyckova, G.; Valencia, P.; Perez-Atayde, A.R.; Argani, P.; Ladanyi, M.; *et al.* Cloning of an Alpha-TFEB fusion in renal tumors harboring the t(6;11)(p21;q13) chromosome translocation. *Proc. Natl. Acad Sci USA* **2003**, *100*, 6051–6056.
49. Kuiper, R.P.; Schepens, M.; Thijssen, J.; van Asseldonk, M.; van den Berg, E.; Bridge, J.; Schuuring, E.; Schoenmakers, E.F.; van Kessel, A.G. Upregulation of the transcription factor TFEB in t(6;11)(p21;q13)-positive renal cell carcinomas due to promoter substitution. *Hum. Mol. Genet.* **2003**, *12*, 1661–1669.
50. Rajaram, V.; Knezevich, S.; Bove, K.E.; Perry, A.; Pfeifer, J.D. DNA sequence of the translocation breakpoints in undifferentiated embryonal sarcoma arising in mesenchymal hamartoma of the liver harboring the t(11;19)(q11;q13.4) translocation. *Genes Chromosomes Cancer* **2007**, *46*, 508–513.
51. Fellenberg, J.; Bernd, L.; Delling, G.; Witte, D.; Zahlten-Hinguranage, A. Prognostic significance of drug-regulated genes in high-grade osteosarcoma. *Mod. Pathol.* **2007**, *20*, 1085–1094.
52. Lai, M.C.; Yang, Z.; Zhou, L.; Zhu, Q.Q.; Xie, H.Y.; Zhang, F.; Wu, L.M.; Chen, L.M.; Zheng, S.S. Long non-coding RNA MALAT-1 overexpression predicts tumor recurrence of hepatocellular carcinoma after liver transplantation. *Med. Oncol.* **2012**, *29*, 1810–1816.
53. Tano, K.; Mizuno, R.; Okada, T.; Rakwal, R.; Shibato, J.; Masuo, Y.; Ijiri, K.; Akimitsu, N. MALAT-1 enhances cell motility of lung adenocarcinoma cells by influencing the expression of motility-related genes. *FEBS Lett.* **2010**, *584*, 4575–4580.
54. Nakagawa, S.; Ip, J.Y.; Shioi, G.; Tripathi, V.; Zong, X.; Hirose, T.; Prasanth, K.V. Malat1 is not an essential component of nuclear speckles in mice. *RNA* **2012**, *18*, 1487–1499.

55. Bickmore, W.A.; Schorderet, P.; Duboule, D. Structural and Functional Differences in the Long Non-Coding RNA Hotair in Mouse and Human. *PLoS Genet.* **2011**, *7*, doi:10.1371/journal.pgen.1002071.
56. Eiβmann, M.; Gutschner, T.; Hämmerle, M.; Günther, S.; Caudron -Herger, M.; Groß, M.; Schirmacher, P.; Rippe, K.; Braun, T.; Zörnig, M.; Diederichs, S. Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. *RNA Biol.* **2012**, *9*, 1076–1087.
57. Gabory, A.; Jammes, H.; Dandolo, L. The *H19* locus: role of an imprinted non-coding RNA in growth and development. *Bioessays* **2010**, *32*, 473–480.
58. Barsyte-Lovejoy, D.; Lau, S.K.; Boutros, P.C.; Khosravi, F.; Jurisica, I.; Andrulis, I.L.; Tsao, M.S.; Penn, L.Z. The c-Myc oncogene directly induces the *H19* noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Res.* **2006**, *66*, 5330–5337.
59. van Bakel, H.; Nislow, C.; Blencowe, B.J.; Hughes, T.R. Most "dark matter" transcripts are associated with known genes. *PLoS Biol.* **2010**, *8*, doi:10.1371/journal.pbio.1000371.
60. Oosumi, T.; Belknap, W.R.; Garlick, B. Mariner transposons in humans. *Nature* **1995**, *378*, 672–672.
61. Tsang, W.P.; Ng, E.K.; Ng, S.S.; Jin, H.; Yu, J.; Sung, J.J.; Kwok, T.T. Oncofetal *H19*-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis* **2010**, *31*, 350–358.
62. Gejman, R.; Batista, D.L.; Zhong, Y.; Zhou, Y.; Zhang, X.; Swearingen, B.; Stratakis, C.A.; Hedley-Whyte, E.T.; Klibanski, A. Selective loss of MEG3 expression and intergenic differentially methylated region hypermethylation in the MEG3/DLK1 locus in human clinically nonfunctioning pituitary adenomas. *J. Clin. Endocrinol. Metab.* **2008**, *93*, 4119–4125.
63. Zhang, X.; Gejman, R.; Mahta, A.; Zhong, Y.; Rice, K.A.; Zhou, Y.; Cheunsuchon, P.; Louis, D.N.; Klibanski, A. Maternally expressed gene 3, an imprinted noncoding RNA gene, is associated with meningioma pathogenesis and progression. *Cancer Res*. **2010**, *70*, 2350–2358.
64. Zhou, Y.; Zhang, X.; Klibanski, A. MEG3 noncoding RNA: A tumor suppressor. *J. Mol. Endocrinol.* **2012**, *48*, R45–R53.
65. Benetatos, L.; Vartholomatos, G.; Hatzimichael, E. MEG3 imprinted gene contribution in tumorigenesis. *Int. J. Cancer* **2011**, *129*, 773–779.
66. Coccia, E.M.; Cicala, C.; Charlesworth, A.; Ciccarelli, C.; Rossi, G.B.; Philipson, L.; Sorrentino, V. Regulation and expression of a growth arrest-specific gene (gas5) during growth,differentiation, and development. *Mol. Cell. Biol.* **1992**, *12*, 3514–3521.
67. Kino, T.; Hurt, D.E.; Ichijo, T.; Nader, N.; Chrousos, G.P. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci Signal* **2010**, *3*, ra8.
68. Mourtada-Maarabouni, M.; Pickard, M.R.; Hedge, V.L.; Farzaneh, F.; Williams, G.T. GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene* **2009**, *28*, 195–208.
69. Diehl, J.A. Cycling to Cancer with Cyclin D1. *Cancer Biol. Ther.* **2002**, *1*, 226–231.
70. Huarte, M.; Guttman, M.; Feldser, D.; Garber, M.; Koziol, M.J.; Kenzelmann-Broz, D.; Khalil, A.M.; Zuk, O.; Amit, I.; Rabani, M.; *et al.* A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **2010**, *142*, 409–419.

71. Amaral, P.P.; Clark, M.B.; Gascoigne, D.K.; Dinger, M.E.; Mattick, J.S. lncRNAdb: A reference database for long noncoding RNAs. *Nucleic Acids Res.* **2011**, *39*, D146–D151.
72. He, H.; Nagy, R.; Liyanarachchi, S.; Jiao, H.; Li, W.; Suster, S.; Kere, J.; de la Chapelle, A. A susceptibility locus for papillary thyroid carcinoma on chromosome 8q24. *Cancer Res.* **2009**, *69*, 625–631.
73. Chen, W.; Böcker, W.; Brosius, J.; Tiedge, H. Expression of neural BC200 RNA in human tumours. *J. Pathol.* **1997**, *183*, 345–351.
74. Iacoangeli, A.; Lin, Y.; Morley, E.J.; Muslimov, I.A.; Bianchi, R.; Reilly, J.; Weedon, J.; Diallo, R.; Bocker, W.; Tiedge, H. BC200 RNA in invasive and preinvasive breast cancer. *Carcinogenesis* **2004**, *25*, 2125–2133.
75. Chung, S.; Nakagawa, H.; Uemura, M.; Piao, L.; Ashikawa, K.; Hosono, N.; Takata, R.; Akamatsu, S.; Kawaguchi, T.; Morizono, T.; *et al.* Association of a novel long non-coding RNA in 8q24 with prostate cancer susceptibility. *Cancer Sci.* **2011**, *102*, 245–252.
76. Hibi, K.; Nakamura, H.; Hirai, A.; Fujikake, Y.; Kasai, Y.; Akiyama, S.; Ito, K.; Takagi, H. Loss of *H19* imprinting in esophageal cancer. *Cancer Res.* **1996**, *56*, 480–482.
77. Fellig, Y.; Ariel, I.; Ohana, P.; Schachter, P.; Sinelnikov, I.; Birman, T.; Ayesh, S.; Schneider, T.; de Groot, N.; Czerniak, A.; *et al. H19* expression in hepatic metastases from a range of human carcinomas. *J. Clin. Pathol.* **2005**, *58*, 1064–1068.
78. Matouk, I.J.; de Groot, N.; Mezan, S.; Ayesh, S.; Abu-lail, R.; Hochberg, A.; Galun, E. The *H19* non-coding RNA is essential for human tumor growth. *PLoS One* **2007**, *2*, doi:10.1371/journal.pone.0000845.
79. Arima, T.; Matsuda, T.; Takagi, N.; Wake, N. Association of IGF2 and *H19* imprinting with choriocarcinoma development. *Cancer Genet. Cytogenet.* **1997**, *93*, 39–47.
80. Berteaux, N.; Lottin, S.; Monte, D.; Pinte, S.; Quatannens, B.; Coll, J.; Hondermarck, H.; Curgy, J.J.; Dugimont, T.; Adriaenssens, E. *H19* mRNA-like noncoding RNA promotes breast cancer cell proliferation through positive control by E2F1. *J. Biol. Chem.* **2005**, *280*, 29625–29636.
81. Matouk, I.J.; Abbasi, I.; Hochberg, A.; Galun, E.; Dweik, H.; Akkawi, M. Highly upregulated in liver cancer noncoding RNA is overexpressed in hepatic colorectal metastasis. *Eur. J. Gastroenterol. Hepatol.* **2009**, *21*, 688–692.
82. Panzitt, K.; Tschernatsch, M.M.; Guelly, C.; Moustafa, T.; Stradner, M.; Strohmaier, H.M.; Buck, C.R.; Denk, H.; Schroeder, R.; Trauner, M.; *et al.* Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. *Gastroenterology* **2007**, *132*, 330–342.
83. Pasic, I.; Shlien, A.; Durbin, A.D.; Stavropoulos, D.J.; Baskin, B.; Ray, P.N.; Novokmet, A.; Malkin, D. Recurrent focal copy-number changes and loss of heterozygosity implicate two noncoding RNAs and one tumor suppressor gene at chromosome 3q13.31 in osteosarcoma. *Cancer Res*. **2010**, *70*, 160–171.
84. Poliseno, L.; Salmena, L.; Zhang, J.; Carver, B.; Haveman, W.J.; Pandolfi, P.P. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **2010**, *465*, 1033–1038.

85. Khaitan, D.; Dinger, M.E.; Mazar, J.; Crawford, J.; Smith, M.A.; Mattick, J.S.; Perera, R.J. The melanoma-upregulated long noncoding RNA SPRY4-IT1 modulates apoptosis and invasion. *Cancer Res.* **2011**, *71*, 3852–3862.
86. Wang, F.; Li, X.; Xie, X.; Zhao, L.; Chen, W. UCA1, a non-protein-coding RNA up-regulated in bladder carcinoma and embryo, influencing cell growth and promoting invasion. *FEBS Lett.* **2008**, *582*, 1919–1927.
87. Wang, X.S.; Zhang, Z.; Wang, H.C.; Cai, J.L.; Xu, Q.W.; Li, M.Q.; Chen, Y.C.; Qian, X.P.; Lu, T.J.; Yu, L.Z.; *et al.* Rapid identification of UCA1 as a very sensitive and specific unique marker for human bladder carcinoma. *Clin. Cancer Res.* **2006**, *12*, 4851–4858.
88. Dallosso, A.R.; Hancock, A.L.; Malik, S.; Salpekar, A.; King-Underwood, L.; Pritchard-Jones, K.; Peters, J.; Moorwood, K.; Ward, A.; Malik, K.T.; *et al.* Alternately spliced WT1 antisense transcripts interact with WT1 sense RNA and show epigenetic and splicing defects in cancer. *RNA* **2007**, *13*, 2287–2299.
89. de Kok, J.B.; Verhaegh, G.W.; Roelofs, R.W.; Hessels, D.; Kiemeney, L.A.; Aalders, T.W.; Swinkels, D.W.; Schalken, J.A. DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. *Cancer Res.* **2002**, *62*, 2695–2689.
90. Leygue, E. Steroid receptor RNA activator (SRA1): Unusual bifaceted gene products with suspected relevance to breast cancer. *Nucl. Recept. Signal.* **2007**, *4*, doi:10.1621/nrs.05006.
91. Qi, P.; Du, X. The long non-coding RNAs, a new cancer diagnostic and therapeutic gold mine. *Mod. Pathol.* **2012**, doi:10.1038/modpathol.2012.160.
92. Gibb, E.A.; Brown, C.J.; Lam, W.L. The functional role of long non-coding RNA in human carcinomas. *Mol. Cancer* **2011**, *10*, doi:10.1186/1476-4598-10-38.
93. Morris, K.V. RNA-directed transcriptional gene silencing and activation in human cells. *Oligonucleotides* **2009**, *19*, 299–306.
94. Schöler, N.; Langer, C.; Döhner, H.; Buske, C.; Kuchenbauer, F. Serum microRNAs as a novel class of biomarkers: A comprehensive review of the literature. *Exp. Hematol.* **2010**, *38*, 1126–1130.
95. Xie, Y.; Todd, N.W.; Liu, Z.; Zhan, M.; Fang, H.; Peng, H.; Alattar, M.; Deepak, J.; Stass, S.A.; Jiang, F. Altered miRNA expression in sputum for diagnosis of non-small cell lung cancer. *Lung Cancer* **2010**, *67*, 170–176.
96. Xing, L.; Todd, N.W.; Yu, L.; Fang, H.; Jiang, F. Early detection of squamous cell lung cancer in sputum by a panel of microRNA markers. *Mod. Pathol.* **2010**, *23*, 1157–1164.
97. Kosaka, N.; Iguchi, H.; Ochiya, T. Circulating microRNA in body fluid: a new potential biomarker for cancer diagnosis and prognosis. *Cancer Sci.* **2010**, *101*, 2087–2092.
98. Zhu, Y.; Yu, M.; Li, Z.; Kong, C.; Bi, J.; Li, J.; Gao, Z. *ncRAN*, a newly identified long noncoding RNA, enhances human bladder tumor growth, invasion, and survival. *Urology* **2011**, *77*, 510.e1–510.e5.
99. Prensner, J.R.; Iyer, M.K.; Balbin, O.A.; Dhanasekaran, S.M.; Cao, Q.; Brenner, J.C.; Laxman, B.; Asangani, I.A.; Grasso, C.S.; Kominsky, H.D.; *et al.* Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. *Nat. Biotechnol.* **2011**, *29*, 742–749.

100. Hessels, D.; Klein Gunnewiek, J.M.T.; van Oort, I.; Karthaus, H.F.M.; van Leenders, G.J.L.; van Balken, B.; Kiemeney, L.A.; Witjes, J.A.; Schalken, J.A. DD3PCA3-based Molecular Urine Analysis for the Diagnosis of Prostate Cancer. *Eur. Urol.* **2003**, *44*, 8–16.
101. Tinzl, M.; Marberger, M.; Horvath, S.; Chypre, C. DD3PCA3 RNA Analysis in Urine – A New Perspective for Detecting Prostate Cancer. *Eur. Urol.* **2004**, *46*, 182–187.
102. Muro, E.M.; Andrade-Navarro, M.A. Pseudogenes as an alternative source of natural antisense transcripts. *BMC Evol. Biol.* **2010**, *10*, doi:10.1186/1471-2148-10-338.
103. Morris, K.V.; Santoso, S.; Turner, A.M.; Pastori, C.; Hawkins, P.G. Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. *PLoS Genet.* **2008**, *4*, doi:10.1371/journal.pgen.1000258.
104. Lyle, R.; Watanabe, D.; te Vruchte, D.; Lerchner, W.; Smrzka, O.W.; Wutz, A.; Schageman, J.; Hahner, L.; Davies, C.; Barlow, D.P. The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. *Nat. Genet.* **2000**, *25*, 19–21.

Reprinted from *IJMS*. Cite as: Li, X.; Zhang, Z.; Yu, M.; Li, L.; Du, G.; Xiao, W.; Yang, H. Involvement of miR-20a in Promoting Gastric Cancer Progression by Targeting Early Growth Response 2 (EGR2). *Int. J. Mol. Sci.* **2013**, *14*, 16226-16239.

Article

Involvement of miR-20a in Promoting Gastric Cancer Progression by Targeting Early Growth Response 2 (EGR2)

Xiangsheng Li, Zhichao Zhang, Ming Yu, Liqi Li, Guangsheng Du, Weidong Xiao * and Hua Yang *

Department of General Surgery, Xinqiao Hospital, Third Military Medical University, Xinqiao Road, Chongqing 400037, China; E-Mails: xiangsonlee@163.com (X.L.); zhichao_zhang@163.com (Z.Z.); yumimianbao@163.com (M.Y.); liliqi198610@163.com (L.L.); guangsheng_du@126.com (G.D.)

* Authors to whom correspondence should be addressed; E-Mails: weidong.xiao@126.com (W.X.); hwbyang@163.com (H.Y.); Tel./Fax: +86-023-6875-5114 (W.X. & H.Y.).

Received: 5 June 2013; in revised form: 18 July 2013 / Accepted: 18 July 2013 / Published: 6 August 2013

Abstract: Gastric cancer (GC) is one of the most common cancers, with high incidences in East Asia. microRNAs (miRNAs) play essential roles in the carcinogenesis of GC. miR-20a was elevated in GC, while the potential function of miR-20a was poorly understood. miR-20a expression was examined in GC tissues and cell lines. The effects of miR-20a on the growth, migration, invasion, and chemoresistance of GC cells were examined. Luciferase reporter assay and Western blot were used to screen the target of miR-20a. miR-20a was increased in GC tissues and cell lines. miR-20a promoted the growth, migration and invasion of GC cells, enhanced the chemoresistance of GC cells to cisplatin and docetaxel. Luciferase activity and Western blot confirmed that miR-20a negatively regulated EGR2 expression. Overexpression of EGR2 significantly attenuated the oncogenic effect of miR-20a. miR-20a was involved in the carcinogenesis of GC through modulation of the EGR2 signaling pathway.

Keywords: gastric cancer; miR-20a; early growth response 2; carcinogenesis; chemoresistance

1. Introduction

Gastric cancer (GC) is one of the most common cancers and the second most common malignancy of cancer death worldwide, especially in East Asia [1]. Despite a steady decrease in GC incidence and

mortality during the last decade, GC still ranked second in global cancer mortality. The carcinogenesis of GC was complicated involving dysregualtion of oncogenes and tumor suppressors [2]. Recently, emerging evidence found that a new group of RNAs, known as microRNAs (miRNAs), regulated a large variety of genes, including both oncogenes and tumor suppressor genes [3].

miRNAs were a family of endogenous, non-coding small RNAs (approximately 20–25 nucleotides), which negatively regulated gene expression by inhibiting translation or inducing mRNA degradation via binding to the 3' untranslated region (3' UTR) of target mRNAs. miRNAs played critical roles in the development, proliferation, differentiation, metabolism, and apoptosis. In addition, aberrant expression of miRNAs was related to carcinogenesis [4]. Abnormal expression profiles of miRNAs had been reported in numerous cancers, including breast, colon, lung, prostate, and GC [5–8]. miRNAs functioned as oncogenes or tumor suppressors by regulating different target gene expression in different cancers. miR-9, miR-22, and miR-146a had all been shown to act as tumor suppressors [9–11], whereas miR-19, miR-23a, and miR-301a had been shown to function as oncogenes [12–14]. These studies suggested that dysregulation of miRNAs might be involved in carcinogenesis and cancer progression.

miR-20a belonged to the miR-17-92 cluster, which were widely overexpressed oncogenes in diverse cancer subtypes [14]. Previous studies had shown that certain cancer suppressors, such as BH3-only protein (BIM), Phosphatase and tensin homolog (PTEN), were direct targets of the miR-17-92 cluster [12,15]. In human cervical cancer cells, miR-20a was reported to promote cancer cell migration and invasion by targeting Tankyrase 2 (TNKS2) [16]. Similarly, in osteosarcoma, miR-20a increased the metastatic potential of osteosarcoma cells by targeting Fas, indicating an oncogenic function of miR-20a [17]. However, the role of miR-20a in the progression of GC and its underlying mechanism remained unclear.

In this study, we explored the biological effects and the potential mechanisms of miR-20a in GC by detecting the expression of miR-20a in GC, validating previous finding that miR-20a was elevated in GC. Overexpression of miR-20a promoted the growth, migration, invasion, as well as chemoresistance of GC cells. By bioinformatics analysis, we predicted the tumor-suppressor, early growth response 2 (EGR2), as a putative target of miR-20a. Subsequent luciferase activity assay and Western blot confirmed that miR-20a repressed the expression of EGR2 by inducing EGR2 mRNA decay. Overexpression of EGR2 significantly attenuated the oncogenic effect of miR-20a.

2. Results

2.1. miR-20a Was Increased in GC Tissue Samples and Cell Lines

To validate the expression levels of miR-20a, we conducted quantitative real time PCR (qRT-PCR) in 28 GC tissues and the corresponding normal tissues and 3 GC cell lines and a gastric epithelial cell line. Expression of miR-20a was significantly increased in GC tissues and cell lines (Figure 1A,B). The clinicopathological features of 28 patients with GC were shown in Table 1, and miR-20a was correlated with the metastasis of GC patients, while the miR-20a expression had no relationship with other clinicopathological parameters. Thus, it was suggested that elevation of miR-20a in GC might have essential roles in GC carcinogenesis.

Figure 1. miR-20a was increased in gastric cancer (GC) tissue samples and cell lines. (**A**) qRT-PCR for miR-20a was performed using 28 GC tissue samples and matched with adjacent non-tumour normal tissues; (**B**) qRT-PCR for miR-20a was performed using 3 GC cell lines and human gastric mucosa cell line GES-1. The data represented triplicate measurements. * $p < 0.05$, ** $p < 0.01$ compared with control.

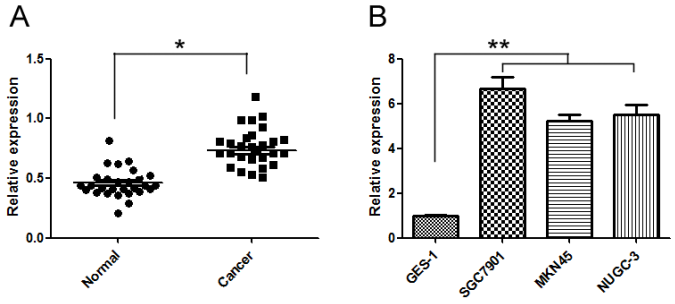


Table 1. Correlation of the expression of miR-20a with clinicopathologic features.

Clinicopathologic features	No.	Relative expression of miR-20a [a]	*p*-value [b]
Gender			0.727
Male	18	0.72 (0.51–1.33)	
Female	10	0.70 (0.52–1.02)	
Site of tumor			0.810
Upper stomach	8	0.68 (0.52–1.00)	
Middle stomach	6	0.70 (0.51–1.22)	
Lower stomach	14	0.71 (0.51–1.32)	
Differentiation			0.655
Poor	12	0.73 (0.59–1.33)	
Moderate	16	0.72 (0.52–1.12)	
Metastasis			0.003
N0	4	0.55 (0.51–0.61)	
N1	5	0.62 (0.51–0.64)	
N2	9	0.74 (0.52–1.09)	
N3	9	0.92 (0.51–1.33)	

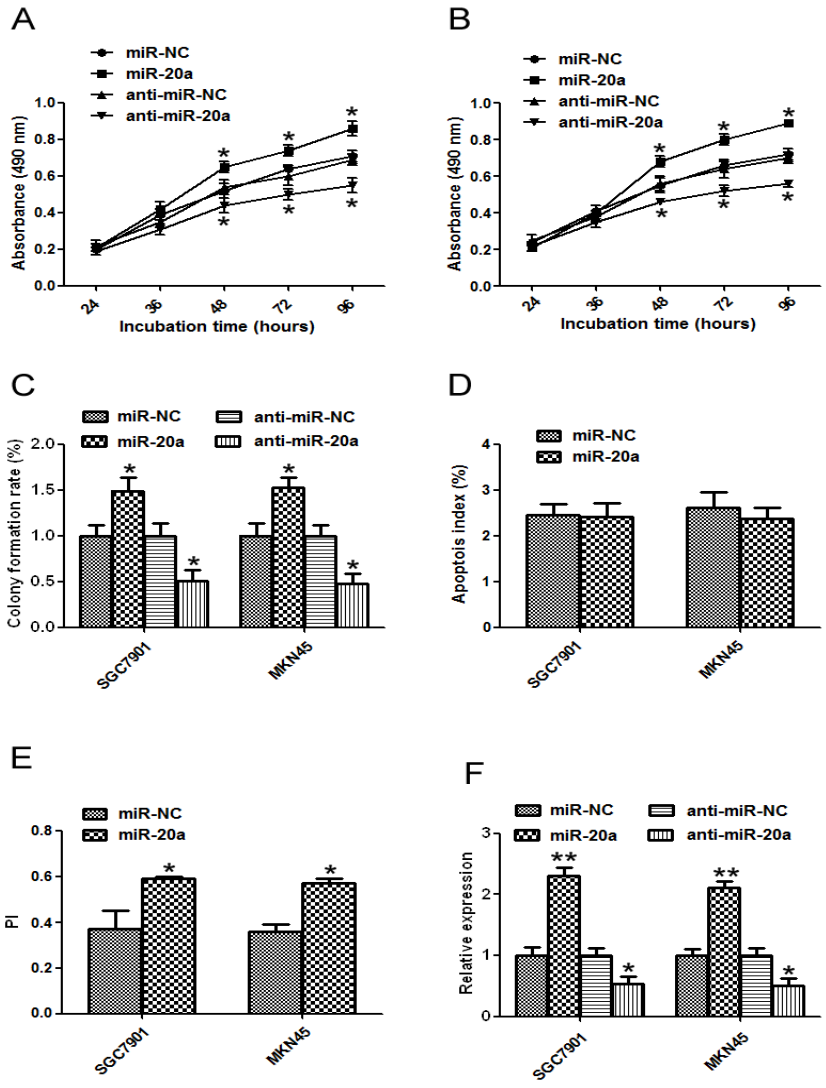
[a] Median of relative expression; [b] Mann-Whitney U test between two groups and Kruskall-Wallis test for three groups.

2.2. miR-20a Promoted Growth of GC Cells

As miR-20a was markedly increased in GC, it might function as a cancer promoter or suppressor. Therefore, we tested the role of miR-20a by gain- and loss- function experiments in SGC7901 and MKN45 cells. In a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide (MTT) assay, cells transfected with miR-20a precursor grew more rapidly than the control group, while miR-20a inhibitor inhibited the growth (Figure 2A,B). To further study the effect of miR-20a on the growth of GC cells, colony formation assay was performed. GC cells transfected with miR-20a precursor showed higher colony formation than cells transfected with control. However, cells transfected with miR-20a inhibitor showed lower colony formation than cells transfected with control (Figure 2C). We further used Flow cytometric analysis to determine the effect of miR-20a on apoptosis of GC cells; no significant

difference was detected between miR-20a precursor and the control in these cells (Figure 2D). Results from cell cycle assay indicated that overexpression of miR-20a precursor had less cells in G0/G1 phase. Furthermore, overexpression group had more cells in S and G2M phases (Figure 2E). Proliferation index (PI) = (S + G2M)/(G0 + GS + G2M). The PI was higher in the overexpression group than that in the control group. These data suggested that miR-20a might promote SGC7901 and MKN45 cell proliferation *in vitro*. The transfection efficiency was detected by qRT-PCR (Figure 2F).

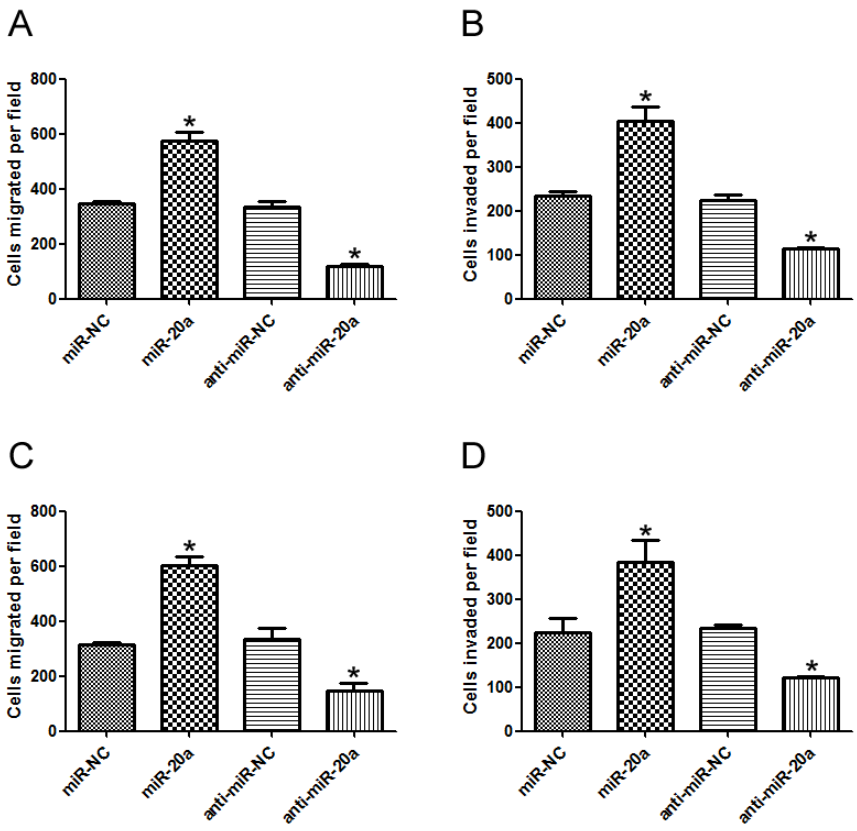
Figure 2. MiR-20a promoted growth of GC cell lines. SGC7901 or MKN45 cells were transfected with miR-20a precursor/inhibitor or the corresponding control, respectively. (**A**) At 24, 48, 36, 72, or 96 h after transfection, MTT assay was performed to examine SGC7901 proliferation; (**B**) MTT assay of MKN45 cells; (**C**) Representative results of colony formation assay in SGC7901 and MKN45 cells transfected with miR-20a precursor/inhibitor or the corresponding control, respectively; (**D**) The apoptosis of SGC7901 and MKN45 cells after miR-20a precursor transfection; (**E**) The proliferation index of SGC7901 and MKN45 cells transfected with miR-20a precursor; (**F**) qRT-PCR for miR-20a was performed in SGC7901 and MKN45 cells transfected with miR-20a precursor/inhibitor or the corresponding control, respectively. The data represented at least four measurements. * $p < 0.05$, ** $p < 0.01$ compared with control.



2.3. miR-20a Promoted Migration and Invasion of GC Cells

To investigate the role of miR-20a in GC metastasis, miR-20a precursor/inhibitor was transfected into SGC-7901 and MKN45 cells and *in vitro* migration and invasion assays were performed. Results showed that miR-20a significantly increased the *in vitro* migration ability of GC cells, while miR-20a inhibitor remarkably decreased the *in vitro* migration ability of GC cells (Figure 3A,C). Similar results were observed in *in vitro* invasion ability of GC cells (Figure 3B,D). Collectively, these data suggested that miR-20a promoted the migration and invasion abilities of GC cells.

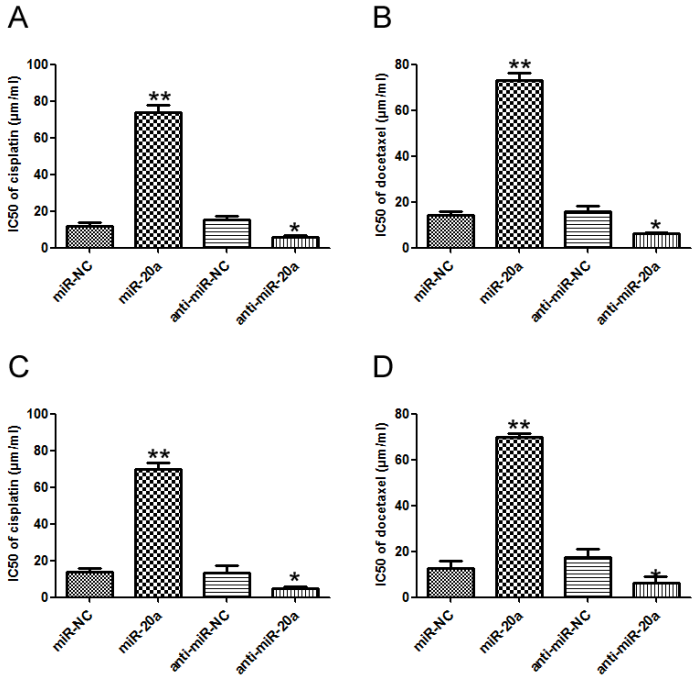
Figure 3. miR-20a promoted migration and invasion of GC cells. (**A**) *In vitro* migration assay of SGC7901 cells transfected with miR-20a precursor/inhibitor or the corresponding control, respectively; (**B**) *In vitro* invasion assay of SGC7901 cells transfected with miR-20a precursor/inhibitor or the corresponding control, respectively; (**C**) *In vitro* migration assay of MKN4 cells; (**D**) *In vitro* invasion assay of MKN4 cells. The data represented at least four measurements. * $p < 0.05$ compared with control.



2.4. miR-20a Promoted Chemoresistance of GC Cells

The effect of miR-20a on the sensitivity of GC cells to chemotherapeutic agents, cisplatin and docetaxel, was investigated. Transfection of miR-20a precursor increased the IC50 value of cisplatin, while inhibition of miR-20a decreased the IC50 value of cisplatin in SGC7901 and MKN45 cells compared with that in control group (Figure 4A,C). Similar results were obtained in docetaxel treated SGC7901 and MKN45 cells (Figure 4B,D). Our data suggested that miR-20a might promote chemoresistance of GC cells.

Figure 4. miR-20a promoted chemoresistance of GC cells. (**A**) Alteration of IC50 values (cisplatin) in SGC7901 cells transfected with miR-20a precursor/inhibitor, and the corresponding control was analyzed by MTT; (**B**) Alteration of IC50 values (docetaxel) in SGC7901 cells; (**C**) Alteration of IC50 values (cisplatin) in MKN45 cells; (**D**) Alteration of IC50 values (docetaxel) in MKN45 cells. The data represented at least four measurements. * $p < 0.05$, ** $p < 0.01$ compared with control.



2.5. EGR2 Was a Direct Target of miR-20a

To investigate the downstream target of miR-20a, Targetscan 6.2 was used to screen its target. Early growth response 2 (EGR2) was predicted to be a target of miR-20a. To confirm that, we amplified the EGR2 3' UTR containing the target sequences, or the mutants, into a luciferase reporter vector (Figure 5A). As shown in Figure 5B, miR-20a suppressed the luciferase activity of the wild type EGR2 3' UTR (WT), while mutation of the miR-20a binding sites (Mut) blocked this suppression in SGC7901 cells. Western blot demonstrated that transfection of miR-20a precursor in SGC7901 cells inhibited EGR2 expression while miR-20a inhibitor elevated EGR2 protein level (Figure 5C). qRT-PCR showed that miR-20a precursor decreased EGR2 mRNA level, while miR-20a inhibitor elevated EGR2 mRNA level (Figure 5D), indicating that miR-20a suppressed EGR2 expression post-transcriptionally.

2.6. miR-20a Was inversely Correlated with EGR2 Expression

qRT-PCR was performed to detect the mRNA levels of EGR2 in 28 GC and adjacent non-tumor normal tissue samples. EGR2 mRNA was significantly decreased in GC group (Figure 6A). Furthermore, EGR2 protein levels were also down-regulated in GC tissues compared with the non-tumor normal tissue samples (Figure 6B). Moreover, we correlated EGR2 mRNA level with miR-20a expression in the same GC tissues. EGR2 mRNA level was inversely correlated with miR-20a expression in GC tissues (Figure 6C).

Figure 5. Early growth response 2 (EGR2) was a direct target of miR-20a. (**A**) Wild type of the mutated sequences of EGR2 3' UTR (nucleotides 458-465); (**B**) SGC7901 cells were co-transfected with miR-20a precursor or negative control (miR-NC) with EGR2 3' UTR fragment with either the miR-20a target sequence (WT), or a mutant (Mut). Luciferase activity was detected; (**C**) EGR2 protein level was detected by Western blot in cells transfected with miR-20a precursor/inhibitor or the corresponding control; (**D**) expression of EGR2 mRNA was detected by qRT-PCR in cells transfected with miR-20a precursor/inhibitor or the corresponding control. GAPDH was used as an internal control. The data represented at least four measurements. * $p < 0.05$ compared with control.

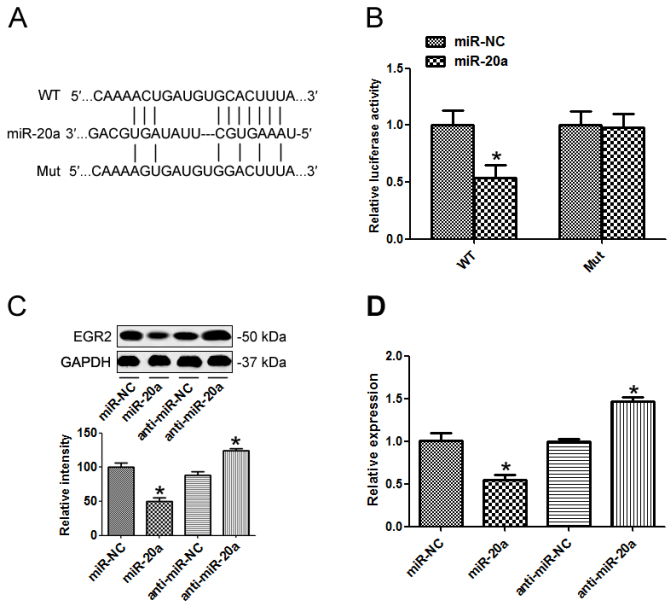
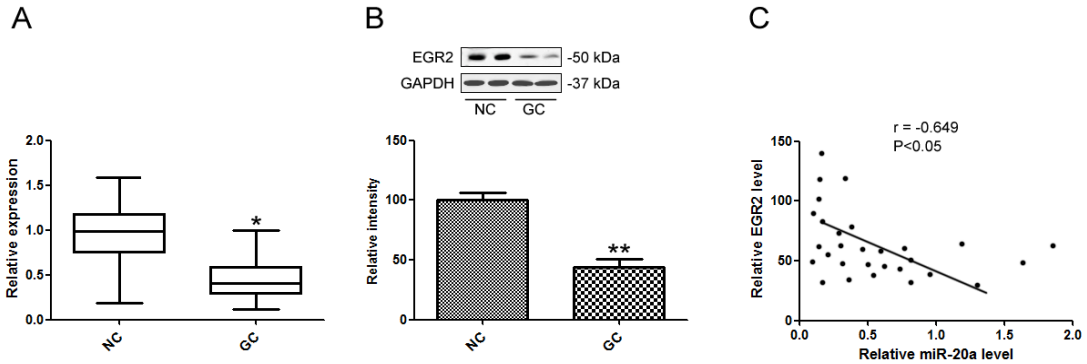


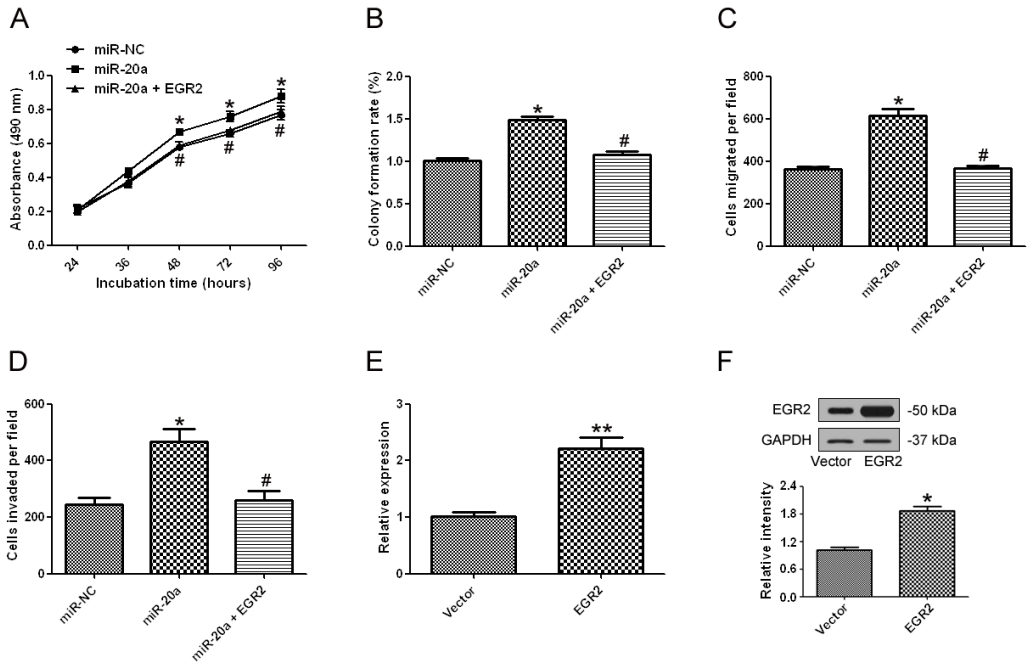
Figure 6. miR-20a was inversely correlated with EGR2 expression. (**A**) EGR2 mRNA levels in GC tissues were analyzed by qPCR; (**B**) EGR2 protein levels in GC tissues were analyzed by Western blot; (**C**) Correlation of miR-20a expression and EGR2 mRNA level was analyzed. The data represented triplicate measurements. * $p < 0.05$, ** $p < 0.01$ compared with control.



2.7. miR-20a Promoted GC Progression by Targeting EGR2

Since EGR2 was found to be a target of miR-20a, we further investigated whether overexpression of EGR2 could attenuate the oncogenic effect of miR-20a. MTT assay (Figure 7A), colony formation assay (Figure 7B), cell migration (Figure 7C) and invasion (Figure 7D) all showed that supplement of EGR2 by an EGR2 overexpression plasmid could significantly attenuate the oncogenic effect of miR-20a. The effect of EGR2 plasmid was validated by qRT-PCR (Figure 7E) and Western blot (Figure 7F). These data suggested that miR-20a promoted GC progression partially by targeting EGR2.

Figure 7. miR-20a promoted GC progression by targeting EGR2. (**A**) SGC7901 cells were co-transfected with miR-20a precursor and EGR2 overexpression plasmid or the control, at 24, 48, 36, 72, or 96 h after transfection, MTT assay was performed to examine SGC7901 proliferation; (**B**) Representative results of colony formation assay in SGC7901 cells co-transfected with miR-20a precursor and EGR2 overexpression plasmid or the control; (**C**) *In vitro* migration assay of SGC7901 cells transfected with miR-20a precursor and EGR2 overexpression plasmid or the control; (**D**) *In vitro* invasion assay of SGC7901 cells transfected with miR-20a precursor and EGR2 overexpression plasmid or the control; (**E**) qRT-PCR was used to detect the mRNA level of EGR2 in in cells transfected with EGR2 overexpression plasmid or the control; (**F**) EGR2 protein level was detected by Western blot in cells transfected with EGR2 overexpression plasmid or the control. GAPDH was used as an internal control. The data represented at least four measurements. * $p < 0.05$, ** $p < 0.01$ compared with control; $^{\#}$ $p < 0.05$, compared with miR-20a precursor transfected group.



3. Discussion

Although numerous miRNAs had been identified in GC carcinogenesis, their underlying molecular mechanisms in GC development still remained poorly understood. Hence, exploring the function of miRNAs specifically involved in GC carcinogenesis would greatly help expand our understanding of GC and screening new targets for its diagnosis and therapy [18]. The aberrant expression of miR-20a in cervical cancer, colorectal cancer, GC, and prostate cancer had been found [16,19–21]. A recent study suggested that miR-20a overexpression in rheumatoid fibroblast-like synoviocytes decreased apoptosis signal-regulating kinase 1 (ASK1) activity, indicating an anti-apoptotic effect [22]. miR-20a was up-regulated in gallbladder carcinoma [23], but down-regulated in hepatocellular carcinoma [24]. The difference indicated that dysregulation of miR-20a in different cancers depended on the cellular microenvironment.

In this study, we validated that miR-20a was increased in a number of GC tissue samples and cell lines SGC7901, MKN45, and NUGC-3 cells compared to GES-1 cells. Ectopic expression of miR-20a promoted proliferation of GC cells, while suppression of miR-20a with inhibitor had the opposite effect. miR-20a was found to inhibit EGR2 expression partially by inducing mRNA decay. Similarly Wang M and colleagues reported that the levels of circulating miR-17-5p/20a might be a molecular marker for GC [20]. They found that overexpression of miR-17-5p/20a promoted GC cell cycle progression and inhibited apoptosis, whereas knockdown of miR-17-5p/20a resulted in cell cycle arrest and increased apoptosis in AGS cells and *in vivo* as well by targeting p21 and tumor protein p53-induced nuclear protein 1 (TP53INP1) [25]. In our study, overexpression of miR-20a promoted proliferation without effect on apoptosis in SGC7901 and MKN45 cell lines *in vitro*. The contradictory results might due to different cell lines, Wang M *et al.* used AGS cell line, and we used SGC7901 and MKN45 cells. Secondly, different molecules were involved in the two studies.

The MTT and colony formation assays in GC cells all suggested that forced miR-20a overexpression promoted carcinogenesis and proliferation of GC cells. Down-regulation of miR-20a expression by miR-20a inhibitor decreased these effects. Similar effects were obtained from *in vitro* migration and invasion assays. Our study suggested that miR-20a overexpression might act as an oncogene in GC.

Drug resistance remained a major obstacle for conventional chemotherapeutic agents. miRNAs had been shown to regulate drug resistance in numerous cancers [26–28]. In the present study, we found that miR-20a overexpression could significantly elevate the IC50 values of two clinical drugs, cisplatin and docetaxel, promoting the chemoresistance of GC cells.

To determine how miR-20a acted as an oncogene, we screened potential target using bioinformatics analysis, Targetscan 6.2. As a tumor suppressor, EGR2 was down-regulated in GC. Luciferase activity assay suggested direct targeting of EGR2 by miR-20a. EGR2 mRNA level was reversely correlated with miR-20a in GC patients. More importantly, supplement of EGR2 by an EGR2 overexpression plasmid could remarkably attenuate the effect of miR-20a on GC progression. EGR2 belongs to a multi-gene family encoding C2H2-type zinc-finger proteins and plays a role in the regulation of cellular proliferation [29]. Recently, some reports suggested an essential role of EGR2 in apoptosis regulation [30]. EGR2 was a target of the p53 family, and overexpression of EGR2 led to apoptosis, while down-regulation of EGR2 attenuated p53-mediated apoptosis [31]. EGR2 could be also

regulated by other miRNAs. Wu Q *et al.* has reported that miR-150 could promote GC proliferation by negatively regulating EGR2 expression [32]. Our study shed new light on the regulation of EGR2.

In summary, our study validated that miR-20a was dramatically increased in GC tissues and cell lines and that ectopic expression of miR-20a promoted proliferation, migration, invasion and chemoresistance of GC cells. Moreover, down-regulation of miR-20a had the opposite effect on GC cells by targeting EGR2. Further studies for the functional and clinical implications of miR-20a and its target EGR2 might contribute to the early diagnosis and treatment of GC.

4. Materials and Methods

4.1. Cell lines and Tissue Samples

Human GC cell lines SGC7901, MKN45, NUGC-3, and human gastric mucosa cell line GES-1 were from the Chinese Academy of Sciences (Shanghai, China). Cells were grown in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum at 37 °C with 5% CO_2. Twenty eight paired tissues of GC and matched normal tissues (located >5 cm away from the tumor) were collected from our department. Informed consent was obtained from each patient and this work was approved by the Ethics Committee of Third Military Medical University.

4.2. Plasmids and Transfection

miR-20a precursor (Catalog# HmiR0202-MR01, shown as miR-20a in figures) and inhibitor (Catalog# HmiR-AN0312-AM01, shown as anti-miR-20a) constructs were purchased from GeneCopoeia (Rockville, MD, USA). pEGFP-N1-EGR2 plasmid was generated by using the following primers, sense 5'-CCCTCGAGATCCCAGGCTCAGTCCAACC-3', antisense 5'-CCAAGCTTAG GTGTCCGGGTCCGAGA-3'. The amplified sequences were inserted into pEGFP-N1 within XhoI/HindIII sites. Transfection was performed using Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA, USA). Transfected cells were harvested after 24 h for migration and invasion assays and after 48 h for RNA isolation and Western blot.

4.3. Quantitative Real Time PCR

Total RNA was extracted using TRIzol (Invitrogen, Carlsbad, CA, USA) then reverse-transcribed into cDNA with a TaqMan MicroRNA Reverse-Transcription Kit (Applied Biosystems, Foster City, CA, USA). PCR reactions were performed using the ABI Stepone plus Detection System (Applied Biosystems, Foster City, CA, USA). The relative expression of miRNA was normalized with U6. Samples were compared by using the relative CT method. The fold change was determined relative to a control after normalizing to a housekeeping gene by using $2^{-\Delta\Delta CT}$, where ACT is (gene of interest CT) minus (GAPDH CT), and ΔΔCT is (ΔCT cancer) minus (ΔCT control). The relative expression of EGR2 was normalized with GAPDH. All experiments were carried out at least in triplicate.

4.4. In Vitro *Cell Proliferation Assay*

Forty eight hours after transfection, cells were seeded into 96-well plates at 6×10^3 cells/well. The MTT assay was used to measure cell viability. Optical densities at 490 nm were measured.

4.5. Colony Formation Assay

Cells transfected with miR-20a precursor/inhibitor or the corresponding control were seeded in a 10 cm dish and maintained in complete culture medium. After 21 days, SGC7901 cells were fixed with methanol and then stained with 0.1% crystal violet. Colonies were manually counted.

4.6. In Vitro *Migration and Invasion Assay*

In vitro migration assay was performed using 8 μm pore size Transwell plates (Millipore, Billerica, MA, USA). After transfection, the cells (1×10^5 cells/100 μL serum-free medium) were added to the upper chamber. RPMI 1640 containing 10% FCS was added to the bottom chamber as a chemoattractant. After 24 h, cells on the upper surface were removed, while cells attached on the bottom were fixed and stained with 0.1% crystal violet. The images of invaded cells were counted with a photomicroscope (Olympus, Tokyo, Japan). For the *in vitro* invasion assay, the plates were coated with matrigel (BD Biosciences, San Jose, CA, USA) diluted in serum-free medium and performed the same as migration assay.

4.7. Luciferase Reporter Assay

The 3' UTR of human EGR2 was PCR-amplified and cloned into psiCHECK-2 vector. These constructs (1 μg) were co-transfected with 1 μg of control or miR-20a precursor into SGC7901 cell. Luciferase activity was assayed 48 h after transfection by using the Dual-luciferase activity assay system (Promega, Madison, WI, USA). All the experiments were performed at least in four times.

4.8. Western Blot

Cells were harvested 48 h after transfection, and then proteins were extracted and separated by 10% SDS-PAGE. Gels were transferred to nitrocellulose filter membrane and probed with anti-EGR2 or anti-GAPDH. Horseradish peroxidase-conjugated secondary antibodies were applied and detected by enhanced chemiluminescence (Thermo Scientific, Rockford, IL, USA).

4.9. Flow Cytometric Analysis of Apoptosis

Cells were harvested at the above indicated time points, at least 5×10^5 cells were recovered by centrifugation for evaluation of apoptotic cells with the use of double staining with annexin V-fluoresein isothiocyanate (annexin V-FITC) and propidium iodide (PI) (BioVision, St. Pete Beach, FL, USA) according to the manufacturer's instructions, followed by flow cytometric analysis with the use of Cell Quest software (Version 5.1, Becton, Rutherford, NJ, USA).

4.10. Cell Cycle Analysis by Flow Cytometry

Cells were harvested 48 h after transfection and fixed in 75% ethanol. Then, cells were treated with RNase A and PI (50 μg/mL) and incubated for 30 min. Cell cycle analysis was performed by Flow cytometry.

4.11. Statistical Analysis

All experiments were repeated independently at least in triplicate, and the results were expressed as the mean ± SD. The results were assessed by a one-way ANOVA, or Student *t* test. A value of $p < 0.05$ was accepted to indicate statistical significance.

5. Conclusions

Our work demonstrated that miR-20a was involved in the carcinogenesis of GC through modulation of the EGR2 signaling pathway.

Conflicts of Interest

The authors have no conflict of interest.

References

1. Jemal, A.; Bray, F.; Center, M.M.; Ferlay, J.; Ward, E.; Forman, D. Global cancer statistics. *CA Cancer J. Clin.* **2011**, *61*, 69–90.
2. Ali, Z.; Deng, Y.; Ma, C. Progress of research in gastric cancer. *J. Nanosci. Nanotechnol.* **2012**, *12*, 8241–8248.
3. Link, A.; Kupcinskas, J.; Wex, T.; Malfertheiner, P. Macro-role of microRNA in gastric cancer. *Dig. Dis.* **2012**, *30*, 255–267.
4. Song, J.H.; Meltzer, S.J. microRNAs in pathogenesis, diagnosis, and treatment of gastroesophageal cancers. *Gastroenterology* **2012**, *143*, 35–47.
5. Liu, H. microRNAs in breast cancer initiation and progression. *Cell. Mol. Life Sci.* **2012**, *69*, 3587–3599.
6. Rossi, S.; di Narzo, A.F.; Mestdagh, P.; Jacobs, B.; Bosman, F.T.; Gustavsson, B.; Majoie, B.; Roth, A.; Vandesompele, J.; Rigoutsos, I.; *et al*. microRNAs in colon cancer: A roadmap for discovery. *FEBS Lett.* **2012**, *586*, 3000–3007.
7. Du, L.; Pertsemlidis, A. microRNA regulation of cell viability and drug sensitivity in lung cancer. *Expert Opin. Biol. Ther.* **2012**, *12*, 1221–1239.
8. O'Kelly, F.; Marignol, L.; Meunier, A.; Lynch, T.H.; Perry, A.S.; Hollywood, D. microRNAs as putative mediators of treatment response in prostate cancer. *Nat. Rev. Urol.* **2012**, *9*, 397–407.
9. Guo, M.M.; Hu, L.H.; Wang, Y.Q.; Chen, P.; Huang, J.G.; Lu, N.; He, J.H.; Liao, C.G. miR-22 is down-regulated in gastric cancer, and its overexpression inhibits cell migration and invasion via targeting transcription factor Sp1. *Med. Oncol.* **2013**, *30*, 542.
10. Yao, Q.; Cao, Z.; Tu, C.; Zhao, Y.; Liu, H.; Zhang, S. microRNA-146a acts as a metastasis suppressor in gastric cancer by targeting WASF2. *Cancer Lett.* **2013**, *335*, 219–224.
11. Zheng, L.; Qi, T.; Yang, D.; Qi, M.; Li, D.; Xiang, X.; Huang, K.; Tong, Q. microRNA-9 suppresses the proliferation, invasion and metastasis of gastric cancer cells through targeting cyclin D1 and Ets1. *PLoS One* **2013**, *8*, e55719.

12. Wang, F.; Li, T.; Zhang, B.; Li, H.; Wu, Q.; Yang, L.; Nie, Y.; Wu, K.; Shi, Y.; Fan, D. microRNA-19a/b regulates multidrug resistance in human gastric cancer cells by targeting PTEN. *Biochem. Biophys. Res. Commun.* **2013**, *434*, 688–694.
13. Wang, M.; Li, C.; Yu, B.; Su, L.; Li, J.; Ju, J.; Yu, Y.; Gu, Q.; Zhu, Z.; Liu, B. Overexpressed miR-301a promotes cell proliferation and invasion by targeting RUNX3 in gastric cancer. *J. Gastroenterol.* **2013**, [Epub ahead of print].
14. An, J.; Pan, Y.; Yan, Z.; Li, W.; Cui, J.; Jiao, Y.; Tian, L.; Xing, R.; Lu, Y. miR-23a in Amplified 19p13.13 loci targets metallothionein 2A and promotes growth in gastric cancer cells. *J. Cell. Biochem.***2013**, *114*, 2160–2169.
15. Yan, H.J.; Liu, W.S.; Sun, W.H.; Wu, J.; Ji, M.; Wang, Q.; Zheng, X.; Jiang, J.T.; Wu, C.P. miR-17-5p inhibitor enhances chemosensitivity to gemcitabine via upregulating Bim expression in pancreatic cancer cells. *Dig. Dis. Sci.* **2012**, *57*, 3160–3167.
16. Kang, H.W.; Wang, F.; Wei, Q.; Zhao, Y.F.; Liu, M.; Li, X.; Tang, H. miR-20a promotes migration and invasion by regulating TNKS2 in human cervical cancer cells. *FEBS Lett.* **2012**, *586*, 897–904.
17. Huang, G.; Nishimoto, K.; Zhou, Z.; Hughes, D.; Kleinerman, E.S. miR-20a encoded by the miR-17-92 cluster increases the metastatic potential of osteosarcoma cells by regulating Fas expression. *Cancer Res.* **2012**, *72*, 908–916.
18. Gigek, C.O.; Chen, E.S.; Calcagno, D.Q.; Wisnieski, F.; Burbano, R.R.; Smith, M.A. Epigenetic mechanisms in gastric cancer. *Epigenomics* **2012**, *4*, 279–294.
19. Li, X.; Pan, J.H.; Song, B.; Xiong, E.Q.; Chen, Z.W.; Zhou, Z.S.; Su, Y.P. Suppression of CX43 expression by miR-20a in the progression of human prostate cancer. *Cancer Biol. Ther.* **2012**, *13*, 890–898.
20. Wang, M.; Gu, H.; Wang, S.; Qian, H.; Zhu, W.; Zhang, L.; Zhao, C.; Tao, Y.; Xu, W. Circulating miR-17-5p and miR-20a: Molecular markers for gastric cancer. *Mol. Med. Rep.* **2012**, *5*, 1514–1520.
21. Chai, H.; Liu, M.; Tian, R.; Li, X.; Tang, H. miR-20a targets BNIP2 and contributes chemotherapeutic resistance in colorectal adenocarcinoma SW480 and SW620 cell lines. *Acta Biochim. Biophys. Sin.* **2011**, *43*, 217–225.
22. Philippe, L.; Alsaleh, G.; Pichot, A.; Ostermann, E.; Zuber, G.; Frisch, B.; Sibilia, J.; Pfeffer, S.; Bahram, S.; Wachsmann, D.; *et al*. miR-20a regulates ASK1 expression and TLR4-dependent cytokine release in rheumatoid fibroblast-like synoviocytes. *Annu. Rheum. Dis.* **2013**, *72*, 1071–1079.
23. Chang, Y.; Liu, C.; Yang, J.; Liu, G.; Feng, F.; Tang, J.; Hu, L.; Li, L.; Jiang, F.; Chen, C.; *et al*. miR-20a triggers metastasis of gallbladder carcinoma. *J. Hepatol.* **2013**, doi:10.1016/j.jhep.2013.04.034.
24. Fan, M.Q.; Huang, C.B.; Gu, Y.; Xiao, Y.; Sheng, J.X.; Zhong, L. Decrease expression of microRNA-20a promotes cancer cell proliferation and predicts poor survival of hepatocellular carcinoma. *J. Exp. Clin. Cancer Res.* **2013**, *32*, 21.
25. Wang, M.; Gu, H.; Qian, H.; Zhu, W.; Zhao, C.; Zhang, X.; Tao, Y.; Zhang, L.; Xu, W. miR-17-5p/20a are important markers for gastric cancer and murine double minute 2 participates in their functional regulation. *Eur. J. Cancer* **2013**, *49*, 2010–2021.

26. Wang, F.; Song, X.; Li, X.; Xin, J.; Wang, S.; Yang, W.; Wang, J.; Wu, K.; Chen, X.; Liang, J.; *et al*. Noninvasive visualization of microRNA-16 in the chemoresistance of gastric cancer using a dual reporter gene imaging system. *PLoS One* **2013**, *8*, e61792.
27. Tekiner, T.A.; Basaga, H. Role of microRNA deregulation in breast cancer cell chemoresistance and stemness. *Curr. Med. Chem.* **2013**, [Epub ahead of print].
28. Wang, Y.Q.; Guo, R.D.; Guo, R.M.; Sheng, W.; Yin, L.R. microRNA-182 promotes cell growth, invasion, and chemoresistance by targeting programmed cell death 4 (PDCD4) in human ovarian carcinomas. *J. Cell. Biochem.* **2013**, *114*, 1464–1473.
29. Joseph, L.J.; le Beau, M.M.; Jamieson, G.A., Jr.; Acharya, S.; Shows, T.B.; Rowley, J.D.; Sukhatme, V.P. Molecular cloning, sequencing, and mapping of EGR2, a human early growth response gene encoding a protein with "zinc-binding finger" structure. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 7164–7168.
30. Dzialo-Hatton, R.; Milbrandt, J.; Hockett, R.D., Jr.; Weaver, C.T. Differential expression of Fas ligand in Th1 and Th2 cells is regulated by early growth response gene and NF-AT family members. *J. Immunol.* **2001**, *166*, 4534–4542.
31. Yokota, I.; Sasaki, Y.; Kashima, L.; Idogawa, M.; Tokino, T. Identification and characterization of early growth response 2, a zinc-finger transcription factor, as a p53-regulated proapoptotic gene. *Int. J. Oncol.* **2010**, *37*, 1407–1416.
32. Wu, Q.; Jin, H.; Yang, Z.; Luo, G.; Lu, Y.; Li, K.; Ren, G.; Su, T.; Pan, Y.; Feng, B.; *et al*. miR-150 promotes gastric cancer proliferation by negatively regulating the pro-apoptotic gene EGR2. *Biochem. Biophys. Res. Commun.* **2010**, *392*, 340–345.

Reprinted from *IJMS*. Cite as: Nishizawa, T.; Suzuki, H. The Role of microRNA in Gastric Malignancy. *Int. J. Mol. Sci.* **2013**, *14*, 9487-9496.

Review

The Role of microRNA in Gastric Malignancy

Toshihiro Nishizawa [1,2] and Hidekazu Suzuki [1,*]

[1] Division of Gastroenterology and Hepatology, Department of Internal Medicine, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan; E-Mail: nisizawa@kf7.so-net.ne.jp

[2] Division of Gastroenterology, National Hospital Organization Tokyo Medical Center, 2-5-1 Higashigaoka, Meguro-ku, Tokyo 152-8902, Japan

* Author to whom correspondence should be addressed; E-Mail: hsuzuki@a6.keio.jp; Tel.: +81-3-5363-3914; Fax: +81-3-5363-3967.

Received: 27 March 2013; in revised form: 19 April 2013 / Accepted: 25 April 2013 / Published: 29 April 2013

Abstract: *Helicobacter pylori* (*H. pylori*) infection is the main cause of gastritis, gastro-duodenal ulcer, and gastric cancer. MicroRNAs (miRNAs) are small noncoding RNAs that function as endogenous silencers of numerous target genes. Many miRNA genes are expressed in a tissue-specific manner and play important roles in cell proliferation, apoptosis, and differentiation. Recent discoveries have shed new light on the involvement of miRNAs in gastric malignancy. However, at the same time, several miRNAs have been associated with opposing events, leading to reduced inflammation, inhibition of malignancy, and increased apoptosis of transformed cells. The regulation of miRNA expression could be a novel strategy in the chemoprevention of human gastric malignancy. In this article, the biological importance of miRNAs in gastric malignancy is summarized.

Keywords: microRNA; *Helicobacter pylori*; gastric cancer

1. Introduction

Helicobacter pylori (*H. pylori*) infection is one of the most prevalent infectious diseases worldwide, and is estimated to affect 40%–50% of the world population [1]. *H. pylori* has been identified as a group 1 carcinogen by the World Health Organization and is associated with the development of gastric cancer [2]. *H. pylori* eradication has been shown to have a prophylactic effect on gastric cancer [3,4].

MicroRNAs (miRNAs) are noncoding RNAs comprising 18–24 nucleotides that can post-transcriptionally downregulate various target genes [5]. It is estimated that the human genome encodes more than one thousand miRNAs, targeting 30%–60% of all protein-coding genes [6]. They are expressed in a tissue-specific manner, and play important roles in cell proliferation, apoptosis, and differentiation [5,7]. Moreover, recent studies have shown a connection between aberrant expression of miRNAs and the development of cancer. In this article, the biological importance of miRNAs in gastric malignancy is summarized.

2. *Helicobacter pylori* and miRNA

CagA of *H. pylori* is a bacterium-derived oncogenic protein closely associated with the development of gastric cancers [8]. After injection into host cells using a type IV secretion system, CagA is phosphorylated at tyrosine residues by the c-Src and Lyn kinases. Phosphorylated CagA then activates the Src homology-2 domain-containing phosphatase 2 (SHP2), which activates the Erk1/2 pathway. CagA translocated into CD44v9-positive gastric cancer stem-like cells is thought to escape from reactive oxygen species-dependent autophagy, resulting in gastric carcinogenesis [9].

Table 1. MicroRNAs (miRNAs) change in response to *Helicobacter pylori.*

miRNAs	Change	Target mRNAs	Biological process targeted
let-7a	↓	**RAB40C** **HMGA2**	**Cell cycle progression** **Invasion**
let-7b/d/e/f	↓	HMGA2	Invasion
miR101	↓	**MCL1**	**Apoptosis**
miR-106b	↓	p21 BIM	Cell cycle progression Apoptosis
miR-125a	↓	ERBB2	Proliferation
miR-141	↓	**FGFR2**	**Proliferation**
miR-200a	↓	ZEB1, ZEB2	Epithelial to mesenchymal transition
miR-200b/c	↓	BCL2, XIAP	Apoptosis
miR-203	↓	**ABL1**	**Proliferation, Invasion**
miR-204	↓	EZR	Proliferation
miR-218	↓	**ROBO1**	**Invasion, Metastasis**
miR-375	↓	**PDK1, 12-3-3** **JAK2**	**Apoptosis** **Proliferation**
miR-429	↓	BCL2, XIAP MYC	Apoptosis Proliferation
miR-17	↑	**p21**	**Cell cycle progression**
miR-20a	↑	**p21**	**Cell cycle progression**
miR-21	↑	**PTEN** **RECK**	**Proliferation** **Metastasis**
miR-146a	↑	**IRAK1, TRAF6** **SMAD4**	**Proliferation, Immune response** **Apoptosis**
miR-155	↑	**IKK-ε, SMAD4** **FADD, PLIα**	**Immune response** **Apoptosis**
miR-223	↑	**EPB41L3**	**Invasion, Metastasis**

↓: miRNA is downregulated in response to *H. pylori*; ↑: miRNA is upregulated in response to *H. pylori*; Bold indicates miRNA changes in the same way in gastric cancer.

miRNA changes in response to *H. pylori* infection are summarized in Table 1. *miR-584* and *miR-1290* expression are upregulated in CagA-transformed cells. The *miR-584* and *miR-1290* target is Foxa1, and knockdown of Foxa1 promotes the epithelial-mesenchymal transition (EMT). Overexpression of *miR-584* and *miR-1290* induces intestinal metaplasia of gastric epithelial cells. These results indicate that *miR-584* and *miR-1290* interfere with cell differentiation and lead to remodeling of the gastric mucosal tissues [10].

CagA enhances *c-myc* and DNA methyltransferase 3B, and attenuates *miR-6a* and *miR-101* expression, which results in the attenuation of *let-7* expression by histone and DNA methylation. CagA induces aberrant epigenetic silencing of *let-7* expression, leading to Ras pathway activation. Thus, the miRNA pathway is a new pathogenic mechanism for CagA.

Shiotani *et al.* [11] reported that the expression of oncogenic miRNAs (*miR-17/92* and the *miR-106b-93-25* cluster, *miR-21, miR-194*, and *miR-196*) is significantly higher in the intestinal metaplasia than in the non-intestinal metaplasia. *H. pylori* eradication improves miRNA deregulation, but not in the intestinal metaplasia. Long-term colonization of *H. pylori* might induce an epigenetic modification of gastric mucosal genes, including the promoter of tumor suppressor miRNAs, but this is not completely reversible by bacterial eradication alone. Epigenetic therapy in severe atrophic or metaplastic gastric mucosa after *H. pylori* eradication might be a possible option for gastric cancer prevention.

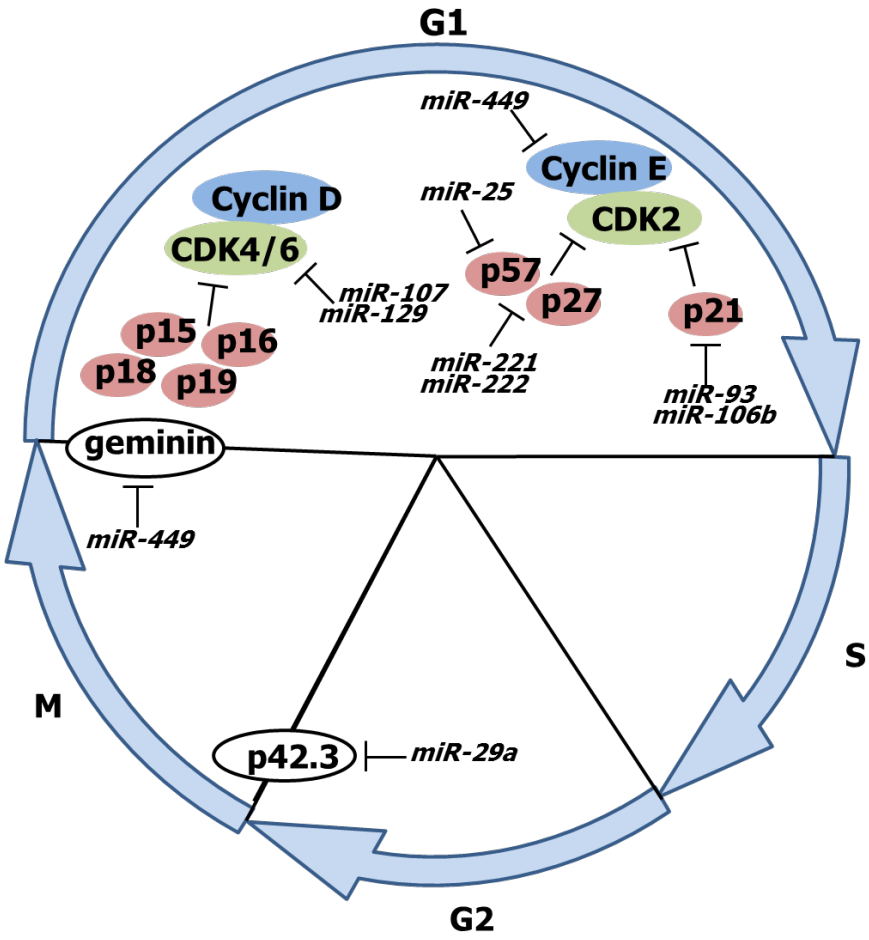
3. Cell Cycle Progression and miRNA

The dysregulation of cell cycle progression is a hallmark of malignancy. Cyclin-CDK (cyclin-dependent kinase) complexes regulate this progression through the cell cycle. miRNA dysregulation promotes cell cycle progression by upregulating cyclin expression or downregulating expression of CDK inhibitors (p57, p21, *etc.*) in numerous malignancies (Figure 1). *miR-449* is downregulated in *H. pylori*-infected gastric mucosa and in gastric cancer and targets cyclin E2 and geminin. Both cyclin E2 and geminin are overexpressed in various malignancies and promote G1/S and M/G1 cell cycle progression [6]. Consequently, downregulation of *miR-449*, as occurs following *H. pylori* infection, promotes cell cycle progression and proliferation through the upregulation of cyclin E2 and geminin.

G2/M cell cycle progression and proliferation in gastric cancer cells are regulated by p42.3 [12]; *miR-29a* is significantly downregulated in gastric cancer and targets p42.3 [13] (Figure 1). Thus, the downregulation of *miR-29a* results in a reciprocal increase in p42.3 expression, promoting increased cell cycle progression and proliferation.

Both *miR-93* and *miR-106b* directly target p21, resulting in its transcriptional silencing and impairment of its tumor-suppressing activity [14]. In addition, *miR-25* targets p57, while *miR-221* and *miR-222* target p27 and p57 [15] (Figure 1). These oncogenic miRNA clusters are also significantly upregulated in gastric cancer [16]. Overexpression of most of these miRNAs results in activation of CDK2, thereby promoting G1/S phase progression.

Figure 1. Regulation of cell cycle progression. Cyclins and cyclin-dependent kinases (CDKs) determine a cell's progress through the cell cycle.

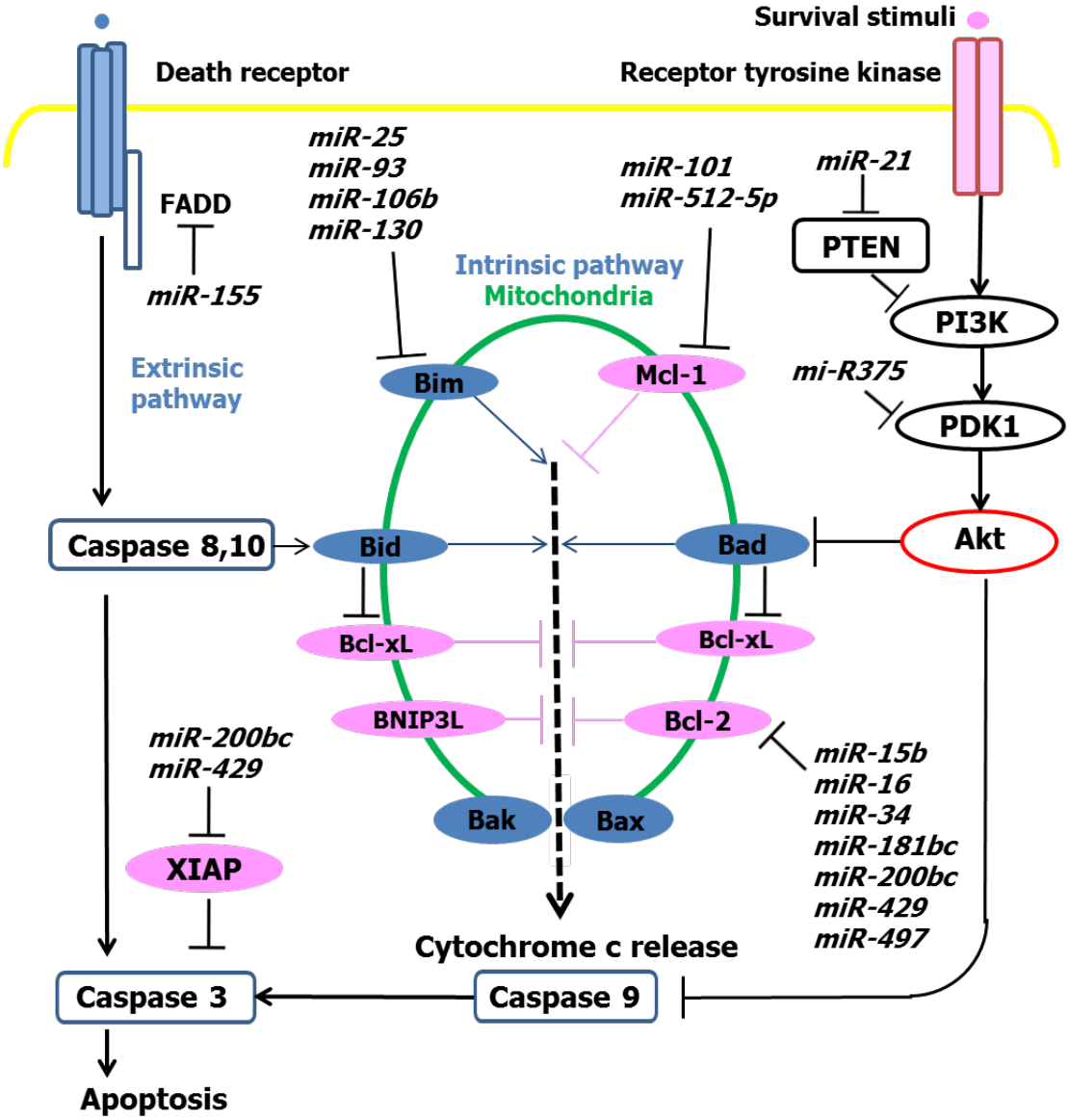


4. Inhibition of Apoptosis and miRNA

Evasion of apoptosis is a common feature of malignancy. Apoptosis is classifiable as either intrinsic or extrinsic pathway-dependent. The extrinsic apoptosis pathway is initiated on the cell surface through the activation of specific pro-apoptotic death receptors. Tumor necrosis factors are cytokines produced mainly by activated macrophages that bind the death receptors as their ligands. Ligand binding induces receptor clustering and the recruitment of the adaptor protein Fas-associated death domain (FADD), leading to induction of caspases and ultimately cell death. *miR-155* targets *FADD*, leading to decreased expression of this key adaptor molecule (Figure 2) [17]. Therefore, the upregulation of *miR-155* by *H. pylori* and during carcinogenesis results in the downregulation of FADD and the inhibition of apoptosis.

The intrinsic pathway is initiated within cells and hinges on the balance of activity between pro-apoptotic (e.g., Bax, Bak, Bim, BNIP3L, and Bid) and anti-apoptotic (e.g., Bcl-2, Bcl-xL, and Mcl-1) proteins from the Bcl-2 (B cell lymphoma 2) superfamily. Some miRNAs overexpressed in gastric cancer function as oncogenic miRNAs by targeting members of the pro-apoptotic proteins. *miR-25*, *miR-93*, *miR-106b*, and *miR-130* inhibit apoptosis by preventing the expression of the pro-apoptotic protein, Bim (Figure 2) [14].

Figure 2. Signaling cascades that regulate the intrinsic and extrinsic pathways of apoptosis. Receptor tyrosine kinase detects survival stimuli such as growth factors and induces phosphoinositide 3-kinase (PI3K)/Akt signaling cascades that ultimately result in the inhibition of the pro-apoptotic protein, Bad. Pro-apoptotic and anti-apoptotic proteins govern the intrinsic cell death pathway, which results in the release of cytochrome *c* from the mitochondria and induction of the caspase cascade. Signaling through death receptors initiates the extrinsic pathway of apoptosis.



Tumor suppressor miRNAs (*miR-15b*, *miR-16*, *miR-34*, *miR-181b*, *miR-181c,* and *miR-497*) target anti-apoptotic Bcl-2. These miRNA clusters are downregulated in gastric cancer cells, leading to increased expression of Bcl-2 and inhibition of apoptosis. The *miR-200bc/429* cluster is downregulated in *H. pylori*-infected gastric mucosa, and these miRNAs directly target Bcl-2 and XIAP (x-linked inhibitor of apoptosis) [18]. *miR-101* and *miR-515-5p* target Mcl-1, which are downregulated in gastric cancer, lead to increased levels of Mcl-1 and an anti-apoptotic phenotype (Figure 2).

In addition to targeting proteins directly involved in the intrinsic and extrinsic cell death pathways, miRNAs target other factors that ultimately lead to apoptosis inhibition and increased proliferation.

miR-21 targets PTEN (phosphatase and tensin homolog), a tumor suppressor and negative regulator of the PI3K/Akt signaling pathway. *miR-21* is upregulated in gastric cancer, and its overexpression shifts the balance between proliferation and apoptosis, by increasing cellular proliferation and inhibiting apoptosis.

miR-375 targets 3-phophoinositide dependent protein kinase (PDK1), a kinase that directly phosphorylates Akt, thereby regulating the PI3K/Akt signaling pathway. Overexpression of *miR-375* reduces cell viability and *miR-375* is downregulated in gastric cancer (Figure 2) [19].

5. Metastasis and miRNA

Some miRNAs that are known to regulate cell cycle progression and apoptosis pathways are also involved in invasion and metastasis. *miR-181b* is aberrantly overexpressed in *H. pylori* infection and gastric cancer tissues [20]. Cell proliferation, migration, and invasion in the gastric cancer cells were significantly increased after *miR-181b* transfection, and the number of apoptotic cells was also increased. Furthermore, overexpression of *miR-181b* downregulated the protein levels of tissue inhibitor of metalloproteinase 3 (TIMP). The upregulation of *miR-181b* may play an important role in the progression of gastric cancer and *miR-181b* may be a potential molecular target for gastric cancer therapies.

miR-218 is reduced significantly in gastric cancer tissues, *H. pylori*-infected gastric mucosa, and *H. pylori*-infected AGS cells [21]. *miR-218*, a tumor suppressor miRNA, is downregulated in gastric cancer, which correlates with increased metastasis and cancer invasion [22]. This downregulation is thought to occur through the direct targeting of roundabout homolog (ROBO1), which leads to enhanced signaling through the ROBO1 receptor. The SLIT/ROBO signaling pathway is implicated in many biological responses through regulating cell migration. Thus, disruption of this signaling cascade can result in increased invasion and metastasis. *miR-21* also targets RECK (reversion-inducing-cysteine-rich protein with kazal motifs), a tumor and metastasis suppressor that inhibits tumor metastasis and angiogenesis through modulation of matrix metalloproteinases. Recently, Li *et al.* indicated that *miR-21*, *miR-218*, and *miR-223* may be potential biomarkers for gastric cancer detection [23].

miR-148a is downregulated in gastric cancer. The protein interaction network regulated by *miR-148a* is associated with metastasis-related function, such as integrin-mediated signaling, cell-matrix adhesion, and blood coagulation [24]. A single miRNA can provoke a chain reaction and further affect the protein interaction network. This interactive network-based approach could provide insight into carcinogenesis.

6. miRNA and Anticancer Therapy

miRNAs are promising molecular targets for anticancer therapeutics in gastric cancer. Kim *et al.* reported that *miR-10b* was silenced in gastric cancer cells by promoter methylation. *miR-10b* targets the oncogene that encodes microtubule-associated protein, RP/EB family member 1. After 5-aza-2'-deoxycytidine treatment of gastric cancer cells, *miR-10b* methylation is significantly decreased, and the expression of *miR-10b* is restored. The modulation of *miR-10b* may represent a therapeutic approach for treating gastric cancer [25].

Runx3 is an important tumor suppressor that is inactivated in gastric cancer, and promoter hypermethylation of Runx3 is frequent [26]. 5-aza-2'-deoxycytidine treatment reactivates the expression of Runx3. Lai *et al.* reported that *miR-130b* expression is upregulated in gastric cancer, and this is inversely associated with Runx3 hypermethylation. *miR-130b* overexpression increases cell viability, reduces cell death, and decreases the expression of Bim in TGF-beta mediated apoptosis, subsequent to the downregulation of Runx3 protein expression. The attenuation of Runx3 protein levels by miRNA may reduce the growth suppressive potential of Runx3 and contribute to tumorigenesis [27]. Wang *et al.* reported that *miR-301a* is upregulated in gastric cancer, and directly downregulates Runx3 expression [28].

The selective cyclooxygenase-2 (COX-2) inhibitor celecoxib is a potential drug for the treatment of gastrointestinal tumors. We investigated the role of miRNAs in gastric carcinogenesis and the feasibility of a new therapeutic approach for gastric cancer [29]. miRNA microarray analysis revealed that *miR-29c* is significantly downregulated in gastric cancer tissues relative to non-tumor gastric mucosa [29]. *miR-29c* is significantly activated by celecoxib in gastric cancer cells (AGS) [29]. Celecoxib activation of *miR-29c* induces suppression of the oncogene *Mcl-1*, a target of *miR-29c* and apoptosis in gastric cancer cells. These results suggest that the downregulation of the *miR-29c* tumor suppressor plays a critical role in the progression of gastric cancer. As such, selective COX-2 inhibitors may be a clinical option for the treatment of gastric cancer via restoration of *miR-29c*.

7. MALT Lymphoma and miRNA

Gastric B-cell lymphoma of the mucosa-associated lymphoid tissue (MALT lymphoma) develops in the chronically inflamed mucosa of patients infected with the bacterial pathogen. In 60%–80% of these cases, the *H. pylori*-positive gastric MALT lymphoma regresses after *H. pylori* eradication. The t(11;18) (q21;q21) translocation is associated with *API2-MALT1* fusion, and this translocation responds only rarely or not at all to *H. pylori* eradication.

We previously reported that a hematopoietic-specific miRNA, *miR-142*, and an oncogenic miRNA, *miR-155*, are overexpressed in MALT lymphoma lesions [30]. *miR-142-5p* and *miR-155* suppress the proapoptotic gene *TP53INP1* as their target [30]. The expression levels of *miR-142-5p* and *miR-155* are significantly increased in MALT lymphomas that do not respond to *H. pylori* eradication [30]. The expression levels of *miR-142-5p* and *miR-155* are associated with the clinical courses of gastric MALT lymphoma cases and these miRNAs may have a potential application as novel biomarkers for gastric MALT lymphoma.

Craig *et al.* reported the strong downregulation of the putative tumor suppressor miRNA, *miR-203*, in human MALT lymphoma samples, which results from extensive promoter hypermethylation of the *miR-203* locus and coincides with the deregulation of its target, *ABL1*. Treatment of lymphoma B cells with demethylating agents leads to increased *miR-203* expression and concomitant downregulation of ABL1, confirming the effectiveness of epigenetic regulation of this miRNA. These results show that the transformation from gastritis to MALT lymphoma is epigenetically regulated by *miR-203* promoter methylation and identifies ABL1 as a novel target for treatment [31].

Although generally considered an indolent disease, MALT lymphoma may have the ability to transform into gastric diffuse large B-cell lymphoma (gDLBCL). Craig *et al.* reported that Myc

overexpression is detected in 80% of gDLBCLs, but only 20% of MALT lymphomas are spotted on a tissue microarray. FoxP1 overexpression is detectible in gDLBCL, but not in gastric MALT lymphoma. *miR-34a* is downregulated in malignant lymphoma, as are the targets of *miR-34a*, Myc and FoxP1 and *miR-34a* shows strong antiproliferative properties when overexpressed in DLBCL cells. *miR-34a* replacement therapy is therefore a promising strategy in lymphoma treatment [32].

8. Conclusions

As we are just beginning to understand the relationship between miRNAs and gastric malignancies and the number of identified miRNA genes is increasing, there is a potential for a large number of therapeutic targets and biomarkers in this area. Further studies are necessary to investigate whether miRNA-oriented therapy is an effective strategy for the chemoprevention of gastric malignancies.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Suzuki, H.; Iwasaki, E.; Hibi, T. *Helicobacter pylori* and gastric cancer. *Gastric Cancer* **2009**, *12*, 79–87.
2. Suzuki, H.; Nishizawa, T.; Hibi, T. *Helicobacter pylori* eradication therapy. *Future Microbiol.* **2010**, *5*, 639–648.
3. Nishizawa, T.; Suzuki, H.; Nakagawa, I.; Minegishi, Y.; Masaoka, T.; Iwasaki, E.; Hibi, T. Early *Helicobacter pylori* eradication restores sonic hedgehog expression in the gastric mucosa of Mongolian gerbils. *Digestion.* **2009**, *79*, 99–108.
4. Fukase, K.; Kato, M.; Kikuchi, S.; Inoue, K.; Uemura, N.; Okamoto, S.; Terao, S.; Amagai, K.; Hayashi, S.; Asaka, M. Effect of eradication of *Helicobacter pylori* on incidence of metachronous gastric carcinoma after endoscopic resection of early gastric cancer: An open-label, randomised controlled trial. *Lancet* **2008**, *372*, 392–397.
5. Saito, Y.; Suzuki, H.; Hibi, T. The role of microRNAs in gastrointestinal cancers. *J. Gastroenterol.* **2009**, *44*, 18–22.
6. Noto, J.M.; Peek, R.M. The role of microRNAs in *Helicobacter pylori* pathogenesis and gastric carcinogenesis. *Front Cell Infect Microbiol.* **2011**, *1*, 21.
7. Saito, Y.; Suzuki, H.; Tsugawa, H.; Suzuki, S.; Matsuzaki, J.; Hirata, K.; Hibi, T. Dysfunctional gastric emptying with down-regulation of muscle-specific microRNAs in *Helicobacter pylori*-infected mice. *Gastroenterology* **2011**, *140*, 189–198.
8. Suzuki, H.; Nishizawa, T.; Tsugawa, H.; Mogami, S.; Hibi, T. Roles of oxidative stress in stomach disorders. *J. Clin. Biochem. Nutr.* **2012**, *50*, 35–39.
9. Tsugawa, H.; Suzuki, H.; Saya, H.; Hatakeyama, M.; Hirayama, T.; Hirata, K.; Nagano, O.; Matsuzaki, J.; Hibi, T. Reactive oxygen species-induced autophagic degradation of *Helicobacter pylori* caga is specifically suppressed in cancer stem-like cells. *Cell Host Microbe* **2012**, *12*, 764–777.

10. Zhu, Y.; Jiang, Q.; Lou, X.; Ji, X.; Wen, Z.; Wu, J.; Tao, H.; Jiang, T.; He, W.; Wang, C.; *et al.* MicroRNAs up-regulated by CagA of *Helicobacter pylori* induce intestinal metaplasia of gastric epithelial cells. *PLoS One* **2012**, *7*, e35147.
11. Shiotani, A.; Uedo, N.; Iishi, H.; Murao, T.; Kanzaki, T.; Kimura, Y.; Kamada, T.; Kusunoki, H.; Inoue, K.; Haruma, K. *H. pylori* eradication did not improve dysregulation of specific oncogenic miRNAs in intestinal metaplastic glands. *J. Gastroenterol.* **2012**, *47*, 988–998.
12. Xu, X.; Li, W.; Fan, X.; Liang, Y.; Zhao, M.; Zhang, J.; Tong, W.; Wang, J.; Yang, W.; Lu, Y. Identification and characterization of a novel p42.3 gene as tumor-specific and mitosis phase-dependent expression in gastric cancer. *Oncogene* **2007**, *26*, 7371–7379.
13. Cui, Y.; Su, W.Y.; Xing, J.; Wang, Y.C.; Wang, P.; Chen, X.Y.; Shen, Z.Y.; Cao, H.; Lu, Y.Y.; Fang, J.Y. MiR-29a inhibits cell proliferation and induces cell cycle arrest through the downregulation of p42.3 in human gastric cancer. *PLoS One* **2011**, *6*, e25872.
14. Kan, T.; Sato, F.; Ito, T.; Matsumura, N.; David, S.; Cheng, Y.; Agarwal, R.; Paun, BC.; Jin, Z.; Olaru, A.V.; *et al.* The miR-106b-25 polycistron, activated by genomic amplification, functions as an oncogene by suppressing p21 and Bim. *Gastroenterology* **2009**, *136*, 1689–1700.
15. Kim, Y.K.; Yu, J.; Han, T.S.; Park, S.Y.; Namkoong, B.; Kim, D.H.; Hur, K.; Yoo, M.W.; Lee, H.J.; Yang, H.K.; *et al.* Functional links between clustered microRNAs: Suppression of cell-cycle inhibitors by microRNA clusters in gastric cancer. *Nucleic Acids Res.* **2009**, *37*, 1672–1681.
16. Volinia, S.; Calin, G.A.; Liu, C.G.; Ambs, S.; Cimmino, A.; Petrocca, F.; Visone, R.; Iorio, M.; Roldo, C.; Ferracin, M.; *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 2257–2261.
17. Xiao, B.; Liu, Z.; Li, B.S.; Tang, B.; Li, W.; Guo, G.; Shi, Y.; Wang, F.; Wu, Y.; Tong, W.D.; *et al.* Induction of microRNA-155 during *Helicobacter pylori* infection and its negative regulatory role in the inflammatory response. *J. Infect Dis.* **2009**, *200*, 916–925.
18. Zhu, W.; Xu, H.; Zhu, D.; Zhi, H.; Wang, T.; Wang, J.; Jiang, B.; Shu, Y.; Liu, P. miR-200bc/429 cluster modulates multidrug resistance of human cancer cell lines by targeting BCL2 and XIAP. *Cancer Chemother. Pharmacol.* **2012**, *69*, 723–731.
19. Tsukamoto, Y.; Nakada, C.; Noguchi, T.; Tanigawa, M.; Nguyen, L.T.; Uchida, T.; Hijiya, N.; Matsuura, K.; Fujioka, T.; Seto, M.; *et al.* MicroRNA-375 is downregulated in gastric carcinomas and regulates cell survival by targeting PDK1 and 14-3-3zeta. *Cancer Res.* **2010**, *70*, 2339–2349.
20. Guo, J.X.; Tao, Q.S.; Lou, P.R.; Chen, X.C.; Chen, J.; Yuan, G.B. miR-181b as a potential molecular target for anticancer therapy of gastric neoplasms. *Asian Pac. J. Cancer Prev.* **2012**, *13*, 2263–2267.
21. Gao, C.; Zhang, Z.; Liu, W.; Xiao, S.; Gu, W.; Lu, H. Reduced microRNA-218 expression is associated with high nuclear factor kappa B activation in gastric cancer. *Cancer* **2010**, *116*, 41–49.
22. Tie, J.; Pan, Y.; Zhao, L.; Wu, K.; Liu, J.; Sun, S.; Guo, X.; Wang, B.; Gang, Y.; Zhang, Y.; *et al.* MiR-218 inhibits invasion and metastasis of gastric cancer by targeting the Robo1 receptor. *PLoS Genet.* **2010**, *6*, e1000879.
23. Li, B.S.; Zhao, Y.L.; Guo, G.; Zhu, E.D.; Luo, X.; Mao, X.H.; Zou, Q.M.; Yu, P.W.; Zuo, Q.F.; Li, N.; *et al.* Plasma microRNAs, miR-223, miR-21 and miR-218, as novel potential biomarkers for gastric cancer detection. *PLoS One* **2012**, *7*, e41629.

24. Tseng, C.; Lin, C.C.; Chen, C.N.; Huang, H.C.; Juan, H.F. Integrative network analysis reveals active microRNA and their functions in gastric cancer. *BioMed. Central Syst. Biol.* **2011**, *5*, 1–11.
25. Kim, K.; Lee, H.C.; Park, J.L.; Kim, M.; Kim, S.Y.; Noh, S.M.; Song, K.S.; Kim, J.C.; Kim, Y.S. Epigenetic regulation of microRNA-10b and targeting of oncogenic MAPRE1 in gastric cancer. *Epigenetics* **2011**, *6*, 740–751.
26. Suzuki, M.; Suzuki, H.; Minegishi, Y.; Ito, K.; Nishizawa, T.; Hibi, T. *H. pylori*-eradication therapy increases runx3 expression in the glandular epithelial cells in enlarged-fold gastritis. *J. Clin. Biochem. Nutr.* **2010**, *46*, 259–264.
27. Lai, K.W.; Koh, K.X.; Loh, M.; Tada, K.; Subramaniam, M.M.; Lim, X.Y.; Vaithilingam, A.; Salto-Tellez, M.; Iacopetta, B.; Ito, Y.; *et al.* MicroRNA-130b regulates the tumour suppressor RUNX3 in gastric cancer. *Eur. J. Cancer* **2010**, *46*, 1456–1463.
28. Wang, M.; Li, C.; Yu, B.; Su, L.; Li, J.; Ju, J.; Yu, Y.; Gu, Q.; Zhu, Z.; Liu, B.; Overexpressed miR-301a promotes cell proliferation and invasion by targeting RUNX3 in gastric cancer. *J. Gastroenterol.* **2013**, in press.
29. Saito, Y.; Suzuki, H.; Imaeda, H.; Matsuzaki, J.; Hirata, K.; Tsugawa, H.; Hibino, S.; Kanai, Y.; Saito, H.; Hibi, T. The tumor suppressor microRNA-29c is downregulated and restored by celecoxib in human gastric cancer cells. *Int. J. Cancer* **2013**, *132*, 1751–1760.
30. Saito, Y.; Suzuki, H.; Tsugawa, H.; Imaeda, H.; Matsuzaki, J.; Hirata, K.; Hosoe, N.; Nakamura, M.; Mukai, M.; Saito, H.; *et al.* Overexpression of miR-142-5p and miR-155 in Gastric Mucosa-Associated Lymphoid Tissue (MALT) lymphoma resistant to *Helicobacter pylori* eradication. *PLoS One* **2012**, *7*, e47396.
31. Craig, V.J.; Cogliatti, S.B.; Rehrauer, H.; Wundisch, T.; Muller, A. Epigenetic silencing of microRNA-203 dysregulates ABL1 expression and drives *Helicobacter*-associated gastric lymphomagenesis. *Cancer Res.* **2011**, *71*, 3616–3624.
32. Craig, V.J.; Cogliatti, S.B.; Imig, J.; Renner, C.; Neuenschwander, S.; Rehrauer, H.; Schlapbach, R.; Dirnhofer, S.; Tzankov, A.; Muller, A. Myc-mediated repression of microRNA-34a promotes high-grade transformation of B-cell lymphoma by dysregulation of FoxP1. *Blood* **2011**, *117*, 6227–6236.

Reprinted from *IJMS*. Cite as: Garitano-Trojaola, A.; Agirre, X.; Prósper, F.; Fortes, P. Long Non-Coding RNAs in Haematological Malignancies. *Int. J. Mol. Sci.* **2013**, *14*, 15386-15422.

Review

Long Non-Coding RNAs in Haematological Malignancies

Andoni Garitano-Trojaola [1], Xabier Agirre [1], Felipe Prósper [1,2] and Puri Fortes [3,*]

[1] Laboratory of Myeloproliferative Syndromes, Oncology Area, Foundation for Applied Medical Research, University of Navarra, Pamplona 31008, Spain; E-Mails: agaritano@alumni.unav.es (A.G.-T.); xaguirre@unav.es (X.A.); fprosper@unav.es (F.P.)

[2] Hematology Service and Area of Cell Therapy, University of Navarra Clinic, University of Navarra, Pamplona 31008, Spain

[3] Department of Hepatology and Gene Therapy, Foundation for Applied Medical Research, University of Navarra, Pamplona 31008, Spain

***** Author to whom correspondence should be addressed; E-Mail: pfortes@unav.es; Tel.: +34-948-194-700; Fax: +34-948-194-718.

Received: 31 May 2013; in revised form: 28 June 2013 / Accepted: 9 July 2013 / Published: 24 July 2013

Abstract: Long non-coding RNAs (lncRNAs) are functional RNAs longer than 200 nucleotides in length. LncRNAs are as diverse as mRNAs and they normally share the same biosynthetic machinery based on RNA polymerase II, splicing and polyadenylation. However, lncRNAs have low coding potential. Compared to mRNAs, lncRNAs are preferentially nuclear, more tissue specific and expressed at lower levels. Most of the lncRNAs described to date modulate the expression of specific genes by guiding chromatin remodelling factors; inducing chromosomal loopings; affecting transcription, splicing, translation or mRNA stability; or serving as scaffolds for the organization of cellular structures. They can function in *cis*, cotranscriptionally, or in *trans*, acting as decoys, scaffolds or guides. These functions seem essential to allow cell differentiation and growth. In fact, many lncRNAs have been shown to exert oncogenic or tumor suppressor properties in several cancers including haematological malignancies. In this review, we summarize what is known about lncRNAs, the mechanisms for their regulation in cancer and their role in leukemogenesis, lymphomagenesis and hematopoiesis. Furthermore, we discuss the potential of lncRNAs in diagnosis, prognosis and therapy in cancer, with special attention to haematological malignancies.

Keywords: lncRNAs; leukemia; hematologic malignancies

1. Introduction

Transcriptome analysis by tiling arrays and RNA sequencing has led to the amazing conclusion that while 70%–90% of the genome is transcribed, only 2% is dedicated to the transcription of protein coding sequences [1]. This result has caused a great impression in a scientific community that is deeply proteocentric, *i.e.*, is dedicated to the study of proteins and generally does not pay much attention to other molecules such as lipids or RNAs.

Most cellular RNA is composed of highly expressed non-coding RNAs whose relevance in cell functionality has been well-known for years. However, their transcription requires a relatively small proportion of the genome. These housekeeping non-coding RNAs include transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs), required for mRNA translation; small nuclear RNAs (snRNAs), essential for splicing; and small nucleolar RNAs (snoRNAs), involved in RNA modification. More recently, several small RNAs have been described as playing essential roles in gene expression and transposon silencing. These include microRNAs (miRNAs), small interfering RNAs (siRNAs) and piwi interacting RNAs (piRNAs). Less clear is the role and the molecular mechanisms involved in the function of other small RNAs derived from retrotransposons or 3' untranslated regions or associated with transcription start sites, promoters, termini or repeats. All these non-coding RNAs, with the exception of some of the housekeeping RNAs (some rRNAs and a few snRNAs and snoRNAs), share the common characteristic of being smaller than 200 nts. Therefore the remaining non-coding RNAs, longer than 200 nts, have been grouped under the name of long non-coding RNAs (lncRNAs).

LncRNAs have a terrible name. They are not really long, just longer than the limit of 200 nts imposed by small RNAs. In fact, the average size of coding mRNAs is near 2500 nts while the average length of all the lncRNAs recently described by the Encode project is less than 600 nts [2]. Thus, most of the long non-coding RNAs are shorter than the coding mRNAs, even if some of the lncRNAs may be longer than 100 kbs. Apart from not being really long, it is difficult to determine whether lncRNAs are indeed non-coding. Traditionally, lncRNAs have been characterized by what they do not have: they lack open reading frames (ORFs) longer than 100 amino acids, conserved codons and homology to protein databases [3,4]. Therefore, they have poor coding potential, although they could still code for small open reading frames or non-conserved peptides. Some authors have also analyzed coding capacities of specific lncRNAs by matching their sequences with ribosome footprints or peptide fragments from mass spectrometry analysis. Hits would indicate translation [5–8]. In spite of these efforts, it should be borne in mind that what makes lncRNAs interesting for most scientists is not whether they can encode for proteins or not but the fact that they are functional as RNA molecules. The demonstration of function as an RNA should be required for annotation as an lncRNA, as a functional long RNA is the best definition for lncRNAs. To complicate things further, there are several cases of coding mRNAs that contain regulatory RNA elements and act as bifunctional RNAs; on one hand they code for a protein (p53, for instance) and on the other hand they have a function as RNAs [9–14]. Furthermore, several coding genes are transcribed to non-coding alternative splicing variants.

Functional or lncRNA genes are very similar to coding genes at the DNA and chromatin level as they share the same epigenetic marks. Similar to mRNAs, most lncRNAs are transcribed from RNA polymerase II, are capped at the 5' end, contain introns and approximately 40% are polyadenylated at the 3' end [15]. The lncRNAs recently described by Encode show a bias for having just one intron and a trend for less-efficient cotranscriptional splicing [8,16]. It has been estimated that there could be as many lncRNA genes as coding genes, but the number of lncRNAs is still growing and some authors consider that it could increase from ~20,000 to ~200,000 [17,18]. Compared to mRNAs, most lncRNAs localize preferentially to the nucleus, are more cell type specific and are expressed at lower levels [19]. In fact, there is less than one copy per cell of many lncRNAs. The low expression levels and the fact that the sequence of lncRNAs is poorly conserved have convinced many scientists that they are not relevant for cell functionality. However, although lncRNAs are under lower selective pressure than protein-coding genes, sequence analysis shows that lncRNAs are under higher selective pressure than ancestral repeat sequences with neutral selection. Moreover, promoters of lncRNAs have similar selection levels than promoters of protein coding genes [8]. Even in the absence of strong sequence conservation, the genomic location and structure of many lncRNAs is conserved together with short stretches of sequences, suggesting that lncRNAs could be under selective pressure to maintain a functional RNA structure rather than a linear sequence [8].

Recent publications in the field have led to the hypothesis that many lncRNAs may be key regulators of development and may play relevant roles in cell homeostasis and proliferation. In fact, several lncRNAs have been described that function as oncogenes or tumor suppressors [20]. It is expected that for cell biology the role of lncRNAs could be as revolutionary as the role of small non-coding RNAs such as miRNAs. miRNA studies have highlighted the relevance of gene regulation in cell homeostasis, differentiation and proliferation and may impact the clinic with new therapies and new diagnostic and prognostic tools for many diseases. The relevance of miRNAs has been clearly established for haematological malignancies [21,22]. In this review we will summarize what is known about lncRNAs in normal haematopoiesis and in haematological tumors. Even though many more studies need to be done, the results obtained thus far suggest that several lncRNAs may be key molecules in haematopoiesis and in the pathogenesis of haematological malignancies.

2. Classes of lncRNAs and lncRNA Functionality

2.1. Classification by Genomic Location

Under the name of lncRNAs there are RNAs with many different characteristics, which complicates classification. Therefore a well accepted method is based on genomic location rather than on functionality, conservation or origin. From a genetic point of view lncRNAs can be classified into one or more of the following categories: (a) sense, when overlapping with one or more exons of another transcript in the same strand; (b) antisense, when overlapping with one or more exons of another transcript in the opposite strand; (c) intronic, when derived from an intron of another transcript; (d) divergent or bidirectional, when they share a promoter with another transcript in the opposite strand and therefore are coregulated; (e) intergenic, when they are independent, located in between two other genes. Long intergenic non coding RNAs (lincRNAs) are a special class of intergenic lncRNAs whose

genes have histone mark signatures of active transcription (trimethylation in lysine 4 and lysine 36 of histone 3: H3K4m3, H3K36m3) [23].

In the case of antisense transcripts, classification based on genomic location helps to predict functionality. 50%–70% of sense transcripts have natural antisense partners (NATs) [24–26]. NATs are generally involved in the regulation in *cis* of the corresponding sense RNA by mechanisms that act at the transcriptional and posttranscriptional level. NATs can induce transcriptional interference or recruit chromatin modifiers and remodelers to establish a local transcriptionally active or inactive chromatin conformation [27]. Posttranscriptionally, examples of NATs exist that regulate imprinting, RNA editing, splicing, by blocking binding of the spliceosome to the 5' splice site of an intron leading to intron retention [28–32] or translation and stability by forming a duplex with the sense RNA that masks the binding site for miRNAs [33]. Thus, NATs can modify processing and induce or reduce the expression or the translation of their sense counterpart. Some intronic lncRNAs also regulate the expression of their genomic partners. Intronic lncRNAs may be generated by stabilization of the intron after splicing of the host gene but, more commonly, they are produced from independent transcription. Some intronic non-coding RNAs are associated with polycomb-related repressive histone marks along the promoter region and gene body of their host gene, which results in local transcriptional silencing [34].

2.2. Classification by Specific Characteristics

Most lncRNAs with special characteristics cannot be easily classified into a single group according to genomic location. These include enhancer RNAs (eRNAs), lncRNA-activating (lncRNA-a) genes, transcribed ultraconserved regions (T-UCRs), pseudogenes, telomere-associated ncRNAs (TERRAs), circular RNAs, *etc.*

eRNAs are transcribed by RNA polymerase II at active enhancer regions, characterized by H3 Lys4 monomethylation or Lys27 acetylation and binding of the regulatory protein p300 [35–39]. eRNAs are not polyadenylated. Many are bidirectional and poorly expressed [38,40], but expression of several eRNAs seems to be tightly regulated [38,39]. Although many eRNAs were thought to be by-products of the presence of RNA pol II in enhancers, recent evidence suggests that some may function to control the expression of neighbouring genes [41]. eRNAs are also important in the formation of the chromosomal loopings that bring enhancers closer to promoters [39], and in the induction of, for example, p53-dependent enhancer activity and transcription [42].

lncRNA-a genes generally transcribe intergenic RNAs which are involved in the expression of neighbouring genes [41]. Thus, downregulation of the lncRNA-a results in downregulation of the neighbour gene. This effect requires expression of the Mediator complex and it has been shown that interaction of the lncRNA-a with Mediator is required for the upregulation of nearby genes [43].

T-UCRs and pseudogenes are lncRNAs that share sequence similarity to other mammalian genomes or other regions of the same genome, respectively. There are 481 UCRs longer than 200 bp that are absolutely conserved between human, rat, and mouse genomes [44]. Most are transcribed or T-UCRs in normal human tissues, both ubiquitously and tissue specifically. The high degree of conservation across species implies that T-UCRs may be essential, but deletion of some of these regions in knockout mice has not been associated with a detectable phenotype [45]. One possible function of some T-UCRs is miRNA control, as many T-UCRs have significant antisense complementarity with particular

miRNAs and there is a negative correlation between expression of specific T-UCRs and predicted antisense miRNAs targets [46,47]. In fact, some T-URCs have been shown to be targeted by miRNAs.

Pseudogenes originated from duplication of ancestor or parental coding genes (duplicated pseudogenes) or through retrotransposition of processed RNAs transcribed from ancestor genes (processed pseudogenes). Subsequently, they have lost their coding capacity as a result of the accumulation of mutations. When pseudogenes are expressed, they may regulate the expression and function of their parental gene by several mechanisms [48,49]. For instance, pseudogenes may act as miRNA decoys that lead to increased stability and translation of their parental gene [50–53].

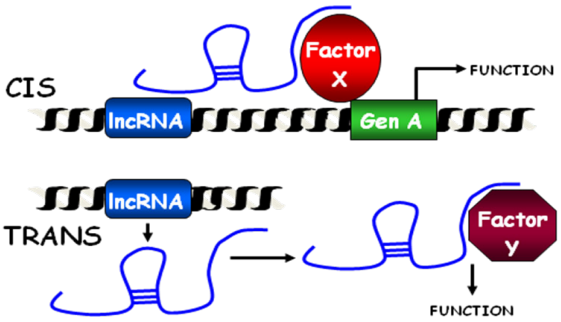
Circular RNAs, newcomers to the RNA list, can also function as RNA decoys [54–56]. It is generally accepted that circular RNAs originate from reverse splicing, where the acceptor splice site located downstream binds to an upstream donor splice site. This causes the circularization of the RNA and a tremendous increase in RNA stability, as circular RNAs lack 5' or 3' ends and therefore, are resistant to exonucleases. The increased stability of circular RNAs may lead to long-term functionality by miRNA sequestration [57].

2.3. Classification as cis *or* trans-*Acting Molecules*

LncRNAs can also be classified according to their functionality as *cis* and/or *trans* acting molecules (Figure 1). *Trans*-acting lncRNAs function away from the site of synthesis while *cis*-acting lncRNAs function at the site of transcription to affect the expression of neighbouring genes. Several *cis*-acting lncRNAs guide epigenetic regulators to their site of transcription while they are being transcribed. Thus, lncRNA transcription is critical and rapidly creates an anchor to recruit proteins involved in chromatin remodelling [58–61]. This molecular mechanism has tremendous advantages: (i) it responds very fast, as it only requires transcription of an RNA and a proper accumulation of nuclear chromatin remodelers; (ii) it is very specific, as the targeting does not involve RNA-DNA interactions other than those required for lncRNA transcription and (iii) it may function with just a single molecule of lncRNA per locus. This may explain the low abundance of *cis*-acting lncRNAs and the relatively high concentration of lncRNAs close to developmental genes whose expression is strictly controlled [62]. Thus, *cis*-acting lncRNAs control the epigenetic regulation of some imprinted genes. Imprinting depends on the parental origin of the imprinted genes, which play critical roles in mammalian development and therefore, their expression must be tightly regulated [63]. Many imprinted gene loci express lncRNAs that appear to regulate the expression of neighbouring imprinted protein-coding genes in *cis*, allele specifically [64]. The lncRNA *AIR*, for example, silences the neighbouring imprinted genes *SLC22A3*, *SLC22A2* and *IGF2R* [65].

The clear division between *cis* and *trans* acting lncRNAs has been blurred by recent experiments, where exogenously expressed lncRNAs that normally work in *cis*, are able to find their target sites. Thus, even *cis*-acting lncRNAs may have the capacity to act in *trans* [65]. Furthermore, when considering *cis*-acting lncRNAs, the 3D organization of the genome should be taken into consideration. A *cis*-acting lncRNA may control the expression of neighbour genes brought into proximity by chromosome looping.

Figure 1. Schematic representation of *cis* and *trans*-acting lncRNAs. *cis*-acting lncRNAs function at the site of transcription and affect the expression of neighbouring genes. *Trans*-acting lncRNAs function away from the site of synthesis.



Trans-acting lncRNAs regulate gene expression on a genome-wide scale. A good example is *HOTAIR*, which binds the chromatin-modifying complexes PRC2, LSD1 and CoREST/REST [66–69]. Guiding chromatin remodelers to specific sites is easier to conceive for *cis*-acting lncRNAs. Targeting mediated by *trans*-acting lncRNAs would probably require RNA:DNA:DNA triplex formation via Hoogsteen base-pairing, as has been shown *in vitro* for a promoter-associated lncRNA [70]. However, such interactions may expose the genome to deamination and damage [71,72]. Furthermore, lncRNAs could form secondary and tertiary structures that behave similary to DNA-binding domains from proteins or that bind proteins that mediate DNA binding. This is what has been described for the *XIST* lncRNA, which binds YY1 transcription factor to reach specific sites in the X chromosome [73]. Theoretically, lncRNAs could also form an RNA:DNA hybrid that displaces a single strand of DNA (the so-called R-loop) or an RNA:RNA hybrid of lncRNA with a nascent transcript [74–76].

2.4. lncRNA Functionality

Guiding chromatin remodelling factors seems to be the predominant function exerted by lncRNAs. In fact, it has been estimated that 20% of all lncRNAs may bind PRC2 [66]. Several lncRNAs have also been shown to bind to PRC1, the CoREST/REST repressor complex [66], the histone methyltransferase associated with the activating trithorax complex, MLL1 [77,78], and H3-K9 methyltransferase, G9a [65,79]. However, lncRNAs have also been shown to exert several other functions in the cell nucleus and cytoplasm, including regulation of DNA bending and insulation, RNA transcription, splicing, translation and stability, organization of subnuclear structures and protein localization, among others.

DNA looping. CTCF can induce chromosomal bending and protect specific genes from the effects of distal enhancers and regulatory elements. The lncRNA *SRA* can interact with and enhance the function of CTCF [80]. Also, endogenous but not exogenous nascent *HOTTIP* lncRNA, binds target genes via chromosomal looping [81].

Transcription. LncRNAs may activate or inhibit transcription of specific targets. Some lncRNAs act as coactivators that bind transcription factors and enhance their transcriptional activity [82–84]. This is the function of *SRA* lncRNA in the progestin steroid hormone receptor [85,86]. However, some lncRNAs act as decoys of transcription factors [87] and may move them to the cytoplasm to keep them away from their nuclear targets [88]. Thus, p53-induced lncRNA *PANDA* binds transcription factor

NF-YA and prevents NF-YA activation of cell death genes [89]. *DHFR* lncRNA forms a triplex structure which sequesters the general transcription factor IIB and prevents transcription of the *DHFR* coding gene [90]. Finally, the act of lncRNA transcription may interfere with transcription initiation, elongation or termination of another sense or antisense gene [91]. Transcriptional interference can also lead to activation of gene expression by inhibiting the action of repressor elements.

Organization of subnuclear structures. LncRNAs can recruit protein factors to nuclear structures. This is the case of lncRNA *MALAT1* and *NEAT-1*. *MALAT1* recruits serine/arginine–rich splicing factors to nuclear speckles [92]. More importantly, *NEAT-1* is an essential structural component of paraspeckles, subnuclear structure implicated in RNA splicing and editing [93,94]. Depletion of *NEAT-1* leads to loss of paraspeckles while overexpression of *NEAT-1* causes an increase in the number of paraspeckles [95–97]. *MALAT1* and *NEAT-1* are genomic neighbours overexpressed in several tumors compared to healthy tissues. Surprisingly the mouse knockouts of either *NEAT-1* or *MALAT1* had no detectable phenotype, suggesting that there could be redundant or compensatory molecules [98–101].

Splicing. Splicing can be inhibited by lncRNAs antisense to intron sequences that impede spliceosome binding causing intron retention [28–32]. Furthermore, alternative splicing can be altered by lncRNA-mediated sequestration or modification of splicing factors. Thus, *MALAT1* binds splicing factors present in nuclear speckles and modulates the activity of SR proteins, involved in the selection of splice sites, and therefore regulates the splicing of many pre-mRNAs [92]. Some snoRNA-containing lncRNAs (sno-lncRNAs) are retained close to their sites of transcription where the splicing factor Fox2 is enriched. Changes in the level of the sno-lncRNA lead to a nuclear redistribution of Fox2 and to changes in alternative splicing. Thus, the sno-lncRNAs could function as a regulator of splicing in specific subnuclear domains [102].

Translation. LncRNAs have been described that increase or inhibit translation of specific targets [103,104]. Expression of antisense *UCHL1* lncRNA leads to an increase in Uchl1 protein level without any change at the Uchl1 mRNA level. A repetitive SINEB2 sequence is required for this function. Under cap dependent translation inhibition due to stress, *UCHL1* lncRNA moves from the nucleus to the cytoplasm, binds to Uchl1 mRNA and allows its cap-independent translation. Thus, *UCHL1* lncRNA could behave as a mobile internal ribosomal entry sequence.

Stability. LncRNAs have been described that increase or decrease stability of specific targets [105,106]. Binding of lncRNAs containing ancestral Alu repeats to complementary Alu sequences in the 3´UTR of coding mRNAs forms a dsRNA recognized by the dsRNA binding protein Stau1, which induces Stau-mediated RNA decay [106]. Instead, lncRNA *TINCR* localizes to the cytoplasm, where it interacts with Stau1 and promotes the stability of mRNAs containing the TINCR box motif [105].

miRNA binding. LncRNAs can regulate mRNA stability and translation by binding to miRNAs and preventing their action. Besides the already described role of some pseudogenes and circular lncRNAs in miRNA sequestration, other lncRNAs such as *linc-MD1*, have been shown to serve as “sponge” for miRNAs. *Linc-MD*1 binds two miRNAs, which downregulate transcription factors involved in muscle differentiation and therefore muscle differentiation is induced upon *Linc-MD*1 expression [107].

LncRNAs have been implicated in many other different functions. LncRNA *NRON* is a repressor of NFAT by binding β-importins and regulating the nuclear trafficking of NFAT [88]. *TERC* is a

well-known telomerase-associated lncRNA that serves as a template for the synthesis of chromosome ends. The dsRNA-protein kinase PKR may be activated by binding to a lncRNA [108]. It is expected that in the near future novel and unexpected mechanisms of lncRNA functionality will be discovered. For instance, to date few lncRNAs have been described to have catalytic properties.

The high number of lncRNAs and their heterogeneity helps them to exert such a myriad of functions. In fact, all lncRNA functions respond to just three different mechanisms: decoys, scaffolds and guides [109]. Decoy-acting lncRNAs impede the access of proteins such as transcription factors and RNAs such as miRNAs to their targets. LncRNAs *MD-1* and *PANDA* act as decoys for miRNAs and transcription factors, respectively [89,107]. Scaffold-acting lncRNAs serve as adaptors to bring two or more factors into discrete ribonucleoproteins (RNPs) [110]. LncRNA *TERC*, *HOTAIR* or *NEAT-1* act as scaffolds to form the telomerase complex [111], a silencing complex [69] or the paraspeckle, respectively [93,94]. Guide-acting lncRNAs are required to localize protein complexes at specific positions. *XIST* or *AIR* lncRNAs act as guides to target gene silencing activity in an allele-specific manner. Guide lncRNAs such as *HOTAIR*, can also behave as scaffolds.

It is conceivable that lncRNAs may function through linear or structured domains. Linear domains may bind proteins but also RNA or, possibly, DNA sequences by perfect (e.g., antisense lncRNAs with their sense counterpart) or imperfect complementarity. Novel linear domains able to bind and regulate mRNAs, miRNAs or other lncRNAs could be very easily created evolutionarily. In many cases though, the secondary and tertiary structure of lncRNAs dictates their function. Thus, lncRNAs generally have complex structures with higher folding energies than those observed in mRNAs [112]. Proteins are expected to be the major partners of lncRNAs to form functional RNP particles. RNA binding proteins represent more than 15% of the total amount of proteins [113]. In several cases studied to date, interaction between proteins and RNAs results in conformational changes to the protein, the RNA or both, which could endow the complex with a novel ability.

LncRNA function impacts cell behaviour. LncRNAs have specially emerged as regulators of development. Some transcription factors involved in pluripotency bind promoter regions of more than 100 mouse lncRNAs [15]. 26 lincRNAs have already been described as being required for the maintenance of pluripotency in mouse [114]. Two lncRNAs regulated by pluripotency transcription factors such as Oct4 and Nanog are essential for pluripotency maintenance, as they, in turn, control the expression of Oct4 and Nanog [115]. Therefore, these lncRNAs participate in positive regulatory loops. Similarly, several lncRNAs have been implicated in human disease, including several cancers [116]. Dysregulated lncRNAs have been described in heart disease, Alzheimer disease, psoriasis, spinocerebellar ataxia and fragile X syndrome [33,117–121] and in several tumours including breast, brain, lung, colorectal, prostate and liver cancers, melanoma, leukaemia and others [46,68,116,122–128]. LncRNAs have been described that function as oncogenes [129], tumour suppressors [23,130] or drivers of metastatic transformation, such as *HOTAIR* in breast cancer [68]. In this review we will concentrate on those lncRNAs whose expression is altered in haematological malignancies.

3. LncRNAs Deregulated in Haematological Malignancies

The impact of non-coding RNAs on haematological malignancies has been well described for microRNAs [131,132]. The list of lncRNAs involved in the initiation and progression of blood tumors is still very short and expected to grow exponentially in the near future. Some of the lncRNAs that

play a role in haematological malignancies (Table 1) are in fact host genes of miRNAs with oncogenic or tumour suppressor properties. Others endow oncogenic or tumour suppressor properties in the long non-coding RNA molecule. The mechanism of action of few of them has been studied in some detail.

3.1. Host Genes of Small RNAs

3.1.1. BIC and C13ORF25

Some lncRNAs were described to have oncogenic properties in blood cells before the discovery of miRNAs. This is the case of the B cell Integration cluster (*BIC*) or host gene *mir-155* (*MIR155HG*) (Figure 2A). *BIC* and *miR-155* expression is increased in Hodgkin lymphoma, Acute Myeloid Leukemia (AML) and Chronic Lymphocytic Leukemia (CLL) but it is not detected in healthy samples [133]. Increased expression of BIC and *miR-155* results from transcription activation by the MYB transcription factor [134] and leads to *miR-155*-mediated downregulation of several tumor suppressor genes [135]. In this case, the lncRNA *BIC* plays an important role in the regulation of *miR-155* which is directly involved in the lymphomagenesis or leukemogenesis. Similarly, *C13ORF25* or host gene *mir-17* (*MIR17HG*) encodes the *miR-17-92* cluster and its expression is increased in B-cell lymphoma [136], Mantle Cell Lymphoma (MCL) [137] and other tumors [138,139].

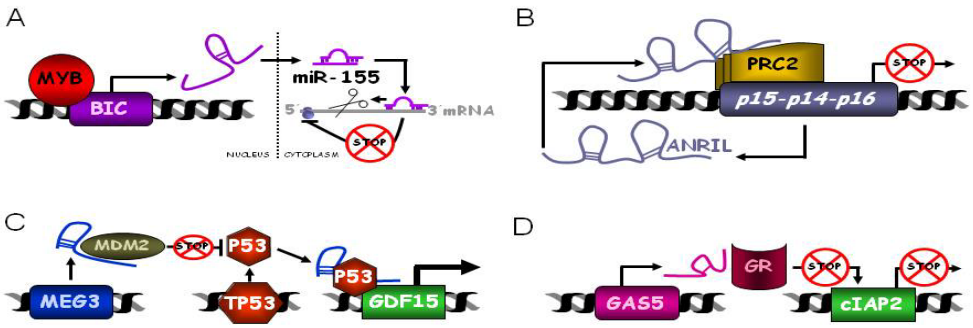
3.1.2. nc886 or *vtRNA2-1*

vtRNA2-1, previously known as *pre-miR-886*, is a short ncRNA suppressed in a wide range of cancer cells that inhibits activation of protein kinase R (PKR) [140]. Even if nc886 is shorter than 200 nts and therefore is not a lncRNA, its relevance in AML merits a short description. *vtRNA2-1* is transcribed from the long arm of chromosome 5 region whose deletion is associated with poor outcome in AML. Furthermore, decreased expression by monoallelic or biallelic DNA methylation correlates with a worse outcome in AML patients [141]. Thus, *vtRNA2-1* could be a tumour suppressor for AML and its role could be mediated by PKR.

Table 1. lncRNAs in hematopoiesis and hematological malignancies.

LncRNAs	LOCATION	HEMATOLOGIC DISEASE/SYSTEM	FUNCTION	MOLECULAR MECHANISM	MECHANISMS INVOLVED IN DYSREGULATION	CITATIONS
MIR155HG BIC	21q21.3	Burkitt, Hodgkin lymphoma, AML, CLL	Host of miRNAs	miR-155	Target MYB and NFKB	[134]
MIR17HG	13q31.3	B-cell lymphoma, MCL	Host of miRNAs	miR-17-92	Target MYC	[136,137,142]
vtRNA2-1	5q31.1	AML (poor prognosis)		PKR inhibition	DNA methylation Deletion 5q	[140,141]
PVT1	8q24.21	MM, Burkitt Lymphoma, T-cell Leukemia, CLL	Oncogene and host of miRNAs	miR-1204 MYC activation	Translocation t(8;14)(q24;q11) t(2;8)(p11;q24) t(8;22)(q24;q11)	[143–148]
CDKN2B-AS1/ANRIL	9p21.3	AML, ALL	Oncogene	PRC1 and PRC2 targeting	rs3731217-G SNP Deletion, hypermethylation	[26,128,149–151]
MEG3	14q32.2	AML,MM	Tumor suppressor	PRC2 binding to control DLK1 imprinting. p53 activation.	DNA methylation	[149,150,152–155]
DLEU1/DLEU2	13q14.2	CLL, MM, Lymphoma	Tumor suppressor	*hsa-miR-16-1* and *15a* BCL2 targeting. NFKB activation	Histone modification, DNA methylation, deletion	[156]
GAS5	1q25.1	B-cell Lymphoma, Leukemia	Tumor suppressor	Glucorticoid receptor repression. Regulated by mTOR pathway.	Translocation (1;3)(q25;q27)	[87]
H19	11p15.5	AML, CML, MPN, T-cell Leukemia, Lymphoma	Oncogene/tumor suppressor	Activated by Myc and down-regulated by p53. miR-675 targeting Rb	Loss of imprinting	[157]
T-UCRs		CLL (prognosis marker)	Oncogene/tumor suppressor	miR control		[46]
lincRNA-p21	Not annotated in human	ALL, CML	Tumor suppressor	Activated by p53 binds hnRNP K to induce apoptosis	Not known	[158]
TCL-6	14q32.13	T cell leukemia	Poorly characterized	Not described	Translocation and inversions with TCR	[151]
WT1-AS	11p13	AML, ALL	Poorly characterized	WT-1 control	Not known	[159]
CRNDE	16q12.2	AML, MM, T-cell leukemia	Oncogene	PRC2 and COREST binding	Not known	[160]
RMRP	9p13.3	Non-Hodgkin lymphoma	Poorly characterized	Not described	Mutation	[161]
SNHG5	6q14.3	B-cell Lymphoma	Poorly characterized	snoRNA host	Translocation (1;3)(q25;q27)	[162]
HOXA-AS2	7p15.2	APL	Poorly characterized		Not known	[163]
HOTAIRM1	7p15.2	Hematopoietic regulator	Regulator of myelopoiesis	HOX A genes.		[164]
EGOT	3p26.1	Hematopoietic regulator	Regulator of eosinophil development			[165]
PU.1-AS	11p11.2 Non annotated	Hematopoietic regulator	*PU.1-AS* regulate the hematopoiesis regulator PU.1	PU.1 control		[166]
EPS	Mouse 4qC7	Hematopoietic regulator	Regulator of erytropoyesis	Pycard repression		[167]
ThyncR1	1q23.1	Hematopoietic regulator	Regulator of T cell selection and maduration.	Riboregulator		[168]

Figure 2. Schematic representation of the function of lncRNAs deregulated in haematological malignancies. (**A**) *BIC*. Myb transcription factor increases the expression of *BIC* in several leukemias and lymphomas. This results in increased levels of *miR-155* and *miR-155*-mediated downregulation of several tumor suppressor genes; (**B**) *ANRIL*. The INK4 *p15INK4b-p14ARF-p16INK4a* cluster transcribes for an antisense transcript named *ANRIL*; PcG complex (PRC2) is targeted to the INK4 locus by *ANRIL*, and locus expression is inhibited; (**C**) *MEG3*. *MEG3*, among other functions, stimulates p53-dependent tumor suppressor pathways by several mechanisms. *MEG3* down-regulates MDM2 expression, therefore decreasing p53 MDM2-mediated degradation. *MEG3* increases p53 protein levels and stimulates p53-dependent transcription. *MEG3* enhances p53 binding to some target promoters such as *GDF15*; (**D**) *GAS5*. *GAS5* binds the DNA binding domain of glucocorticoid receptors (GR) and impedes GR binding to DNA and induction of GR-dependent genes such as *cIAP2*.



3.1.3. *PVT1*

It is not clear whether the role of Plasmacytoma variant translocation 1 (*PVT1*) lncRNA in haematological malignancies depends exclusively on being a miRNA host gene. The *PVT1* gene is transcribed to several mature RNAs by alternative splicing, including a cluster of seven miRNAs, six of them annotated in the miRBase as *miR-1204*, *miR-1205*, *miR-1206*, *miR-1207-5p*, *miR-1207-3p*, and *miR-1208*. The function of these miRNAs is unknown with the exception of *miR-1204*. *miR-1204* has been involved in different roles related to development, differentiation and senescence [146,169]. On one hand *miR-1204* has been described as increasing p53 levels and causing cell death [148]. In fact *PVT1* expression is induced in response to p53 [148]. On the other hand, *miR-1204* has been shown to activate Myc and cell proliferation in mouse pre- B cell lines [146,147].

PVT1 is located in chromosome region 8q24.21, relatively close to the transcription factor c-Myc. Translocations within c-Myc or *PVT1*, which cause the overexpression of these two oncogenes compared to healthy cells, are characteristics associated with B cell malignancies including Burkitt Lymphoma (BL), AIDs, Non-Hodgkin lymphoma, mouse plasmacytoma (Pct) and multiple myeloma (MM) [147]. Furthermore, *PVT1* is in a susceptibility locus for classical Hodgkin's lymphoma [145] and a SNP that causes increased *PVT1* expression is associated with prostate cancer risk [170]. *PVT1* is overexpressed, compared to healthy tissues, in breast and ovarian cancer, pediatric malignant astrocytomas, AML and Hodgkin lymphoma [171], suggesting that *PVT1* could be an oncogene. In fact, upregulation of *PVT1* contributes to tumor survival and chemoresistance [171–174] while its downregulation

inhibits cell proliferation and induces a strong apoptotic response [171]. It has been proposed that *PVT1* regulates c-Myc expression but also that *PVT1* is regulated by c-Myc [175]. However, some authors suggest that Myc and *PVT1* contribute to cancer by different mechanisms [147,171]. Further studies are required to understand the role of *PVT1* in tumorigenesis and to determine whether the miRNAs encoded by *PVT1* mediate its functionality.

3.2. LncRNAs with Oncogenic Properties

ANRIL or *CDKN2B-AS1*

Antisense Non-coding RNA in the INK4 Locus (*ANRIL*) or *CDKN2B-AS1* is transcribed antisense to the *p15INK4b-p14ARF-p16INK4a* cluster, whose members are key effectors of oncogene-induced senescence (Figure 2B). The INK4 proteins are induced during aging and in premalignant lesions, limiting tumor progression. Therefore, expression of the *INK4b-ARF-INK4a* locus is tightly controlled and the Polycomb group (PcG) complexes are required to initiate and maintain silencing of this locus [176,177]. PcG complexes are targeted to the locus by *ANRIL* [178]. Depletion of *ANRIL* disrupts binding of the PRC2 component SUZ12 to the locus, increases the expression of p15INK4b and inhibits cellular proliferation. *ANRIL*, as a pol II nascent transcript, also controls cellular lifespan by targeting the PRC1 component CBX7 to the INK4 locus [27].

Genome-wide association studies revealed that *ANRIL* is located in a genetic susceptibility locus (9p21) associated with several diseases, including coronary artery disease (CAD), atherosclerosis, intracranial aneurysm, type 2 diabetes, and several cancers, such as glioma, basal cell carcinoma, nasopharyngeal carcinoma, and breast cancer [179]. Several single nucleotide polymorphisms (SNP) in this locus alter *ANRIL* structure [180] and *ANRIL* gene expression [181,182], mediating susceptibility to disease. There is a statistically significant association between an *ANRIL* polymorphism and Philadelphia positive Acute Lymphoblastic Leukemia (Ph+ ALL) [183]. Furthermore, 69% of samples (n = 16) from patients with ALL and AML showed relatively increased expression of *ANRIL* and downregulated p15 compared to controls [130]. The expression of *ANRIL*, CBX7, and EZH2 is coordinated and elevated in preneoplastic and neoplastic tissues, leading to decreased p16INK4a expression and decreased senescence [27]. In fact, the *INK4b-ARF-INK4a* locus is subject to frequent deletion or hypermethylation in cancers, including leukemia, melanoma, lung and bladder cancers [177].

3.3. LncRNAs with Tumor Suppressor Properties

3.3.1. *MEG3*

The maternally expressed gene 3 (*MEG3*) was the first lncRNA proposed to function as a tumor suppressor (Figure 2C). *MEG3* is a paternally imprinted polyadenylated RNA, expressed in many normal human tissues as several alternative splicing variants [184,185]. *MEG3* expression was decreased compared to healthy tissues in various brain cancers (pituitary adenomas, glioma and the majority of meningiomas and meningioma cell lines) [149,154], bladder, lung, renal, breast, cervix, colon and prostate cancers and haematological malignancies such as MM, AML or myelodysplastic

syndromes. Surprisingly *MEG3* is overexpressed in Wilms tumor and may be increased or decreased in different hepatocellular carcinomas *versus* healthy livers [186].

The last intron of *MEG3* lncRNA encodes the evolutionarily conserved *miR-770* [187] and *MEG3* isoforms can contain several small open reading frames that are not required for *MEG3* function [152,153]. Instead, the *MEG3* secondary structure, rather than primary sequence, is critical to maintaining function [152]. *MEG3* lncRNA localizes to the nucleus, although some cytoplasmic *MEG3* transcripts have been detected [184,188,189]. In the nucleus, *MEG3* binds to PRC2 to control the imprinting of the *DLK1* locus, where *MEG3* belongs. Furthermore, *MEG3* stimulates both p53-dependent and p53-independent tumor suppressor pathways [149,150,152–155]. *MEG3* activates the tumor suppressor protein p53 at different levels. On one hand *MEG3* down-regulates MDM2 expression, therefore decreasing p53 MDM2-mediated degradation [150]. On the other hand, *MEG3* significantly increases p53 protein levels and stimulates p53-dependent transcription [155]. Finally, *MEG3* enhances p53 binding to some target promoters such as *GDF15* [152,153]. Ectopic expression of *MEG3* RNA leads to p53 accumulation and inhibition of cellular proliferation [153,185]. Inactivation of *MEG3* in the brain increases the expression of genes involved in angiogenesis, suggesting that the tumour suppressor function of *MEG3* works, in part, by inhibiting angiogenesis [190]. In bladder cancer a negative correlation has been shown between *MEG3* expression and autophagy [191].

3.3.2. *DLEU1* and *DLEU2*

Deleted in leukemia 1 (*DLEU1*) and 2 (*DLEU2*) are two genes transcribed head to head in a 30-kb region located in the long arm of chrormosome 13 (13q14), which is lost in more than 50% of patients with CLL and that predicts a poor prognosis [192]. The homozygous loss of this region has great effects on the regulation and control of normal CD5+ B lymphocytes and their homeostasis. Recent studies show that *DLEU1* and *DLEU2* control transcription of their neighbouring candidate tumour suppressor genes, which may act as positive regulators of NF-kB activity [156]. As binding of *DLEU1* and *DLEU2* to chromatin has not been detected, it has been proposed that they regulate neighbouring gene expression by divergent transcription. In addition, the intron 4 of *DLEU2* encodes the miRNAs *hsa-miR-16-1* and *hsa-miR-15a.* This miRNA cluster exerts a crucial role in the tumorigenesis of CLL, in part, regulating the oncogene *BCL2* [193]. Knocking out *hsa-miR-16-1* and *hsa-miR-15a* in mice leads to a lymphoproliferative disease [194]. However the knockout model of *DLEU2*, which includes deletion of *hsa-miR-16-1* and *hsa-miR-15a* as well, shows a more aggressive phenotype than the *hsa-miR-16-1*/*hsa-miR-15a 6* knockout model alone, suggesting that *DLEU2* can participate in CLL development on its own. In fact, increased expression of *DLEU2* leads to reduced proliferation and clonogenicity [195].

3.3.3. *GAS5*

Growth arrest specific 5 (*GAS5*) is induced under starvation conditions and is highly expressed in cells that have arrested growth [196,197]. *GAS5* modulates cell survival and metabolism by antagonizing the glucocorticoid receptor (GR) [87] (Figure 2D). *GAS5* binds the DNA binding domain of GRs directly, preventing GRs from binding to DNA, from functioning as transcription activators and from reducing cell metabolism [87]. *GAS5* could regulate other receptors (androgen, mineralocorticoid

and progesterone but not estrogen receptors) by the same means [87]. Expression of *GAS5* is sufficient to repress GR-induced genes, such as the cellular inhibitor of apoptosis 2 (*cIAP2*) and sensitizes cells to apoptosis [87]. Thus, *GAS5* behaves as a tumor suppressor. *GAS5* expression is decreased in breast cancer and is almost undetectable in growing leukemia cells and increases after density-induced cell cycle arrest [87,196,197]. At the same time, *GAS5* has been shown to be regulated by the mammalian target of rapamycin (mTOR) pathway and to mediate the effect of rapamycin on the cell cycle in T cells [198]. Downregulation of *GAS5* by RNA interference protects leukemic and primary human T cells from the anti-proliferative effect of rapamycin [199].

3.4. LncRNAs with Dual Functions

3.4.1. *H19*

H19 is an imprinted lncRNA located close to the *IGF2* gene. *H19* is expressed form the maternal allele and *IGF2* from the paternal allele [59,200]. A key feature of cancer is the loss of this imprinting, which results in the well documented overexpression of *H19* in cancers of the colon, liver, breast and bladder and in hepatic metastases, compared to healthy tissues [200–204]. Loss of *H19* imprinting has been described in adult T-cell leukaemia/lymphoma (ATL) [157] and decreased *H19* expression was found in the bone marrow of patients with clinically untreated chronic myeloproliferative disorders, including chronic myeloid leukemia (CML), polycythemia vera (PV), essential thrombocythemia (ET), primary myelofibrosis (PMF) and chronic myelomonocytic leukaemia (CMML) [205,206] and AML [207].

H19 can behave as an oncogene or as a tumour suppressor [59]. *H19* expression can be activated by the oncogene c-Myc [200] and downregulated by the tumour suppressor p53 [208,209]. Downregulation of *H19* by RNAi blocks cell growth and clonogenicity of lung cancer cell lines [200] and decreases xenograft tumour growth of a hepatocellular carcinoma cell line [203]. Furthermore, *H19* is the precursor of miR-675, which downregulates the tumor suppressor retinoblastoma in human colorectal cancer [210]. All these results indicate that *H19* is an oncogene [210]. However, depletion of *H19* caused increased polyp count in a mouse model for colorectal cancer [211], larger tumor growth in a mouse teratocarcinoma model and an earlier development of tumours in a mouse hepatocarcinoma model [212]. This dual role as oncogene or tumour suppressor may depend on the cellular environment of the tumour type.

3.4.2. T-UCRs

The expression of many T-UCRs has been described to be significantly altered in tumours such as CLL, colorectal and hepatocellular carcinomas and neuroblastomas [46,162,213,214]. Certain SNPs in T-UCR genes were associated with increased familial breast cancer risk [163]. Moreover, T-UCR transcription profiles can be used to differentiate types of human cancers and predict patient outcome [213]. Some T-UCRs seem tumour specific, such as *UC.73A* and *UC.338*, which are decreased in colon cancer [215]. In fact, some T-UCRs differentially expressed in a particular human cancer locate in fragile sites or cancer-associated genomic regions specifically associated with that type of cancer [216]. This is the case of *UC.349A* and *UC.352*, differentially expressed between

normal and leukemic CD5-positive cells [46] and located within a chromosomal region linked to susceptibility to familial CLL [217]. Moreover, a profile of 19 T-UCRs (8 up- and 11 down-regulated) was able to differentiate between normal, CLL, colorectal, and hepatocarcinoma samples. Expression of five T-UCRs was able to divide a CLL cohort into two prognostic groups [46]. Expression of these diagnostic T-UCRs negatively correlated with a previously defined CLL miRNA signature, suggesting a mechanism for miRNA regulation of these T-UCRs [218].

3.5. LncRNAs Poorly Characterized in Haematological Malignancies

***LincRNA-p21*:** is a p53 activated lncRNA identified in mouse that binds to and guides hnRNP K to target genes. *LincRNA-p21* bound hnRNP K acts as a transcriptional repressor that leads to the induction of apoptosis [23]. As BCR-ABL1 stimulates hnRNP-K expression and stability and promotes tumor progression, it has been suggested that *lincRNA-p21* could play a relevant role in acute or chronic leukemia [219,220]. Furthermore, *lincRNA-p21* can inhibit the translation of target mRNAs [104]. In the absence of HuR, *lincRNA-p21* is stable and interacts with the mRNAs *CTNNB1*, *JUNB* and translational repressor Rck, repressing the translation of the targeted mRNAs [104].

TCL6: T cell Leukemia/Lymphoma 6 (*TCL6*) is transcribed from a locus involved in translocations and inversions with T cell receptor (*TCR*) [221]. These rearrangements in TCR commonly lead to activation of *TCL6* lncRNA and other oncogenes related to T cell leukemogenesis [151].

***WT1-AS*:** is an antisense lncRNA to WT-1, a well-characterized developmental gene that is mutated in Wilms' tumor (WT) and AML. *WT1-AS* has been shown to regulate WT1 protein levels. *WT1-AS* binds the exon 1 of WT1 mRNA in the cytoplasm. It has been suggested that the abnormal splicing of *WT1-AS* in AML could play a role in the development of this malignancy [159].

***CRNDE*:** is overexpressed, compared to healthy tissue, in more than 90% of colorectal adenomas tested, but also in hepatocellular, prostate, brain, kidney and pancreas carcinomas and different haematological neoplasia such as AML, MM and T cell leukemia [160]. *CRNDE* has been described as downregulated in ovarian cancer and tends to be overexpressed in non-differentiated tissues *versus* differentiated controls [160]. *CRNDE* binds PRC2 and the downregulation of *CRNDE* causes upregulation of PRC2 regulated genes, decreases growth and increases apoptosis [66].

RMRP: Ribonuclease mitochondrial RNA processing (*RMRP)* is a lncRNA mutated in Cartilage-Hair Hypoplasia (CHH), an autosomal recessive chondrodysplasia with short stature, which entails a high risk of developing Non-Hodgkin lymphoma disease [161,222].

SNHG5: is a precurssor of snoRNAs, similar to *GAS5*, located at the breakpoint of the chromosomal translocation t(3;6)(q27;q15), involved in diffuse large B-cell lymphoma [223].

***HOXA-AS2: HOXA** Cluster Antisense RNA 2* (*HOXA-AS2*) lncRNA is antisense to *HOX3* and *HOX4* coding genes. In an acute promyelocytic leukemia (APL) cell line, *HOXA-AS2* upregulation correlated with inhibition of apoptosis. Treatment with all-*trans* retinoic acid (ATRA) blocked the expression of *HOXA-AS2* and increased apoptosis of the APL cell line [224].

4. LncRNAs Involved in Hematopoiesis

The best studied lncRNA in hematopoiesis is *HOTAIRM1* (HOX antisense intergenic RNA myeloid 1). *HOTAIRM1* is as an essential regulator of myeloid cell differentiation that locates at the 3'

end of the *HOXA* cluster and controls *HOXA1* expression [164]. HOXA genes are important transcriptional regulators in normal and malignant hematopoiesis and are known to be important for many cancers including leukemias harbouring MLL rearrangements. *HOTAIRM1* is expressed specifically in the myeloid lineage and is induced during the retinoic acid-driven granulocytic differentiation of the NB4 promyelocytic leukaemia cell line and normal human hematopoietic cells. Knockdown of *HOTAIRM1* affects retinoic acid-induced expression of *HOXA1* and *HOXA4* (but not distal *HOXA* genes) and attenuates induction of myeloid differentiation genes [164].

Other lncRNAs involved in hematopoiesis have also been described. *EGO* (or *EGOT* in human) lncRNA was identified in mouse eosinophil differentiation of CD34+HSCs where it stimulated major basic protein and eosinophil-derived neurotoxin mRNA expression [165]. The lncRNA *PU.1-AS* is an antisense transcript of *PU.1* that negatively regulates *PU.1* mRNA translation by a mechanism similar to miRNAs [166]. *PU.1* is a master hematopoietic transcriptional regulator essential for normal hematopoietic development and suppression of leukaemia development. LincRNA erythroid prosurvival (*EPS*) is one of the about 400 lncRNAs whose expression is modulated during red blood cell formation and is required for differentiation during hematopoiesis in mouse [164,165,167]. *EPS* is an erythroid-specific lncRNA that represses expression of *PYCARD*, a proapoptotic gene, and therefore inhibits apoptosis [167,225]. EPS is not well conserved among mammals. It is presently unclear whether a human version of EPS exists. Finally, *THY-ncR1* is a thymus-specific lncRNA expressed in cell lines derived from stage III immature T cells in which the neighbouring *CD1* gene cluster is also specifically activated [168].

5. Regulation of the Expression of lncRNAs Involved in Haematological Malignancies

Altered expression of lncRNAs, similar to that of coding genes, can be the result of genomic alterations, epigenetic regulation or a change in response to transcription factors or stability effectors such as miRNAs.

The presence of mutations in the lncRNA primary sequence correlates highly with human diseases. In fact, most mutations in the genome occur in noncoding regions [226]. Mutations can be large or small. Large-scale mutations are deletions and amplifications of hundreds of nucleotides and chromosomal translocations occurring at fragile sites. Genome-wide analyses looking for fragile sites in lncRNA genes have not yet been performed. However, it is expected that lncRNAs will have a clear association with common chromosomal aberrations similar to that found for miRNAs in human haematological malignancies and carcinomas [46]. In fact, several studies have described lncRNAs affected by large scale mutations. One of the best examples is *ANRIL*, affected by a large germline deletion that includes the complete INK4/ARF locus. This deletion is associated with hereditary cutaneous malignant melanoma and neural system tumors syndrome [179]. *DLEU1* and *DLEU2* lncRNAs also locate in a region commonly deleted in CLL (see above).

Small scale mutations are deletions or insertions of a few nucleotides. The relevance of small scale mutations for lncRNAs is obscured by the fact that little is known about the relevance of the primary sequence in lncRNA functionality and expression. It is expected that small mutations can lead to disease if they affect relevant linear sequences or they alter the structure of domains important in lncRNA functionality or accumulation. In fact, several disease-associated SNPs have been described as

affecting the structure of the 5' and 3' non-translated regions of coding genes [226]. Furthermore, GWAS studies have shown that SNPs in noncoding regions are associated with higher susceptibility to diverse diseases. Germline and somatic mutations in lncRNA genes have been identified in haematological malignancies and colorectal cancers [227]. SNPs that may affect *ANRIL* have been associated with increased risk of type 2 diabetes and increased susceptibility to coronary artery disease and atherosclerosis [228,229]. Some of these mutations did not affect *ANRIL* transcription or stability. Instead, they disrupt *ANRIL* splicing, resulting in a circular transcript, affecting normal *ANRIL* function and influencing *INK4/ARF* locus expression [180]. Moreover, genetic aberrations of the *GAS5* locus have been found in melanoma, breast and prostate cancers [230–232].

Several lncRNAs are regulated at the transcriptional level. Thus, lncRNAs, such as *lincRNA-P21*, are activated in response to DNA damage by the direct binding of the tumour-suppressor protein p53 to the promoter [23]. Similarly, the expression of several lincRNAs responds to pluripotency factors or oncogenes.

Epigenetic modifications are key regulators of lncRNA expression. This has been well described for *MEG3* and *DLEU1/DLEU2*. Expression of the *MEG3* locus is regulated by two regions, which are hypermethylated in several solid tumours leading to downregulation of *MEG3* expression [185,233,234]. AML patients with aberrant hypermethylation of the *MEG3* promoter showed decreased overall survival [235,236]. Thus, *MEG3* methylation status may serve as a useful biomarker in this leukemia. A similar *MEG3* hypermethylation was observed in 35% of the patients with myelodysplastic syndrome, but in this case there was no statistically significant correlation between *MEG3* hypermethylation and prognosis [235]. Similarly, conserved CpG islands at the transcriptional start sites of *DLEU1* and *DLEU2* were found to be significantly demethylated in a cohort of 143 patients with CLL [156]. Demethylation correlated with transcriptional deregulation of the neighbouring candidate tumour suppressor genes. T-UCRs expression has also been shown to be repressed by CpG island hypermethylation [47,213].

Finally, the expression of lncRNAs can be regulated by miRNAs. Several miRNAs have been described as regulating T-URC expression. This has been best described for *miR-155*, which is overexpressed in CLL compared to healthy cells. *miR-155* targets T-UCRs both *in vitro* and in CLL patient samples [46]. Interestingly, *miR-29a* has also been shown to regulate *MEG3* expression in hepatocarcinoma cell lines [186].

6. Concluding Remarks

The identification of lncRNAs and the functional relevance of the lncRNAs studied so far has changed the view about genomes, transcriptomes and gene expression regulation. As the lncRNA field is in its infancy, surprising results are still expected, but a tremendous amount of work needs to be done. Firstly, a systematic identification and annotation of lncRNAs and their expression patterns should be performed and made publically available. As most lncRNAs are tissue specific, all tissues should be profiled. Also, as there is poor sequence conservation between lncRNAs of different species, efforts should be devoted to describing a collection of lncRNAs in different species, including human, mouse, rat, zebra fish, fly, *Arabidopsis* and yeast. As the regulation of expression of lncRNAs is tightly controlled, lncRNAs should also be described in cells responding to different stimuli and in diseased cells. These studies will be complicated further by the fact that lncRNA genes may be transcribed to

different transcripts by alternative splicing, polyadenylation and the use of different promoters. It is also necessary to develop a new universal nomenclature that would facilitate routine work with these non coding RNA molecules.

Secondly, functional studies should be performed. Gain and loss of function studies could be carried out to analyze the impact of the lncRNA on the cell phenotype. Transcriptome analysis coupled with gain and loss of function studies could provide clues regarding the cellular pathways affected by the lncRNA, especially if the lncRNA of interest is a regulator of the expression of specific genes. Analysis of lncRNA subcellular localization can also give clues to lncRNA functionality. This can be done with Fish-like techniques that use several labelled oligos at a time. This is essential to detect the expression of lncRNAs, which are generally very structured and not very abundant. The functional domains of lncRNAs should be identified and it should be ruled out that lncRNAs function through the translation of short peptides. Furthermore, it would be desirable to determine the structure of key domains in lncRNAs similarly to what has been done with proteins. This is a major task as there are no reliable methods to determine the secondary structures of lncRNAs with bioinformatic tools. Chemical probing and point mutation studies have been used to determine the structure of many RNAs, but these techniques are very time consuming. Faster results could be obtained by parallel analysis of RNA structure (PARS-Seq) or Frag-Seq, which uses deep sequencing of RNA fragments obtained from RNAs treated with specific RNases that cleave RNA at highly selective structural positions [237]. Furthermore, it would be interesting to identify the factors that bind to relevant lncRNAs. Ideally, specific RNAs should be immunoprecipitated and subjected to mass spectrometry to identify RNA binding proteins. This is not easy, but has been done successfully with pools of cellular RNAs purified by binding to oligodT beads [238]. Theoretically, a lncRNA of interest could be labelled with a domain targeted by a specific protein and the complex could be purified with antibodies specific to the protein. Alternatively, the lncRNA could be immunoprecipitated from cell extracts using biotinylated tiling oligos and streptavidin. Then, lncRNA bound DNA or RNA can be sequenced from the immunoprecipitates. When looking for DNA interactors, this technique has been named Chromatin Isolation by RNA Purification (ChIRP) and has allowed the identification of the natural regions of chromatin that interact with a given lncRNA [58]. Finally, the lncRNA can be transcribed and labelled *in vitro*, incubated with cell extracts and immunoprecipitated with label binding factors.

Finally, detailed analysis of functional lncRNAs will most probably reveal interesting cellular pathways and help to design the architecture of biological tools that may be of interest for biotechnological development. Domains of lncRNAs that function as decoys for miRNAs or transcription factors, mimicking *GAS5* function [239], could be expressed to obtain therapeutic effects. Several lncRNA domains with a specific tertiary structure and a given function could probably be combined to generate lncRNAs with novel functions that could be of therapeutic interest. For instance, an RNA domain involved in the binding to a specific region of the chromatin could be fused to an RNA domain that interacts with factors that silence or activate gene expression or that induce chromosome bendings or genome reorganizations at the specific position. This could be used for silencing of oncogenes or reactivation of tumour suppressor genes. Thus, analysis of the function of lncRNAs is expected to have a tremendous impact on the management of human disease.

Furthermore, strong associations between some lncRNAs and some human diseases have been described. The number of lncRNAs relevant to human diseases is expected to increase as a result of the

systematic identification of lncRNAs whose expression is altered in healthy and diseased cells and by genome-wide association studies. In fact GWAS analysis has identified *ANRIL* as a lncRNA involved in atherosclerosis, coronary artery disease, and type 2 diabetes [179]. In the case of cancer and specifically of haematological malignancies, GWAS results at lncRNA loci may identify patient populations at risk of cancer, may classify patients into aggressive or mild cancer groups and may predict a patient's response to a given therapy [240,241]. Once lncRNAs related to a disease are described, the issue should be addressed whether they are useful signatures for early disease detection, for prognosis or can be used as candidate drug targets for disease intervention [242].

lncRNAs may have specific advantages when used as diagnostic biomarkers, as some show tissue-specific and cancer-specific expression patterns [243]. This is the case of *HULC*, a liver-specific lncRNA highly expressed in primary liver tumours and hepatic metastases of colorectal carcinoma, but not found in primary colon cancers or in non-liver metastases [244,245]. Thus, the expression of *HULC* and other lncRNAs can be used to differentiate between subtypes of the same cancer or to identify unknown primary tumours. Similarly, *PCGEM1*, *PCA3* or *PRNCR1* are three lncRNAs exclusively associated with prostate cancer [123,246,247]. Also, as in the case with miRNAs, some lncRNAs can be detected in body fluids by quantitative reverse transcriptase polymerase chain reaction and therefore enable non-invasive diagnoses. In fact, *HULC* can be detected in the blood of hepatocellular carcinoma patients using qRT-PCR [245]. The ProgensaTM PCA3 urine test, a kit to detect *PCA3* in urine samples from patients with prostate cancer is already being clinically used [248,249]. This specific test can help patients who had a first negative prostate biopsy to avoid unnecessary repeated biopsies [250]. In spite of this fast clinical translation for *PCA3* analysis in prostate cancer, the biological function of *PCA3* is unknown.

LncRNAs can also be used as predictive markers, as lncRNA expression can correlate with patient outcome or response to chemotherapy. Thus, the expression of *HOTAIR* correlates with metastasis and poor outcome in primary breast tumors, gastrointestinal, hepatocellular and colorectal cancers and the expression of *MALAT* correlates with survival in early-stage lung adenocarcinoma [68,124,251–253]. Also, the expression of *XIST* correlates with disease-free survival of Taxol-treated cancer patients [254].

Finally, lncRNAs could be used therapeutically. In cancer, expression of tumour suppressor lncRNAs, such as *GAS5* or *MEG3*, should decrease tumour growth. When the downregulation of tumour suppressor lncRNAs results from aberrant epigenetic mechanisms such as DNA hypermethylation or loss of histone acetylation, demethylating agents or histone deacetylase inhibitors could help to reestablish expression. Otherwise, expression of lncRNAs may require gene therapy delivery systems with viral vectors, which are not efficient in targeting all cells within a tumour. Furthermore, RNA interference can be used to decrease the expression of lncRNAs with oncogenic properties. While many lncRNAs have been silenced using siRNAs, it is generally believed that the secondary structure of lncRNAs hinders siRNA functionality. Instead, expression of lncRNAs with oncogenic or tumour suppressor molecules could be altered with small molecules that affect their promoters. Small molecules, aptamers or stable antisense oligonucleotides could also be identified that target essential structures for oncogenic lncRNA functionality. Thus, preventing the interactions of *HOTAIR* with PRC2, for example, may limit the metastatic potential of breast cancer cells [255]. Even if all these strategies are possible, much investment in this field will be required to transfer lncRNA research to clinical oncology.

Acknowledgments

Supported in part by grants from Ministerio de Ciencia e Innovacion and Instituto de Salud Carlos III (ISCIII) BIO2009/09295, SAF2012-40003, PI10/01691, PI10/02983, RD12/0036/0068, CP07/00215, FEDER funding, FP7 Marie Curie Program PIOF-GA-2012-330598 LincMHeM, funds from the "UTE project CIMA" and by the project RNAREG [CSD2009-00080], funded by the Ministry of Science and Innovation under the programme CONSOLIDER INGENIO 2010.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Mattick, J.S. Non-coding RNAs: The architects of eukaryotic complexity. *EMBO Rep*. **2001**, *2*, 986–991.
2. Dunham, I.; Kundaje, A.; Aldred, S.F.; Collins, P.J.; Davis, C.A.; Doyle, F.; Epstein, C.B.; Frietze, S.; Harrow, J.; Kaul, R.; *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74.
3. Lin, M.F.; Carlson, J.W.; Crosby, M.A.; Matthews, B.B.; Yu, C.; Park, S.; Wan, K.H.; Schroeder, A.J.; Gramates, L.S.; St Pierre, S.E.; *et al.* Revisiting the protein-coding gene catalog of *Drosophila melanogaster* using 12 fly genomes. *Genome Res*. **2007**, *17*, 1823–1836.
4. Lin, M.F.; Deoras, A.N.; Rasmussen, M.D.; Kellis, M. Performance and scalability of discriminative metrics for comparative gene identification in 12 *Drosophila genomes. PLoS Comput. Biol*. **2008**, *4*, e1000067.
5. Pueyo, J.I.; Couso, J.P. Tarsal-less peptides control Notch signalling through the Shavenbaby transcription factor. *Dev. Biol*. **2011**, *355*, 183–193.
6. Ingolia, N.T.; Lareau, L.F.; Weissman, J.S. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **2011**, *147*, 789–802.
7. Banfai, B.; Jia, H.; Khatun, J.; Wood, E.; Risk, B.; Gundling, W.E., Jr.; Kundaje, A.; Gunawardena, H.P.; Yu, Y.; Xie, L.; *et al.* Long noncoding RNAs are rarely translated in two human cell lines. *Genome Res*. **2012**, *22*, 1646–1657.
8. Derrien, T.; Johnson, R.; Bussotti, G.; Tanzer, A.; Djebali, S.; Tilgner, H.; Guernec, G.; Martin, D.; Merkel, A.; Knowles, D.G.; *et al.* The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res*. **2012**, *22*, 1775–1789.
9. Kloc, M.; Wilk, K.; Vargas, D.; Shirato, Y.; Bilinski, S.; Etkin, L.D. Potential structural role of non-coding and coding RNAs in the organization of the cytoskeleton at the vegetal cortex of Xenopus oocytes. *Development* **2005**, *132*, 3445–3457.
10. Wadler, C.S.; Vanderpool, C.K. A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 20454–20459.

11. Dinger, M.E.; Pang, K.C.; Mercer, T.R.; Mattick, J.S. Differentiating protein-coding and noncoding RNA: Challenges and ambiguities. *PLoS Comput. Biol*. **2008**, *4*, e1000176.
12. Leygue, E. Steroid receptor RNA activator (SRA1): Unusual bifaceted gene products with suspected relevance to breast cancer. *Nucl. Recept. Signal*. **2007**, *5*, e006.
13. Jenny, A.; Hachet, O.; Zavorszky, P.; Cyrklaff, A.; Weston, M.D.; Johnston, D.S.; Erdelyi, M.; Ephrussi, A. A translation-independent role of oskar RNA in early *Drosophila oogenesis*. *Development* **2006**, *133*, 2827–2833.
14. Candeias, M.M.; Malbert-Colas, L.; Powell, D.J.; Daskalogianni, C.; Maslon, M.M.; Naski, N.; Bourougaa, K.; Calvo, F.; Fahraeus, R. P53 mRNA controls p53 activity by managing Mdm2 functions. *Nat. Cell Biol*. **2008**, *10*, 1098–1105.
15. Guttman, M.; Amit, I.; Garber, M.; French, C.; Lin, M.F.; Feldser, D.; Huarte, M.; Zuk, O.; Carey, B.W.; Cassady, J.P.; *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **2009**, *458*, 223–227.
16. Tilgner, H.; Knowles, D.G.; Johnson, R.; Davis, C.A.; Chakrabortty, S.; Djebali, S.; Curado, J.; Snyder, M.; Gingeras, T.R.; Guigo, R. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res*. **2012**, *22*, 1616–1625.
17. Gibb, E.A.; Vucic, E.A.; Enfield, K.S.; Stewart, G.L.; Lonergan, K.M.; Kennett, J.Y.; Becker-Santos, D.D.; MacAulay, C.E.; Lam, S.; Brown, C.J.; *et al.* Human cancer long non-coding RNA transcriptomes. *PLoS One* **2011**, *6*, e25915.
18. Mercer, T.R.; Dinger, M.E.; Sunkin, S.M.; Mehler, M.F.; Mattick, J.S. Specific expression of long noncoding RNAs in the mouse brain. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 716–721.
19. Djebali, S.; Davis, C.A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F.; *et al.* Landscape of transcription in human cells. *Nature* **2012**, *489*, 101–108.
20. Hauptman, N.; Glavac, D. Long non-coding RNA in cancer. *Int. J. Mol. Sci*. **2013**, *14*, 4655–4669.
21. Agirre, X.; Jimenez-Velasco, A.; San Jose-Eneriz, E.; Garate, L.; Bandres, E.; Cordeu, L.; Aparicio, O.; Saez, B.; Navarro, G.; Vilas-Zornoza, A.; *et al.* Down-regulation of hsa-miR-10a in chronic myeloid leukemia CD34+ cells increases USF2-mediated cell growth. *Mol. Cancer Res*. **2008**, *6*, 1830–1840.
22. Agirre, X.; Vilas-Zornoza, A.; Jimenez-Velasco, A.; Martin-Subero, J.I.; Cordeu, L.; Garate, L.; San Jose-Eneriz, E.; Abizanda, G.; Rodriguez-Otero, P.; Fortes, P.; *et al.* Epigenetic silencing of the tumor suppressor microRNA Hsa-miR-124a regulates CDK6 expression and confers a poor prognosis in acute lymphoblastic leukemia. *Cancer Res*. **2009**, *69*, 4443–4453.
23. Huarte, M.; Guttman, M.; Feldser, D.; Garber, M.; Koziol, M.J.; Kenzelmann-Broz, D.; Khalil, A.M.; Zuk, O.; Amit, I.; Rabani, M.; *et al.* A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **2010**, *142*, 409–419.
24. Katayama, S.; Tomaru, Y.; Kasukawa, T.; Waki, K.; Nakanishi, M.; Nakamura, M.; Nishida, H.; Yap, C.C.; Suzuki, M.; Kawai, J.; *et al.* Antisense transcription in the mammalian transcriptome. *Science* **2005**, *309*, 1564–1566.

25. Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M.C.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C.; *et al.* The transcriptional landscape of the mammalian genome. *Science* **2005**, *309*, 1559–1563.
26. Galante, P.A.; Vidal, D.O.; de Souza, J.E.; Camargo, A.A.; de Souza, S.J. Sense-antisense pairs in mammals: Functional and evolutionary considerations. *Genome Biol*. **2007**, *8*, R40.
27. Yap, K.L.; Li, S.; Munoz-Cabello, A.M.; Raguz, S.; Zeng, L.; Mujtaba, S.; Gil, J.; Walsh, M.J.; Zhou, M.M. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* **2010**, *38*, 662–674.
28. Krystal, G.W.; Armstrong, B.C.; Battey, J.F. *N*-myc mRNA forms an RNA-RNA duplex with endogenous antisense transcripts. *Mol. Cell Biol*. **1990**, *10*, 4180–4191.
29. Hastings, M.L.; Milcarek, C.; Martincic, K.; Peterson, M.L.; Munroe, S.H. Expression of the thyroid hormone receptor gene, erbAalpha, in B lymphocytes: Alternative mRNA processing is independent of differentiation but correlates with antisense RNA levels. *Nucleic Acids Res*. **1997**, *25*, 4296–4300.
30. Yan, M.D.; Hong, C.C.; Lai, G.M.; Cheng, A.L.; Lin, Y.W.; Chuang, S.E. Identification and characterization of a novel gene Saf transcribed from the opposite strand of Fas. *Hum. Mol. Genet*. **2005**, *14*, 1465–1474.
31. Annilo, T.; Kepp, K.; Laan, M. Natural antisense transcript of natriuretic peptide precursor A (NPPA): Structural organization and modulation of NPPA expression. *BMC Mol. Biol*. **2009**, *10*, 81.
32. Allo, M.; Buggiano, V.; Fededa, J.P.; Petrillo, E.; Schor, I.; de la Mata, M.; Agirre, E.; Plass, M.; Eyras, E.; Elela, S.A.; *et al.* Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat. Struct. Mol. Biol*. **2009**, *16*, 717–724.
33. Faghihi, M.A.; Modarresi, F.; Khalil, A.M.; Wood, D.E.; Sahagan, B.G.; Morgan, T.E.; Finch, C.E.; St Laurent, G., 3rd; Kenny, P.J.; Wahlestedt, C. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat. Med*. **2008**, *14*, 723–730.
34. Guil, S.; Esteller, M. cis-acting noncoding RNAs: Friends and foes. *Nat. Struct. Mol. Biol*. **2012**, *19*, 1068–1075.
35. Heintzman, N.D.; Stuart, R.K.; Hon, G.; Fu, Y.; Ching, C.W.; Hawkins, R.D.; Barrera, L.O.; Van Calcar, S.; Qu, C.; Ching, K.A.; *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet*. **2007**, *39*, 311–318.
36. Heintzman, N.D.; Hon, G.C.; Hawkins, R.D.; Kheradpour, P.; Stark, A.; Harp, L.F.; Ye, Z.; Lee, L.K.; Stuart, R.K.; Ching, C.W.; *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **2009**, *459*, 108–112.
37. Visel, A.; Blow, M.J.; Li, Z.; Zhang, T.; Akiyama, J.A.; Holt, A.; Plajzer-Frick, I.; Shoukry, M.; Wright, C.; Chen, F.; *et al.* ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **2009**, *457*, 854–858.
38. Kim, T.K.; Hemberg, M.; Gray, J.M.; Costa, A.M.; Bear, D.M.; Wu, J.; Harmin, D.A.; Laptewicz, M.; Barbara-Haley, K.; Kuersten, S.; *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **2010**, *465*, 182–187.

39. Wang, D.; Garcia-Bassets, I.; Benner, C.; Li, W.; Su, X.; Zhou, Y.; Qiu, J.; Liu, W.; Kaikkonen, M.U.; Ohgi, K.A.; *et al.* Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature* **2011**, *474*, 390–394.
40. De Santa, F.; Barozzi, I.; Mietton, F.; Ghisletti, S.; Polletti, S.; Tusi, B.K.; Muller, H.; Ragoussis, J.; Wei, C.L.; Natoli, G. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol*. **2010**, *8*, e1000384.
41. Orom, U.A.; Derrien, T.; Beringer, M.; Gumireddy, K.; Gardini, A.; Bussotti, G.; Lai, F.; Zytnicki, M.; Notredame, C.; Huang, Q.; *et al.* Long noncoding RNAs with enhancer-like function in human cells. *Cell* **2010**, *143*, 46–58.
42. Melo, C.A.; Drost, J.; Wijchers, P.J.; van de Werken, H.; de Wit, E.; Oude Vrielink, J.A.; Elkon, R.; Melo, S.A.; Leveille, N.; Kalluri, R.; *et al.* eRNAs are required for p53-dependent enhancer activity and gene transcription. *Mol. Cell*. **2013**, *49*, 524–535.
43. Lai, F.; Orom, U.A.; Cesaroni, M.; Beringer, M.; Taatjes, D.J.; Blobel, G.A.; Shiekhattar, R. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* **2013**, *494*, 497–501.
44. Bejerano, G.; Pheasant, M.; Makunin, I.; Stephen, S.; Kent, W.J.; Mattick, J.S.; Haussler, D. Ultraconserved elements in the human genome. *Science* **2004**, *304*, 1321–1325.
45. Ahituv, N.; Zhu, Y.; Visel, A.; Holt, A.; Afzal, V.; Pennacchio, L.A.; Rubin, E.M. Deletion of ultraconserved elements yields viable mice. *PLoS Biol*. **2007**, *5*, e234.
46. Calin, G.A.; Liu, C.G.; Ferracin, M.; Hyslop, T.; Spizzo, R.; Sevignani, C.; Fabbri, M.; Cimmino, A.; Lee, E.J.; Wojcik, S.E.; *et al.* Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. *Cancer Cell*. **2007**, *12*, 215–229.
47. Lujambio, A.; Portela, A.; Liz, J.; Melo, S.A.; Rossi, S.; Spizzo, R.; Croce, C.M.; Calin, G.A.; Esteller, M. CpG island hypermethylation-associated silencing of non-coding RNAs transcribed from ultraconserved regions in human cancer. *Oncogene* **2010**, *29*, 6390–6401.
48. Hirotsune, S.; Yoshida, N.; Chen, A.; Garrett, L.; Sugiyama, F.; Takahashi, S.; Yagami, K.; Wynshaw-Boris, A.; Yoshiki, A. An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. *Nature* **2003**, *423*, 91–96.
49. Hawkins, P.G.; Morris, K.V. Transcriptional regulation of Oct4 by a long non-coding RNA antisense to Oct4-pseudogene 5. *Transcription* **2010**, *1*, 165–175.
50. Harrison, P.M.; Zheng, D.; Zhang, Z.; Carriero, N.; Gerstein, M. Transcribed processed pseudogenes in the human genome: An intermediate form of expressed retrosequence lacking protein-coding ability. *Nucleic Acids Res*. **2005**, *33*, 2374–2383.
51. Pink, R.C.; Wicks, K.; Caley, D.P.; Punch, E.K.; Jacobs, L.; Carter, D.R. Pseudogenes: Pseudo-functional or key regulators in health and disease? *RNA* **2011**, *17*, 792–798.
52. He, L. Posttranscriptional regulation of PTEN dosage by noncoding RNAs. *Sci. Signal*. **2010**, *3*, pe39.
53. Salmena, L.; Poliseno, L.; Tay, Y.; Kats, L.; Pandolfi, P.P. A ceRNA hypothesis: The Rosetta Stone of a hidden RNA language? *Cell* **2011**, *146*, 353–358.
54. Ledford, H. Circular RNAs throw genetics for a loop. *Nature* **2013**, *494*, 415.
55. Kosik, K.S. Molecular biology: Circles reshape the RNA world. *Nature* **2013**, *495*, 322–324.

56. Memczak, S.; Jens, M.; Elefsinioti, A.; Torti, F.; Krueger, J.; Rybak, A.; Maier, L.; Mackowiak, S.D.; Gregersen, L.H.; Munschauer, M.; *et al.* Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **2013**, *495*, 333–338.
57. Hansen, T.B.; Jensen, T.I.; Clausen, B.H.; Bramsen, J.B.; Finsen, B.; Damgaard, C.K.; Kjems, J. Natural RNA circles function as efficient microRNA sponges. *Nature* **2013**, *495*, 384–388.
58. Chu, C.; Qu, K.; Zhong, F.L.; Artandi, S.E.; Chang, H.Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **2011**, *44*, 667–678.
59. Gabory, A.; Jammes, H.; Dandolo, L. The H19 locus: Role of an imprinted non-coding RNA in growth and development. *Bioessays* **2010**, *32*, 473–480.
60. Mancini-Dinardo, D.; Steele, S.J.; Levorse, J.M.; Ingram, R.S.; Tilghman, S.M. Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes Dev*. **2006**, *20*, 1268–1282.
61. Pauler, F.M.; Koerner, M.V.; Barlow, D.P. Silencing by imprinted noncoding RNAs: Is transcription the answer? *Trends Genet*. **2007**, *23*, 284–292.
62. Engstrom, P.G.; Suzuki, H.; Ninomiya, N.; Akalin, A.; Sessa, L.; Lavorgna, G.; Brozzi, A.; Luzi, L.; Tan, S.L.; Yang, L.; *et al.* Complex Loci in human and mouse genomes. *PLoS Genet*. **2006**, *2*, e47.
63. Li, Y.; Sasaki, H. Genomic imprinting in mammals: Its life cycle, molecular mechanisms and reprogramming. *Cell Res*. **2011**, *21*, 466–473.
64. Mohammad, F.; Mondal, T.; Kanduri, C. Epigenetics of imprinted long noncoding RNAs. *Epigenetics* **2009**, *4*, 277–286.
65. Nagano, T.; Mitchell, J.A.; Sanz, L.A.; Pauler, F.M.; Ferguson-Smith, A.C.; Feil, R.; Fraser, P. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **2008**, *322*, 1717–1720.
66. Khalil, A.M.; Guttman, M.; Huarte, M.; Garber, M.; Raj, A.; Rivea Morales, D.; Thomas, K.; Presser, A.; Bernstein, B.E.; van Oudenaarden, A.; *et al.* Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11667–11672.
67. Rinn, J.L.; Kertesz, M.; Wang, J.K.; Squazzo, S.L.; Xu, X.; Brugmann, S.A.; Goodnough, L.H.; Helms, J.A.; Farnham, P.J.; Segal, E.; *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **2007**, *129*, 1311–1323.
68. Gupta, R.A.; Shah, N.; Wang, K.C.; Kim, J.; Horlings, H.M.; Wong, D.J.; Tsai, M.C.; Hung, T.; Argani, P.; Rinn, J.L.; *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **2010**, *464*, 1071–1076.
69. Tsai, M.C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J.K.; Lan, F.; Shi, Y.; Segal, E.; and Chang, H.Y. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **2010**, *329*, 689–693.
70. Schmitz, K.M.; Mayer, C.; Postepska, A.; Grummt, I. Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev*. **2010**, *24*, 2264–2269.
71. Buske, F.A.; Mattick, J.S.; Bailey, T.L. Potential *in vivo* roles of nucleic acid triple-helices. *RNA Biol*. **2011**, *8*, 427–439.

72. Aguilera, A.; Garcia-Muse, T. R loops: From transcription byproducts to threats to genome stability. *Mol. Cell* **2012**, *46*, 115–124.
73. Jeon, Y.; Lee, J.T. YY1 tethers Xist RNA to the inactive X nucleation center. *Cell* **2011**, *146*, 119–133.
74. Hung, T.; Chang, H.Y. Long noncoding RNA in genome regulation: Prospects and mechanisms. *RNA Biol*. **2010**, *7*, 582–585.
75. Bonasio, R.; Tu, S.; Reinberg, D. Molecular signals of epigenetic states. *Science* **2010**, *330*, 612–616.
76. Sun, Q.; Csorba, T.; Skourti-Stathaki, K.; Proudfoot, N.J.; Dean, C. R-loop stabilization represses antisense transcription at the *Arabidopsis* FLC locus. *Science* **2013**, *340*, 619–621.
77. Bertani, S.; Sauer, S.; Bolotin, E.; Sauer, F. The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. *Mol. Cell* **2011**, *43*, 1040–1046.
78. Dinger, M.E.; Amaral, P.P.; Mercer, T.R.; Pang, K.C.; Bruce, S.J.; Gardiner, B.B.; Askarian-Amiri, M.E.; Ru, K.; Solda, G.; Simons, C.; *et al.* Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res*. **2008**, *18*, 1433–1445.
79. Pandey, R.R.; Mondal, T.; Mohammad, F.; Enroth, S.; Redrup, L.; Komorowski, J.; Nagano, T.; Mancini-Dinardo, D.; Kanduri, C. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell* **2008**, *32*, 232–246.
80. Yao, H.; Brick, K.; Evrard, Y.; Xiao, T.; Camerini-Otero, R.D.; Felsenfeld, G. Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. *Genes Dev*. **2010**, *24*, 2543–2555.
81. Wang, K.C.; Yang, Y.W.; Liu, B.; Sanyal, A.; Corces-Zimmerman, R.; Chen, Y.; Lajoie, B.R.; Protacio, A.; Flynn, R.A.; Gupta, R.A.; *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **2011**, *472*, 120–124.
82. Feng, J.; Bi, C.; Clark, B.S.; Mady, R.; Shah, P.; Kohtz, J.D. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev*. **2006**, *20*, 1470–1484.
83. Lanz, R.B.; McKenna, N.J.; Onate, S.A.; Albrecht, U.; Wong, J.; Tsai, S.Y.; Tsai, M.J.; O'Malley, B.W. A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell* **1999**, *97*, 17–27.
84. Caretti, G.; Schiltz, R.L.; Dilworth, F.J.; di Padova, M.; Zhao, P.; Ogryzko, V.; Fuller-Pace, F.V.; Hoffman, E.P.; Tapscott, S.J.; Sartorelli, V. The RNA helicases p68/p72 and the noncoding RNA SRA are coregulators of MyoD and skeletal muscle differentiation. *Dev. Cell*. **2006**, *11*, 547–560.
85. Watanabe, M.; Yanagisawa, J.; Kitagawa, H.; Takeyama, K.; Ogawa, S.; Arao, Y.; Suzawa, M.; Kobayashi, Y.; Yano, T.; Yoshikawa, H.; *et al.* A subfamily of RNA-binding DEAD-box proteins acts as an estrogen receptor alpha coactivator through the N-terminal activation domain (AF-1) with an RNA coactivator, SRA. *EMBO J*. **2001**, *20*, 1341–1352.
86. Lanz, R.B.; Chua, S.S.; Barron, N.; Soder, B.M.; DeMayo, F.; O'Malley, B.W. Steroid receptor RNA activator stimulates proliferation as well as apoptosis *in vivo*. *Mol. Cell Biol*. **2003**, *23*, 7163–7176.

87. Kino, T.; Hurt, D.E.; Ichijo, T.; Nader, N.; Chrousos, G.P. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci. Signal*. **2010**, *3*, ra8.
88. Willingham, A.T.; Orth, A.P.; Batalov, S.; Peters, E.C.; Wen, B.G.; Aza-Blanc, P.; Hogenesch, J.B.; Schultz, P.G. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **2005**, *309*, 1570–1573.
89. Hung, T.; Wang, Y.; Lin, M.F.; Koegel, A.K.; Kotake, Y.; Grant, G.D.; Horlings, H.M.; Shah, N.; Umbricht, C.; Wang, P.; *et al.* Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat. Genet*. **2011**, *43*, 621–629.
90. Martianov, I.; Ramadass, A.; Serra Barros, A.; Chow, N.; Akoulitchev, A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* **2007**, *445*, 666–670.
91. Mazo, A.; Hodgson, J.W.; Petruk, S.; Sedkov, Y.; Brock, H.W. Transcriptional interference: An unexpected layer of complexity in gene regulation. *J. Cell Sci*. **2007**, *120*, 2755–2761.
92. Tripathi, V.; Ellis, J.D.; Shen, Z.; Song, D.Y.; Pan, Q.; Watt, A.T.; Freier, S.M.; Bennett, C.F.; Sharma, A.; Bubulya, P.A.; *et al.* The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* **2010**, *39*, 925–938.
93. Mao, Y.S.; Sunwoo, H.; Zhang, B.; Spector, D.L. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat. Cell Biol*. **2011**, *13*, 95–101.
94. Bond, C.S.; Fox, A.H. Paraspeckles: Nuclear bodies built on long noncoding RNA. *J. Cell Biol*. **2009**, *186*, 637–644.
95. Clemson, C.M.; Hutchinson, J.N.; Sara, S.A.; Ensminger, A.W.; Fox, A.H.; Chess, A.; Lawrence, J.B. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell* **2009**, *33*, 717–726.
96. Chen, L.L.; Carmichael, G.G. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: Functional role of a nuclear noncoding RNA. *Mol. Cell* **2009**, *35*, 467–478.
97. Sunwoo, H.; Dinger, M.E.; Wilusz, J.E.; Amaral, P.P.; Mattick, J.S.; Spector, D.L. MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res*. **2009**, *19*, 347–359.
98. Eissmann, M.; Gutschner, T.; Hammerle, M.; Gunther, S.; Caudron-Herger, M.; Gross, M.; Schirmacher, P.; Rippe, K.; Braun, T.; Zornig, M.; *et al.* Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. *RNA Biol*. **2012**, *9*, 1076–1087.
99. Ip, J.Y.; Nakagawa, S. Long non-coding RNAs in nuclear bodies. *Dev Growth Differ*. **2012**, *54*, 44–54.
100. Nakagawa, S.; Naganuma, T.; Shioi, G.; Hirose, T. Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. *J. Cell Biol*. **2011**, *193*, 31–39.
101. Zhang, B.; Arun, G.; Mao, Y.S.; Lazar, Z.; Hung, G.; Bhattacharjee, G.; Xiao, X.; Booth, C.J.; Wu, J.; Zhang, C.; *et al.* The lncRNA Malat1 is dispensable for mouse development but its transcription plays a *cis*-regulatory role in the adult. *Cell Rep*. **2012**, *2*, 111–123.
102. Yin, Q.F.; Yang, L.; Zhang, Y.; Xiang, J.F.; Wu, Y.W.; Carmichael, G.G.; Chen, L.L. Long noncoding RNAs with snoRNA ends. *Mol. Cell*. **2012**, *48*, 219–230.

103. Carrieri, C.; Cimatti, L.; Biagioli, M.; Beugnet, A.; Zucchelli, S.; Fedele, S.; Pesce, E.; Ferrer, I.; Collavin, L.; Santoro, C.; *et al.* Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* **2012**, *491*, 454–457.
104. Yoon, J.H.; Abdelmohsen, K.; Srikantan, S.; Yang, X.; Martindale, J.L.; De, S.; Huarte, M.; Zhan, M.; Becker, K.G.; Gorospe, M. LincRNA-p21 suppresses target mRNA translation. *Mol. Cell* **2012**, *47*, 648–655.
105. Kretz, M.; Siprashvili, Z.; Chu, C.; Webster, D.E.; Zehnder, A.; Qu, K.; Lee, C.S.; Flockhart, R.J.; Groff, A.F.; Chow, J.; *et al.* Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **2013**, *493*, 231–235.
106. Gong, C.; Maquat, L.E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **2011**, *470*, 284–288.
107. Cesana, M.; Cacchiarelli, D.; Legnini, I.; Santini, T.; Sthandier, O.; Chinappi, M.; Tramontano, A.; Bozzoni, I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **2011**, *147*, 358–369.
108. Cai, X.; Cullen, B.R. The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* **2007**, *13*, 313–316.
109. Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914.
110. Spitale, R.C.; Tsai, M.C.; Chang, H.Y. RNA templating the epigenome: Long noncoding RNAs as molecular scaffolds. *Epigenetics* **2011**, *6*, 539–543.
111. Zappulla, D.C.; Cech, T.R. RNA as a flexible scaffold for proteins: Yeast telomerase and beyond. *Cold Spring Harb. Symp. Quant. Biol*. **2006**, *71*, 217–224.
112. Kertesz, M.; Wan, Y.; Mazor, E.; Rinn, J.L.; Nutter, R.C.; Chang, H.Y.; Segal, E. Genome-wide measurement of RNA secondary structure in yeast. *Nature* **2010**, *467*, 103–107.
113. Castello, A.; Fischer, B.; Eichelbaum, K.; Horos, R.; Beckmann, B.M.; Strein, C.; Davey, N.E.; Humphreys, D.T.; Preiss, T.; Steinmetz, L.M.; *et al.* Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **2012**, *149*, 1393–1406.
114. Guttman, M.; Donaghey, J.; Carey, B.W.; Garber, M.; Grenier, J.K.; Munson, G.; Young, G.; Lucas, A.B.; Ach, R.; Bruhn, L.; *et al.* lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **2011**, *477*, 295–300.
115. Sheik Mohamed, J.; Gaughwin, P.M.; Lim, B.; Robson, P.; Lipovich, L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *RNA* **2010**, *16*, 324–337.
116. Taft, R.J.; Pang, K.C.; Mercer, T.R.; Dinger, M.; Mattick, J.S. Non-coding RNAs: Regulators of disease. *J. Pathol*. **2010**, *220*, 126–139.
117. Ishii, N.; Ozaki, K.; Sato, H.; Mizuno, H.; Saito, S.; Takahashi, A.; Miyamoto, Y.; Ikegawa, S.; Kamatani, N.; Hori, M.; *et al.* Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction. *J. Hum. Genet*. **2006**, *51*, 1087–1099.
118. Pasmant, E.; Laurendeau, I.; Heron, D.; Vidaud, M.; Vidaud, D.; Bieche, I. Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: Identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res*. **2007**, *67*, 3963–3969.

119. Sonkoly, E.; Bata-Csorgo, Z.; Pivarcsi, A.; Polyanka, H.; Kenderessy-Szabo, A.; Molnar, G.; Szentpali, K.; Bari, L.; Megyeri, K.; Mandi, Y.; *et al.* Identification and characterization of a novel, psoriasis susceptibility-related noncoding RNA gene, PRINS. *J. Biol. Chem*. **2005**, *280*, 24159–24167.
120. Daughters, R.S.; Tuttle, D.L.; Gao, W.; Ikeda, Y.; Moseley, M.L.; Ebner, T.J.; Swanson, M.S.; Ranum, L.P. RNA gain-of-function in spinocerebellar ataxia type 8. *PLoS Genet*. **2009**, *5*, e1000600.
121. Khalil, A.M.; Faghihi, M.A.; Modarresi, F.; Brothers, S.P.; Wahlestedt, C. A novel RNA transcript with antiapoptotic function is silenced in fragile X syndrome. *PLoS One* **2008**, *3*, e1486.
122. Kogo, R.; Shimamura, T.; Mimori, K.; Kawahara, K.; Imoto, S.; Sudo, T.; Tanaka, F.; Shibata, K.; Suzuki, A.; Komune, S.; *et al.* Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res*. **2011**, *71*, 6320–6326.
123. Chung, S.; Nakagawa, H.; Uemura, M.; Piao, L.; Ashikawa, K.; Hosono, N.; Takata, R.; Akamatsu, S.; Kawaguchi, T.; Morizono, T.; *et al.* Association of a novel long non-coding RNA in 8q24 with prostate cancer susceptibility. *Cancer Sci*. **2011**, *102*, 245–252.
124. Yang, Z.; Zhou, L.; Wu, L.M.; Lai, M.C.; Xie, H.Y.; Zhang, F.; Zheng, S.S. Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. *Ann. Surg. Oncol*. **2011**, *18*, 1243–1250.
125. Lai, M.C.; Yang, Z.; Zhou, L.; Zhu, Q.Q.; Xie, H.Y.; Zhang, F.; Wu, L.M.; Chen, L.M.; Zheng, S.S. Long non-coding RNA MALAT-1 overexpression predicts tumor recurrence of hepatocellular carcinoma after liver transplantation. *Med. Oncol*. **2012**, *29*, 1810–1816.
126. Calin, G.A.; Pekarsky, Y.; Croce, C.M. The role of microRNA and other non-coding RNA in the pathogenesis of chronic lymphocytic leukemia. *Best. Pract. Res. Clin. Haematol*. **2007**, *20*, 425–437.
127. Khaitan, D.; Dinger, M.E.; Mazar, J.; Crawford, J.; Smith, M.A.; Mattick, J.S.; Perera, R.J. The melanoma-upregulated long noncoding RNA SPRY4-IT1 modulates apoptosis and invasion. *Cancer Res*. **2011**, *71*, 3852–3862.
128. Huarte, M.; Rinn, J.L. Large non-coding RNAs: Missing links in cancer? *Hum. Mol. Genet*. **2010**, *19*, R152–R161.
129. Li, L.; Feng, T.; Lian, Y.; Zhang, G.; Garen, A.; Song, X. Role of human noncoding RNAs in the control of tumorigenesis. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12956–12961.
130. Yu, W.; Gius, D.; Onyango, P.; Muldoon-Jacobs, K.; Karp, J.; Feinberg, A.P.; Cui, H. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature* **2008**, *451*, 202–206.
131. Lawrie, C.H. MicroRNAs and lymphomagenesis: A functional review. *Br. J. Haematol*. **2013**, *160*, 571–581.
132. Agirre, X.; Martinez-Climent, J.A.; Odero, M.D.; Prosper, F. Epigenetic regulation of miRNA genes in acute leukemia. *Leukemia* **2012**, *26*, 395–403.
133. Elton, T.S.; Selemon, H.; Elton, S.M.; Parinandi, N.L. Regulation of the MIR155 host gene in physiological and pathological processes. *Gene* **2012**, doi:10.1016/j.gene.2012.12.009.

134. Vargova, K.; Curik, N.; Burda, P.; Basova, P.; Kulvait, V.; Pospisil, V.; Savvulidi, F.; Kokavec, J.; Necas, E.; Berkova, A.; *et al.* MYB transcriptionally regulates the miR-155 host gene in chronic lymphocytic leukemia. *Blood* **2011**, *117*, 3816–3825.
135. Nielsen, C.B.; Shomron, N.; Sandberg, R.; Hornstein, E.; Kitzman, J.; Burge, C.B. Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA* **2007**, *13*, 1894–1910.
136. Ji, M.; Rao, E.; Ramachandrareddy, H.; Shen, Y.; Jiang, C.; Chen, J.; Hu, Y.; Rizzino, A.; Chan, W.C.; Fu, K.; *et al.* The miR-17–92 microRNA cluster is regulated by multiple mechanisms in B-cell malignancies. *Am. J. Pathol*. **2011**, *179*, 1645–1656.
137. Rinaldi, A.; Poretti, G.; Kwee, I.; Zucca, E.; Catapano, C.V.; Tibiletti, M.G.; Bertoni, F. Concomitant MYC and microRNA cluster miR-17–92 (C13orf25) amplification in human mantle cell lymphoma. *Leuk Lymphoma* **2007**, *48*, 410–412.
138. Hayashita, Y.; Osada, H.; Tatematsu, Y.; Yamada, H.; Yanagisawa, K.; Tomida, S.; Yatabe, Y.; Kawahara, K.; Sekido, Y.; Takahashi, T. A polycistronic microRNA cluster, miR-17–92, is overexpressed in human lung cancers and enhances cell proliferation. *Cancer Res*. **2005**, *65*, 9628–9632.
139. Humphreys, K.J.; Cobiac, L.; Le Leu, R.K.; Van der Hoek, M.B.; Michael, M.Z. Histone deacetylase inhibition in colorectal cancer cells reveals competing roles for members of the oncogenic miR-17–92 cluster. *Mol. Carcinog*. **2013**, *52*, 459–474.
140. Kunkeaw, N.; Jeon, S.H.; Lee, K.; Johnson, B.H.; Tanasanvimon, S.; Javle, M.; Pairojkul, C.; Chamgramol, Y.; Wongfieng, W.; Gong, B.; *et al.* Cell death/proliferation roles for nc886, a non-coding RNA, in the protein kinase R pathway in cholangiocarcinoma. *Oncogene* **2012**, doi:10.1038/onc.2012.382.
141. Treppendahl, M.B.; Qiu, X.; Sogaard, A.; Yang, X.; Nandrup-Bus, C.; Hother, C.; Andersen, M.K.; Kjeldsen, L.; Mollgaard, L.; Hellstrom-Lindberg, E.; *et al.* Allelic methylation levels of the noncoding VTRNA2–1 located on chromosome 5q31.1 predict outcome in AML. *Blood* **2012**, *119*, 206–216.
142. He, L.; Thomson, J.M.; Hemann, M.T.; Hernando-Monge, E.; Mu, D.; Goodson, S.; Powers, S.; Cordon-Cardo, C.; Lowe, S.W.; Hannon, G.J.; *et al.* A microRNA polycistron as a potential human oncogene. *Nature* **2005**, *435*, 828–833.
143. Nagoshi, H.; Taki, T.; Hanamura, I.; Nitta, M.; Otsuki, T.; Nishida, K.; Okuda, K.; Sakamoto, N.; Kobayashi, S.; Yamamoto-Sugitani, M.; *et al.* Frequent PVT1 rearrangement and novel chimeric genes PVT1-NBEA and PVT1-WWOX occur in multiple myeloma with 8q24 abnormality. *Cancer Res*. **2012**, *72*, 4954–4962.
144. Zeidler, R.; Joos, S.; Delecluse, H.J.; Klobeck, G.; Vuillaume, M.; Lenoir, G.M.; Bornkamm, G.W.; Lipp, M. Breakpoints of Burkitt's lymphoma t(8;22) translocations map within a distance of 300 kb downstream of MYC. *Genes Chromosomes Cancer* **1994**, *9*, 282–287.
145. Enciso-Mora, V.; Broderick, P.; Ma, Y.; Jarrett, R.F.; Hjalgrim, H.; Hemminki, K.; van den Berg, A.; Olver, B.; Lloyd, A.; Dobbins, S.E.; *et al.* A genome-wide association study of Hodgkin's lymphoma identifies new susceptibility loci at 2p16.1 (REL), 8q24.21 and 10p14 (GATA3). *Nat. Genet*. **2010**, *42*, 1126–1130.

146. Huppi, K.; Volfovsky, N.; Runfola, T.; Jones, T.L.; Mackiewicz, M.; Martin, S.E.; Mushinski, J.F.; Stephens, R.; Caplen, N.J. The identification of microRNAs in a genomically unstable region of human chromosome 8q24. *Mol. Cancer Res*. **2008**, *6*, 212–221.
147. Beck-Engeser, G.B.; Lum, A.M.; Huppi, K.; Caplen, N.J.; Wang, B.B.; Wabl, M. Pvt1-encoded microRNAs in oncogenesis. *Retrovirology* **2008**, *5*, 4.
148. Barsotti, A.M.; Beckerman, R.; Laptenko, O.; Huppi, K.; Caplen, N.J.; Prives, C. p53-Dependent induction of PVT1 and miR-1204. *J. Biol. Chem*. **2011**, *287*, 2509–2519.
149. Zhang, X.; Gejman, R.; Mahta, A.; Zhong, Y.; Rice, K.A.; Zhou, Y.; Cheunsuchon, P.; Louis, D.N.; Klibanski, A. Maternally expressed gene 3, an imprinted noncoding RNA gene, is associated with meningioma pathogenesis and progression. *Cancer Res*. **2010**, *70*, 2350–2358.
150. Zhou, Y.; Zhang, X.; Klibanski, A. MEG3 noncoding RNA: A tumor suppressor. *J. Mol. Endocrinol*. **2012**, *48*, R45–R53.
151. Saitou, M.; Sugimoto, J.; Hatakeyama, T.; Russo, G.; Isobe, M. Identification of the TCL6 genes within the breakpoint cluster region on chromosome 14q32 in T-cell leukemia. *Oncogene* **2000**, *19*, 2796–2802.
152. Zhang, X.; Rice, K.; Wang, Y.; Chen, W.; Zhong, Y.; Nakayama, Y.; Zhou, Y.; Klibanski, A. Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: Isoform structure, expression, and functions. *Endocrinology* **2010**, *151*, 939–947.
153. Zhou, Y.; Zhong, Y.; Wang, Y.; Zhang, X.; Batista, D.L.; Gejman, R.; Ansell, P.J.; Zhao, J.; Weng, C.; Klibanski, A. Activation of p53 by MEG3 non-coding RNA. *J. Biol. Chem*. **2007**, *282*, 24731–24742.
154. Gejman, R.; Batista, D.L.; Zhong, Y.; Zhou, Y.; Zhang, X.; Swearingen, B.; Stratakis, C.A.; Hedley-Whyte, E.T.; Klibanski, A. Selective loss of MEG3 expression and intergenic differentially methylated region hypermethylation in the MEG3/DLK1 locus in human clinically nonfunctioning pituitary adenomas. *J. Clin. Endocrinol. Metab*. **2008**, *93*, 4119–4125.
155. Benetatos, L.; Vartholomatos, G.; Hatzimichael, E. MEG3 imprinted gene contribution in tumorigenesis. *Int. J. Cancer*. **2011**, *129*, 773–779.
156. Garding, A.; Bhattacharya, N.; Claus, R.; Ruppel, M.; Tschuch, C.; Filarsky, K.; Idler, I.; Zucknick, M.; Caudron-Herger, M.; Oakes, C.; *et al.* Epigenetic upregulation of lncRNAs at 13q14.3 in leukemia is linked to the in cis downregulation of a gene cluster that targets NF-kB. *PLoS Genet*. **2013**, *9*, e1003373.
157. Takeuchi, S.; Hofmann, W.K.; Tsukasaki, K.; Takeuchi, N.; Ikezoe, T.; Matsushita, M.; Uehara, Y.; Phillip Koeffler, H. Loss of H19 imprinting in adult T-cell leukaemia/lymphoma. *Br. J. Haematol*. **2007**, *137*, 380–381.
158. Li, J.; Rhodes, J.C.; Askew, D.S. Evolutionary conservation of putative functional domains in the human homolog of the murine His-1 gene. *Gene* **1997**, *184*, 169–176.
159. Dallosso, A.R.; Hancock, A.L.; Malik, S.; Salpekar, A.; King-Underwood, L.; Pritchard-Jones, K.; Peters, J.; Moorwood, K.; Ward, A.; Malik, K.T.; *et al.* Alternately spliced WT1 antisense transcripts interact with WT1 sense RNA and show epigenetic and splicing defects in cancer. *RNA* **2007**, *13*, 2287–2299.

160. Ellis, B.C.; Molloy, P.L.; Graham, L.D. CRNDE: A long non-coding RNA involved in CanceR, neurobiology, and development. *Front. Genet*. **2012**, *3*, 270.
161. Taskinen, M.; Ranki, A.; Pukkala, E.; Jeskanen, L.; Kaitila, I.; Makitie, O. Extended follow-up of the Finnish cartilage-hair hypoplasia cohort confirms high incidence of non-Hodgkin lymphoma and basal cell carcinoma. *Am. J. Med. Genet A*. **2008**, *146A*, 2370–2375.
162. Braconi, C.; Valeri, N.; Kogure, T.; Gasparini, P.; Huang, N.; Nuovo, G.J.; Terracciano, L.; Croce, C.M.; Patel, T. Expression and functional role of a transcribed noncoding RNA with an ultraconserved element in hepatocellular carcinoma. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 786–791.
163. Yang, R.; Frank, B.; Hemminki, K.; Bartram, C.R.; Wappenschmidt, B.; Sutter, C.; Kiechle, M.; Bugert, P.; Schmutzler, R.K.; Arnold, N.; *et al.* SNPs in ultraconserved elements and familial breast cancer risk. *Carcinogenesis* **2008**, *29*, 351–355.
164. Zhang, X.; Lian, Z.; Padden, C.; Gerstein, M.B.; Rozowsky, J.; Snyder, M.; Gingeras, T.R.; Kapranov, P.; Weissman, S.M.; Newburger, P.E. A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. *Blood* **2009**, *113*, 2526–2534.
165. Wagner, L.A.; Christensen, C.J.; Dunn, D.M.; Spangrude, G.J.; Georgelas, A.; Kelley, L.; Esplin, M.S.; Weiss, R.B.; Gleich, G.J. EGO, a novel, noncoding RNA gene, regulates eosinophil granule protein transcript expression. *Blood* **2007**, *109*, 5191–5198.
166. Ebralidze, A.K.; Guibal, F.C.; Steidl, U.; Zhang, P.; Lee, S.; Bartholdy, B.; Jorda, M.A.; Petkova, V.; Rosenbauer, F.; Huang, G.; *et al.* PU.1 expression is modulated by the balance of functional sense and antisense RNAs regulated by a shared cis-regulatory element. *Genes Dev*. **2008**, *22*, 2085–2092.
167. Hu, W.; Yuan, B.; Flygare, J.; Lodish, H.F. Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. *Genes Dev*. **2011**, *25*, 2573–2578.
168. Aoki, K.; Harashima, A.; Sano, M.; Yokoi, T.; Nakamura, S.; Kibata, M.; Hirose, T. A thymus-specific noncoding RNA, Thy-ncR1, is a cytoplasmic riboregulator of MFAP4 mRNA in immature T-cell lines. *BMC Mol. Biol*. **2010**, *11*, 99.
169. Marasa, B.S.; Srikantan, S.; Martindale, J.L.; Kim, M.M.; Lee, E.K.; Gorospe, M.; Abdelmohsen, K. MicroRNA profiling in human diploid fibroblasts uncovers miR-519 role in replicative senescence. *Aging (Albany NY)* **2010**, *2*, 333–343.
170. Meyer, K.B.; Maia, A.T.; O'Reilly, M.; Ghoussaini, M.; Prathalingam, R.; Porter-Gill, P.; Ambs, S.; Prokunina-Olsson, L.; Carroll, J.; Ponder, B.A. A functional variant at a prostate cancer predisposition locus at 8q24 is associated with PVT1 expression. *PLoS Genet*. **2011**, *7*, e1002165.
171. Guan, Y.; Kuo, W.L.; Stilwell, J.L.; Takano, H.; Lapuk, A.V.; Fridlyand, J.; Mao, J.H.; Yu, M.; Miller, M.A.; Santos, J.L.; *et al.* Amplification of PVT1 contributes to the pathophysiology of ovarian and breast cancer. *Clin. Cancer Res*. **2007**, *13*, 5745–5755.
172. Palumbo, A.P.; Boccadoro, M.; Battaglio, S.; Corradini, P.; Tsichlis, P.N.; Huebner, K.; Pileri, A.; Croce, C.M. Human homologue of Moloney leukemia virus integration-4 locus (MLVI-4), located 20 kilobases 3' of the myc gene, is rearranged in multiple myelomas. *Cancer Res*. **1990**, *50*, 6478–6482.

173. Borg, A.; Baldetorp, B.; Ferno, M.; Olsson, H.; Sigurdsson, H. c-myc amplification is an independent prognostic factor in postmenopausal breast cancer. *Int. J. Cancer*. **1992**, *51*, 687–691.
174. You, L.; Chang, D.; Du, H.Z.; Zhao, Y.P. Genome-wide screen identifies PVT1 as a regulator of Gemcitabine sensitivity in human pancreatic cancer cells. *Biochem. Biophys. Res. Commun*. **2011**, *407*, 1–6.
175. Carramusa, L.; Contino, F.; Ferro, A.; Minafra, L.; Perconti, G.; Giallongo, A.; Feo, S. The PVT-1 oncogene is a Myc protein target that is overexpressed in transformed cells. *J. Cell. Physiol*. **2007**, *213*, 511–518.
176. Gil, J.; Peters, G. Regulation of the INK4b-ARF-INK4a tumour suppressor locus: All for one or one for all. *Nat. Rev. Mol. Cell Biol*. **2006**, *7*, 667–677.
177. Popov, N.; Gil, J. Epigenetic regulation of the INK4b-ARF-INK4a locus: In sickness and in health. *Epigenetics* **2010**, *5*, 685–690.
178. Kotake, Y.; Nakagawa, T.; Kitagawa, K.; Suzuki, S.; Liu, N.; Kitagawa, M.; Xiong, Y. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. *Oncogene* **2011**, *30*, 1956–1962.
179. Pasmant, E.; Sabbagh, A.; Vidaud, M.; Bieche, I. ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. *FASEB J*. **2011**, *25*, 444–448.
180. Burd, C.E.; Jeck, W.R.; Liu, Y.; Sanoff, H.K.; Wang, Z.; Sharpless, N.E. Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet*. **2010**, *6*, e1001233.
181. Liu, Y.; Sanoff, H.K.; Cho, H.; Burd, C.E.; Torrice, C.; Mohlke, K.L.; Ibrahim, J.G.; Thomas, N.E.; Sharpless, N.E. INK4/ARF transcript expression is associated with chromosome 9p21 variants linked to atherosclerosis. *PLoS One* **2009**, *4*, e5027.
182. Cunnington, M.S.; Santibanez Koref, M.; Mayosi, B.M.; Burn, J.; Keavney, B. Chromosome 9p21 SNPs associated with multiple disease phenotypes correlate with anril expression. *PLoS Genet*. **2010**, *6*, e1000899.
183. Iacobucci, I.; Sazzini, M.; Garagnani, P.; Ferrari, A.; Boattini, A.; Lonetti, A.; Papayannidis, C.; Mantovani, V.; Marasco, E.; Ottaviani, E.; *et al.* A polymorphism in the chromosome 9p21 ANRIL locus is associated to Philadelphia positive acute lymphoblastic leukemia. *Leuk Res*. **2011**, *35*, 1052–1059.
184. Miyoshi, N.; Wagatsuma, H.; Wakana, S.; Shiroishi, T.; Nomura, M.; Aisaka, K.; Kohda, T.; Surani, M.A.; Kaneko-Ishino, T.; Ishino, F. Identification of an imprinted gene, Meg3/Gtl2 and its human homologue MEG3, first mapped on mouse distal chromosome 12 and human chromosome 14q. *Genes Cells* **2000**, *5*, 211–220.
185. Zhang, X.; Zhou, Y.; Mehta, K.R.; Danila, D.C.; Scolavino, S.; Johnson, S.R.; Klibanski, A. A pituitary-derived MEG3 isoform functions as a growth suppressor in tumor cells. *J. Clin. Endocrinol. Metab*. **2003**, *88*, 5119–5126.
186. Braconi, C.; Kogure, T.; Valeri, N.; Huang, N.; Nuovo, G.; Costinean, S.; Negrini, M.; Miotto, E.; Croce, C.M.; Patel, T. microRNA-29 can regulate expression of the long non-coding RNA gene MEG3 in hepatocellular cancer. *Oncogene* **2011**, *30*, 4750–4756.

187. Hagan, J.P.; O'Neill, B.L.; Stewart, C.L.; Kozlov, S.V.; Croce, C.M. At least ten genes define the imprinted Dlk1-Dio3 cluster on mouse chromosome 12qF1. *PLoS One* **2009**, *4*, e4352.
188. Schuster-Gossler, K.; Bilinski, P.; Sado, T.; Ferguson-Smith, A.; Gossler, A. The mouse Gtl2 gene is differentially expressed during embryonic development, encodes multiple alternatively spliced transcripts, and may act as an RNA. *Dev. Dyn*. **1998**, *212*, 214–228.
189. Mondal, T.; Rasmussen, M.; Pandey, G.K.; Isaksson, A.; Kanduri, C. Characterization of the RNA content of chromatin. *Genome Res*. **2010**, *20*, 899–907.
190. Gordon, F.E.; Nutt, C.L.; Cheunsuchon, P.; Nakayama, Y.; Provencher, K.A.; Rice, K.A.; Zhou, Y.; Zhang, X.; Klibanski, A. Increased expression of angiogenic genes in the brains of mouse meg3-null embryos. *Endocrinology* **2010**, *151*, 2443–2452.
191. Ying, L.; Huang, Y.; Chen, H.; Wang, Y., Xia, L.; Chen, Y.; Liu, Y.; Qiu, F. Downregulated MEG3 activates autophagy and increases cell proliferation in bladder cancer. *Mol. Biosyst*. **2013**, *9*, 407–411.
192. Dal Bo, M.; Rossi, F.M.; Rossi, D.; Deambrogi, C.; Bertoni, F.; Del Giudice, I.; Palumbo, G.; Nanni, M.; Rinaldi, A.; Kwee, I.; *et al.* 13q14 deletion size and number of deleted cells both influence prognosis in chronic lymphocytic leukemia. *Genes Chromosomes Cancer* **2011**, *50*, 633–643.
193. Cimmino, A.; Calin, G.A.; Fabbri, M.; Iorio, M.V.; Ferracin, M.; Shimizu, M.; Wojcik, S.E.; Aqeilan, R.I.; Zupo, S.; Dono, M.; *et al.* miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13944–13949.
194. Klein, U.; Lia, M.; Crespo, M.; Siegel, R.; Shen, Q.; Mo, T.; Ambesi-Impiombato, A.; Califano, A.; Migliazza, A.; Bhagat, G.; *et al.* The DLEU2/miR-15a/16–1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. *Cancer Cell* **2010**, *17*, 28–40.
195. Lerner, M.; Harada, M.; Loven, J.; Castro, J.; Davis, Z.; Oscier, D.; Henriksson, M.; Sangfelt, O.; Grander, D.; Corcoran, M.M. DLEU2, frequently deleted in malignancy, functions as a critical host gene of the cell cycle inhibitory microRNAs miR-15a and miR-16–1. *Exp. Cell Res*. **2009**, *315*, 2941–2952.
196. Coccia, E.M.; Cicala, C.; Charlesworth, A.; Ciccarelli, C.; Rossi, G.B.; Philipson, L.; Sorrentino, V. Regulation and expression of a growth arrest-specific gene (gas5) during growth, differentiation, and development. *Mol. Cell Biol*. **1992**, *12*, 3514–3521.
197. Mourtada-Maarabouni, M.; Pickard, M.R.; Hedge, V.L.; Farzaneh, F.; Williams, G.T. GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene* **2009**, *28*, 195–208.
198. Williams, G.T.; Mourtada-Maarabouni, M.; Farzaneh, F. A critical role for non-coding RNA GAS5 in growth arrest and rapamycin inhibition in human T-lymphocytes. *Biochem. Soc. Trans*. **2011**, *39*, 482–486.
199. Mourtada-Maarabouni, M.; Hasan, A.M.; Farzaneh, F.; Williams, G.T. Inhibition of human T-cell proliferation by mammalian target of rapamycin (mTOR) antagonists requires noncoding RNA growth-arrest-specific transcript 5 (GAS5). *Mol. Pharmacol*. **2010**, *78*, 19–28.
200. Barsyte-Lovejoy, D.; Lau, S.K.; Boutros, P.C.; Khosravi, F.; Jurisica, I.; Andrulis, I.L.; Tsao, M.S.; Penn, L.Z. The c-Myc oncogene directly induces the H19 noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Res*. **2006**, *66*, 5330–5337.

201. Hibi, K.; Nakamura, H.; Hirai, A.; Fujikake, Y.; Kasai, Y.; Akiyama, S.; Ito, K.; Takagi, H. Loss of H19 imprinting in esophageal cancer. *Cancer Res*. **1996**, *56*, 480–482.
202. Fellig, Y.; Ariel, I.; Ohana, P.; Schachter, P.; Sinelnikov, I.; Birman, T.; Ayesh, S.; Schneider, T.; de Groot, N.; Czerniak, A.; *et al.* H19 expression in hepatic metastases from a range of human carcinomas. *J. Clin. Pathol*. **2005**, *58*, 1064–1068.
203. Matouk, I.J.; DeGroot, N.; Mezan, S.; Ayesh, S.; Abu-lail, R.; Hochberg, A.; Galun, E. The H19 non-coding RNA is essential for human tumor growth. *PLoS One* **2007**, *2*, e845.
204. Berteaux, N.; Lottin, S.; Monte, D.; Pinte, S.; Quatannens, B.; Coll, J.; Hondermarck, H.; Curgy, J.J.; Dugimont, T.; Adriaenssens, E. H19 mRNA-like noncoding RNA promotes breast cancer cell proliferation through positive control by E2F1. *J. Biol. Chem*. **2005**, *280*, 29625–29636.
205. Nunez, C.; Bashein, A.M.; Brunet, C.L.; Hoyland, J.A.; Freemont, A.J.; Buckle, A.M.; Murphy, C.; Cross, M.A.; Lucas, G.; Bostock, V.J.; *et al.* Expression of the imprinted tumour-suppressor gene H19 is tightly regulated during normal haematopoiesis and is reduced in haematopoietic precursors of patients with the myeloproliferative disease polycythaemia vera. *J. Pathol*. **2000**, *190*, 61–68.
206. Bock, O.; Schlue, J.; Kreipe, H. Reduced expression of H19 in bone marrow cells from chronic myeloproliferative disorders. *Leukemia* **2003**, *17*, 815–816.
207. Tessema, M.; Langer, F.; Bock, O.; Seltsam, A.; Metzig, K.; Hasemeier, B.; Kreipe, H.; Lehmann, U. Down-regulation of the IGF-2/H19 locus during normal and malignant hematopoiesis is independent of the imprinting pattern. *Int. J. Oncol*. **2005**, *26*, 499–507.
208. Dugimont, T.; Montpellier, C.; Adriaenssens, E.; Lottin, S.; Dumont, L.; Iotsova, V.; Lagrou, C.; Stehelin, D.; Coll, J.; Curgy, J.J. The H19 TATA-less promoter is efficiently repressed by wild-type tumor suppressor gene product p53. *Oncogene* **1998**, *16*, 2395–2401.
209. Farnebo, M.; Bykov, V.J.; Wiman, K.G. The p53 tumor suppressor: A master regulator of diverse cellular processes and therapeutic target in cancer. *Biochem. Biophys. Res. Commun*. **2010**, *396*, 85–89.
210. Tsang, W.P.; Ng, E.K.; Ng, S.S.; Jin, H.; Yu, J.; Sung, J.J.; Kwok, T.T. Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis* **2010**, *31*, 350–358.
211. Yoshimizu, T.; Miroglio, A.; Ripoche, M.A.; Gabory, A.; Vernucci, M.; Riccio, A.; Colnot, S.; Godard, C.; Terris, B.; Jammes, H.; *et al.* The H19 locus acts *in vivo* as a tumor suppressor. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 12417–12422.
212. Leighton, P.A.; Saam, J.R.; Ingram, R.S.; Stewart, C.L.; Tilghman, S.M. An enhancer deletion affects both H19 and Igf2 expression. *Genes Dev*. **1995**, *9*, 2079–2089.
213. Scaruffi, P.; Stigliani, S.; Moretti, S.; Coco, S.; De Vecchi, C.; Valdora, F.; Garaventa, A.; Bonassi, S.; Tonini, G.P. Transcribed-Ultra Conserved Region expression is associated with outcome in high-risk neuroblastoma. *BMC Cancer* **2009**, *9*, 441.
214. Mestdagh, P.; Fredlund, E.; Pattyn, F.; Rihani, A.; Van Maerken, T.; Vermeulen, J.; Kumps, C.; Menten, B.; de Preter, K.; Schramm, A.; *et al.* An integrative genomics screen uncovers ncRNA T-UCR functions in neuroblastoma tumours. *Oncogene* **2010**, *29*, 3583–3592.
215. Sana, J.; Hankeova, S.; Svoboda, M.; Kiss, I.; Vyzula, R.; Slaby, O. Expression levels of transcribed ultraconserved regions uc.73 and uc.388 are altered in colorectal cancer. *Oncology* **2012**, *82*, 114–118.

216. Rossi, S.; Sevignani, C.; Nnadi, S.C.; Siracusa, L.D.; Calin, G.A. Cancer-associated genomic regions (CAGRs) and noncoding RNAs: Bioinformatics and therapeutic implications. *Mamm. Genome*. **2008**, *19*, 526–540.
217. Ng, D.; Toure, O.; Wei, M.H.; Arthur, D.C.; Abbasi, F.; Fontaine, L.; Marti, G.E.; Fraumeni, J.F., Jr.; Goldin, L.R.; Caporaso, N.; *et al.* Identification of a novel chromosome region, 13q21.33-q22.2, for susceptibility genes in familial chronic lymphocytic leukemia. *Blood* **2007**, *109*, 916–925.
218. Calin, G.A.; Ferracin, M.; Cimmino, A.; Di Leva, G.; Shimizu, M.; Wojcik, S.E.; Iorio, M.V.; Visone, R.; Sever, N.I.; Fabbri, M.; *et al.* A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N. Engl. J. Med*. **2005**, *353*, 1793–1801.
219. Notari, M.; Neviani, P.; Santhanam, R.; Blaser, B.W.; Chang, J.S.; Galietta, A.; Willis, A.E.; Roy, D.C.; Caligiuri, M.A.; Marcucci, G.; *et al.* A MAPK/HNRPK pathway controls BCR/ABL oncogenic potential by regulating MYC mRNA translation. *Blood* **2006**, *107*, 2507–2516.
220. Du, Q.; Wang, L.; Zhu, H.; Zhang, S.; Xu, L.; Zheng, W.; Liu, X. The role of heterogeneous nuclear ribonucleoprotein K in the progression of chronic myeloid leukemia. *Med. Oncol*. **2010**, *27*, 673–679.
221. Tsujimoto, Y.; Gorham, J.; Cossman, J.; Jaffe, E.; Croce, C.M. The t(14;18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. *Science* **1985**, *229*, 1390–1393.
222. Ridanpaa, M.; van Eenennaam, H.; Pelin, K.; Chadwick, R.; Johnson, C.; Yuan, B.; vanVenrooij, W.; Pruijn, G.; Salmela, R.; Rockas, S.; *et al.* Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia. *Cell* **2001**, *104*, 195–203.
223. Tanaka, R.; Satoh, H.; Moriyama, M.; Satoh, K.; Morishita, Y.; Yoshida, S.; Watanabe, T.; Nakamura, Y.; Mori, S. Intronic U50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t(3;6)(q27;q15) of human B-cell lymphoma. *Genes Cells* **2000**, *5*, 277–287.
224. Zhao, H.; Zhang, X.; Frazao, J.B.; Condino-Neto, A.; Newburger, P.E. HOX antisense lincRNA HOXA-AS2 is an apoptosis repressor in all trans retinoic acid treated NB4 promyelocytic leukemia cells. *J. Cell Biochem*. **2013**, doi:10.1002/jcb.24586.
225. Paralkar, V.R.; Weiss, M.J. A new 'Linc' between noncoding RNAs and blood development. *Genes Dev*. **2011**, *25*, 2555–2558.
226. Halvorsen, M.; Martin, J.S.; Broadaway, S.; Laederach, A. Disease-associated mutations that alter the RNA structural ensemble. *PLoS Genet*. **2010**, *6*, e1001074.
227. Wojcik, S.E.; Rossi, S.; Shimizu, M.; Nicoloso, M.S.; Cimmino, A.; Alder, H.; Herlea, V.; Rassenti, L.Z.; Rai, K.R.; Kipps, T.J.; *et al.* Non-codingRNA sequence variations in human chronic lymphocytic leukemia and colorectal cancer. *Carcinogenesis* **2010**, *31*, 208–215.
228. Scott, L.J.; Mohlke, K.L.; Bonnycastle, L.L.; Willer, C.J.; Li, Y.; Duren, W.L.; Erdos, M.R.; Stringham, H.M.; Chines, P.S.; Jackson, A.U.; *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **2007**, *316*, 1341–1345.
229. Broadbent, H.M.; Peden, J.F.; Lorkowski, S.; Goel, A.; Ongen, H.; Green, F.; Clarke, R.; Collins, R.; Franzosi, M.G.; Tognoni, G.; *et al.* Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. *Hum. Mol. Genet*. **2008**, *17*, 806–814.

230. Morrison, L.E.; Jewell, S.S.; Usha, L.; Blondin, B.A.; Rao, R.D.; Tabesh, B.; Kemper, M.; Batus, M.; Coon, J.S. Effects of ERBB2 amplicon size and genomic alterations of chromosomes 1, 3, and 10 on patient response to trastuzumab in metastatic breast cancer. *Genes Chromosomes Cancer* **2007**, *46*, 397–405.
231. Nupponen, N.N.; Carpten, J.D. Prostate cancer susceptibility genes: Many studies, many results, no answers. *Cancer Metastasis Rev*. **2001**, *20*, 155–164.
232. Smedley, D.; Sidhar, S.; Birdsall, S.; Bennett, D.; Herlyn, M.; Cooper, C.; Shipley, J. Characterization of chromosome 1 abnormalities in malignant melanomas. *Genes Chromosomes Cancer* **2000**, *28*, 121–125.
233. Kagami, M.; O'Sullivan, M.J.; Green, A.J.; Watabe, Y.; Arisaka, O.; Masawa, N.; Matsuoka, K.; Fukami, M.; Matsubara, K.; Kato, F.; *et al.* The IG-DMR and the MEG3-DMR at human chromosome 14q32.2: Hierarchical interaction and distinct functional properties as imprinting control centers. *PLoS Genet*. **2010**, *6*, e1000992.
234. Astuti, D.; Latif, F.; Wagner, K.; Gentle, D.; Cooper, W.N.; Catchpoole, D.; Grundy, R.; Ferguson-Smith, A.C.; Maher, E.R. Epigenetic alteration at the DLK1-GTL2 imprinted domain in human neoplasia: Analysis of neuroblastoma, phaeochromocytoma and Wilms' tumour. *Br. J. Cancer*. **2005**, *92*, 1574–1580.
235. Benetatos, L.; Hatzimichael, E.; Dasoula, A.; Dranitsaris, G.; Tsiara, S.; Syrrou, M.; Georgiou, I.; Bourantas, K.L. CpG methylation analysis of the MEG3 and SNRPN imprinted genes in acute myeloid leukemia and myelodysplastic syndromes. *Leuk Res*. **2010**, *34*, 148–153.
236. Khoury, H.; Suarez-Saiz, F.; Wu, S.; Minden, M.D. An upstream insulator regulates DLK1 imprinting in AML. *Blood* **2010**, *115*, 2260–2263.
237. Rojas, J.J.; Guedan, S.; Searle, P.F.; Martinez-Quintanilla, J.; Gil-Hoyos, R.; Alcayaga-Miranda, F.; Cascallo, M.; Alemany, R. Minimal RB-responsive E1A promoter modification to attain potency, selectivity, and transgene-arming capacity in oncolytic adenoviruses. *Mol. Ther*. **2010**, *18*, 1960–1971.
238. Castello, A.; Horos, R.; Strein, C.; Fischer, B.; Eichelbaum, K.; Steinmetz, L.M.; Krijgsveld, J.; Hentze, M.W. System-wide identification of RNA-binding proteins by interactome capture. *Nat. Protoc*. **2013**, *8*, 491–500.
239. Fan, M.; Li, X.; Jiang, W.; Huang, Y.; Li, J.; Wang, Z. A long non-coding RNA, PTCSC3, as a tumor suppressor and a target of miRNAs in thyroid cancer cells. *Exp. Ther. Med*. **2013**, *5*, 1143–1146.
240. Chung, C.C.; Chanock, S.J. Current status of genome-wide association studies in cancer. *Hum. Genet*. **2011**, *130*, 59–78.
241. Giacomini, K.M.; Brett, C.M.; Altman, R.B.; Benowitz, N.L.; Dolan, M.E.; Flockhart, D.A.; Johnson, J.A.; Hayes, D.F.; Klein, T.; Krauss, R.M.; *et al.* The pharmacogenetics research network: From SNP discovery to clinical drug response. *Clin. Pharmacol. Ther*. **2007**, *81*, 328–345.
242. Mattick, J.S. The genetic signatures of noncoding RNAs. *PLoS Genet*. **2009**, *5*, e1000459.
243. Prensner, J.R.; Iyer, M.K.; Balbin, O.A.; Dhanasekaran, S.M.; Cao, Q.; Brenner, J.C.; Laxman, B.; Asangani, I.A.; Grasso, C.S.; Kominsky, H.D.; *et al.* Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. *Nat. Biotechnol*. **2011**, *29*, 742–749.

244. Matouk, I.J.; Abbasi, I.; Hochberg, A.; Galun, E.; Dweik, H.; Akkawi, M. Highly upregulated in liver cancer noncoding RNA is overexpressed in hepatic colorectal metastasis. *Eur. J. Gastroenterol. Hepatol*. **2009**, *21*, 688–692.
245. Panzitt, K.; Tschernatsch, M.M.; Guelly, C.; Moustafa, T.; Stradner, M.; Strohmaier, H.M.; Buck, C.R.; Denk, H.; Schroeder, R.; Trauner, M.; *et al.* Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. *Gastroenterology* **2007**, *132*, 330–342.
246. De Kok, J.B.; Verhaegh, G.W.; Roelofs, R.W.; Hessels, D.; Kiemeney, L.A.; Aalders, T.W.; Swinkels, D.W.; Schalken, J.A. DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. *Cancer Res*. **2002**, *62*, 2695–2698.
247. Ifere, G.O.;Ananaba, G.A. Prostate cancer gene expression marker 1 (PCGEM1): A patented prostate- specific non-coding gene and regulator of prostate cancer progression. *Recent Pat. DNA Gene Seq*. **2009**, *3*, 151–163.
248. Lee, G.L.; Dobi, A.; Srivastava, S. Prostate cancer: Diagnostic performance of the PCA3 urine test. *Nat. Rev. Urol*. **2011**, *8*, 123–124.
249. Tomlins, S.A.; Aubin, S.M.; Siddiqui, J.; Lonigro, R.J.; Sefton-Miller, L.; Miick, S.; Williamsen, S.; Hodge, P.; Meinke, J.; Blase, A.; *et al.* Urine TMPRSS2:ERG fusion transcript stratifies prostate cancer risk in men with elevated serum PSA. *Sci. Transl. Med*. **2011**, *3*, 94ra72.
250. Durand, X.; Moutereau, S.; Xylinas, E.; de la Taille, A. Progensa PCA3 test for prostate cancer. *Expert Rev. Mol. Diagn*. **2011**, *11*, 137–144.
251. Niinuma, T.; Suzuki, H.; Nojima, M.; Nosho, K.; Yamamoto, H.; Takamaru, H.; Yamamoto, E.; Maruyama, R.; Nobuoka, T.; Miyazaki, Y.; *et al.* Upregulation of miR-196a and HOTAIR drive malignant character in gastrointestinal stromal tumors. *Cancer Res*. **2012**, *72*, 1126–1136.
252. Geng, Y.J.; Xie, S.L.; Li, Q.; Ma, J.; Wang, G.Y. Large intervening non-coding RNA HOTAIR is associated with hepatocellular carcinoma progression. *J. Int. Med. Res*. **2011**, *39*, 2119–2128.
253. Ji, P.; Diederichs, S.; Wang, W.; Boing, S.; Metzger, R.; Schneider, P.M.; Tidow, N.; Brandt, B.; Buerger, H.; Bulk, E.; *et al.* MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* **2003**, *22*, 8031–8041.
254. Huang, K.C.; Rao, P.H.; Lau, C.C.; Heard, E.; Ng, S.K.; Brown, C.; Mok, S.C.; Berkowitz, R.S.; Ng, S.W. Relationship of XIST expression and responses of ovarian cancer to chemotherapy. *Mol. Cancer Ther*. **2002**, *1*, 769–776.
255. Tsai, M.C.; Spitale, R.C.; Chang, H.Y. Long intergenic noncoding RNAs: New links in cancer progression. *Cancer Res*. **2011**, *71*, 3–7.

Reprinted from *IJMS*. Cite as: Schwarzenbacher, D.; Balic, M.; Pichler, M. The Role of MicroRNAs in Breast Cancer Stem Cells. *Int. J. Mol. Sci.* **2013**, *14*, 14712-14723.

Review

The Role of MicroRNAs in Breast Cancer Stem Cells

Daniela Schwarzenbacher, Marija Balic and Martin Pichler *

Division of Clinical Oncology, Department of Medicine, Medical University of Graz, Auenbruggerplatz 15, 8036 Graz, Austria; E-Mails: daniela.schwarzenbacher@medunigraz.at (D.S.); marija.balic@medunigraz.at (M.B.)

* Author to whom correspondence should be addressed; E-Mail: martin.pichler@medunigraz.at; Tel.: +43-316-385-81320; Fax: +43-316-385-13355.

Received: 30 May 2013; in revised form: 25 June 2013 / Accepted: 2 July 2013 / Published: 15 July 2013

Abstract: The concept of the existence of a subset of cancer cells with stem cell-like properties, which are thought to play a significant role in tumor formation, metastasis, resistance to anticancer therapies and cancer recurrence, has gained tremendous attraction within the last decade. These cancer stem cells (CSCs) are relatively rare and have been described by different molecular markers and cellular features in different types of cancers. Ten years ago, a novel class of molecules, small non-protein-coding RNAs, was found to be involved in carcinogenesis. These small RNAs, which are called microRNAs (miRNAs), act as endogenous suppressors of gene expression that exert their effect by binding to the 3'-untranslated region (UTR) of large target messenger RNAs (mRNAs). MicroRNAs trigger either translational repression or mRNA cleavage of target mRNAs. Some studies have shown that putative breast cancer stem cells (BCSCs) exhibit a distinct miRNA expression profile compared to non-tumorigenic breast cancer cells. The deregulated miRNAs may contribute to carcinogenesis and self-renewal of BCSCs via several different pathways and can act either as oncomirs or as tumor suppressive miRNAs. It has also been demonstrated that certain miRNAs play an essential role in regulating the stem cell-like phenotype of BCSCs. Some miRNAs control clonal expansion or maintain the self-renewal and anti-apoptotic features of BCSCs. Others are targeting the specific mRNA of their target genes and thereby contribute to the formation and self-renewal process of BCSCs. Several miRNAs are involved in epithelial to mesenchymal transition, which is often implicated in the process of formation of CSCs. Other miRNAs were shown to be involved in the increased chemotherapeutic resistance of BCSCs. This review highlights the recent findings and crucial role of miRNAs in the maintenance, growth and

behavior of BCSCs, thus indicating the potential for novel diagnostic, prognostic and therapeutic miRNA-based strategies.

Keywords: microRNAs; breast cancer; tumor stem cells

1. Introduction

1.1. Breast Cancer and Breast Cancer Stem Cells

Breast cancer is the most frequently diagnosed cancer and the leading cause of cancer-related death among women worldwide [1]. According to the American Cancer Society, an estimated number of 232,340 new cases of breast cancer will be diagnosed in women and approximately 39,620 female breast cancer deaths are estimated in the United States for 2013. Thus, it is expected that breast cancer will account for 29% of all new cases of cancer in 2013 among women [2]. Hence, it is essential to gain a better understanding of the molecular mechanisms of breast cancer formation to ensure more efficient cancer treatments [3]. Since human breast tumors are very heterogeneous regarding time since diagnosis, histological pattern and clinical course, breast cancer can be classified into several subtypes based on distinct gene expression profiles [4]. In general, heterogeneity within and among several subtypes of cancers can arise in various ways. One common model to explain the usually observed heterogeneity of tumors is the cancer stem cell model [5]. According to the cancer stem cell hypothesis, tumors are hierarchically organized with cancer stem cells (CSCs) at the top [6] and the non-tumorigenic cell population forming the bulk of the tumor [7]. The term CSC indicates that only a subset of cancer cells in the tumor has self-renewal (asymmetric and symmetric division) capacity and the ability to produce all types of cancer cells within the tumor [6,8]. Targeting CSCs is of great interest as CSCs are considered to be more resistant to radiotherapy and chemotherapy, and are also thought to be responsible for the dissemination and growth of metastases [6].

Breast cancer stem cells (BCSCs) were originally described by Al-Hajj *et al.* in 2003. They isolated a tumorigenic subset of cancer cells from human breast tumors based on the expression of the surface markers $CD44^+$, $CD24^{-/low}$ and ESA^+ (CD is short for cluster of differentiation, ESA is short for epithelial specific antigen).This was the first evidence for the existence of CSCs in breast cancer and they were the first to show that only the minority of breast cancer cells with a $CD44^+$, $CD24^{-/low}$ and ESA^+ phenotype have the ability to form new tumors in NOD/SCID mice [9]. In 2007, Ginestier *et al.* indicated that high aldehyde dehydrogenase 1 (ALDH1) expression is also characteristic for BCSCs and therefore extended the BCSC phenotype on $CD44^+$, $CD24^{-/low}$, ESA^+ and alternatively $ALDH^+$ [10]. Dontu and colleagues developed an *in vitro* cell culture system under non-adherent conditions for human mammary epithelial cells. Under these conditions only cells with stem cell-like properties are able to survive. These cells can proliferate and build so called mammospheres, which are multicellular formations and are thought to contain high numbers of mammary stem cells as well as progenitor cells [11]. These mammosphere cultures are commonly used in experimental studies to enrich BCSCs. However, other studies indicate that the currently used markers for BCSCs remain controversial. Lehmann *et al.* discovered that some markers used to identify putative breast tumor initiating cells do

not correlate with *in vivo* tumorigenicity. Therefore it may be essential to determine other markers and/or factors that affect the increased tumorigenicity of BCSCs [12]. Another recently published study also revealed that these markers alone might not be sufficient to distinguish tumorigenic from non-tumorigenic cells. They demonstrated that tumor cells which are negative to the common CSC markers are also capable of inducing tumor growth *in vivo* [13].

1.2. Epithelial to Mesenchymal Transition in Cancer

Epithelial to mesenchymal transition (EMT) is an essential process during embryonic development in many species of mammals [14]. The transformation of epithelial to mesenchymal cells has also been associated with cancer progression, because the EMT program often becomes activated during cancer invasion and metastasis. This process is characterized by the loss of the epithelial marker E-cadherin, loss of cell–cell contact and cell polarity as well as an increased cell motility [15]. EMT has also been directly linked with the CSC phenotype. Induction of EMT in breast cancer cells leads to generation of cells with stem cell like properties [16].

1.3. Function and Biogenesis of miRNAs

It is now clear, that miRNAs together with other non-coding RNAs (long non-coding RNAs, small nucleolar RNAs and ultraconserved regions) contribute to carcinogenesis. Aberrantly expressed miRNAs are involved in initiation and progression of cancer. MiRNAs are small non-coding RNAs with a length of approximately 22 nucleotides (nt), which act as endogenous inhibitors of gene function. They modulate the expression of their target genes by either degrading their target mRNA or inhibiting their translation [17] through pairing of miRNA sequences to complementary bases on the target mRNA [18]. MiRNAs can function both as oncogenes and as tumor suppressors [19,20] and are considered as emerging potential candidates for improved cancer diagnosis, prognosis and therapy [21–24].

Biogenesis of miRNAs is a complex process. Most miRNA genes are transcribed by RNA polymerase II as long primary transcripts containing a stemloop structure. This pri-miRNA is cleaved by the RNase III endonuclease Drosha and the double-stranded RNA-binding domain (dsRBD) protein DGCR8/Pasha in the nucleus. The cleavage produces a ~70 nt hairpin precursor miRNA (pre-miRNA) with a 2-nt 3' overhang. The 3' overhang is recognized by Exportin-5, which transports the pre-miRNA into the cytoplasm. There, the pre-miRNA is cleaved by another RNase III endonuclease, Dicer. Dicer interacts with the dsRBD proteins TRBP/Loquacious and cleaving produces the mature ~22 nt miRNA: miRNA* duplex. The miRNA strand is usually incorporated to a RNA-induced silencing complex (RISC), a ribonuclein complex, while the miRNA* strand is typically degraded. When the miRNA is bound to RISC, the miRNA and its target mRNA can interact by base-pairing. The target mRNAs can then be cleaved and degraded or repressed in their translation [25].

In different breast cancer subtypes (basal, luminal cancers) miRNAs are differentially expressed and some miRNAs are associated with a specific ER, PR and Her2/neu status in human breast cancers [26]. MiRNAs can also function as potential targets of anticancer therapies. In breast cancer several miRNAs may possibly play a key role in cancer progression. Different studies have shown that silencing or overexpression of particular miRNAs can have an effect on the process of invasion and development of metastases in human breast cancers [27], showing a potential therapeutic application

of miRNAs in breast cancer. The aim of this review is to summarize the involvement of different miRNAs in the formation and regulation of human BCSCs. Table 1 gives an overview of the roles of different miRNAs in BCSCs which are described in this review.

Table 1. Roles of different miRNAs in BCSCs.

miRNA	Roles in BCSCs
let-7 family	downregulated in BCSCs targets RAS and HMGA2 acts as tumor suppressor Lin28 blocks *let-7* biogenesis and promotes tumorigenic activity in breast cancer influences mammosphere formation and proliferation *in vitro* affects tumor formation ability and metastatic potential *in vivo* reduced *let-7* expression inhibits differentiation, maintains proliferation and promotes EMT
miR-200 family	downregulated in BCSCs targets Bmi-1 and Suz12 regulation of EMT relevant for stem cell functions in cancer cells (self-renewal, clonal expansion, differentiation) *in vitro* induces stem-like properties
miR-30 family	downregulated in BCSCs targets Ubc9, ITGB3 and AVEN influences self-renewal capacity and anti-apoptotic features important for modulation of the stem-like properties of BCSCs regulates non-attachment growth of mammospheres and mammosphere formation ability controls genes involved in apoptosis and proliferation in BCSCs
miR-128	downregulated in BCSCs targets Bmi-1 and ABCC5 link to chemotherapeutical resistance and survival rates of breast cancer patients influences number and size of mammospheres *in vitro* reduced tumor growth and induced apoptosis *in vivo*
miR-34c	downregulated in BCSCs targets Notch4 influences self-renewal and EMT acts on mammosphere formation *in vitro* is epigenetically regulated via methylation controls migration of tumor cells
miR-16	downregulated in BCSCs targets Wip1 influences number and size of mammospheres and cell proliferation responsible for sensitivity to chemotherapeutic drug doxorubicin
miR-181	upregulated in BCSCs targets ATM TGF-β induces mammosphere formation by upregulation of *miR-181*
miR-495	upregulated in BCSCs targets REDD1 leads to downregulation of E-cadherin promotes colony formation leads to increased tumor formation *in vivo* responsible for maintaining a stem-cell line phenotype

2. Particular miRNAs and Their Role in Tumor-Initiating BCSCs

2.1. miRNAs Down-Regulated in BCSCs

2.1.1. *let-7* Family

Yu and colleagues were the first to investigate the expression of miRNAs in BCSCs in 2007. They compared the miRNA expression profile in self-renewing BCSCs and differentiated cells from breast cancer cell lines as well as in samples from primary breast tumors. They enriched BCSCs of a human breast cancer cell line (SKBR3) by passaging them in NOD/SCID mice treated with chemotherapeutical agents. The tumors contained a high percentage of $CD44^{+}CD24^{-/low}$ cells and showed high mammosphere formation ability *in vitro*. The miRNA *let-7* was found to be the most consistently down-regulated miRNA in tumor-initiating cells (SK-3rd) compared to the non-self-renewing population of cancer cells. *let-7* expression increased when the cells differentiated to non-tumorigenic cancer cells. The *let-7* family functions as a well-known tumor suppressor and targets the oncogenes rat sarcoma (RAS) and high mobility group AT-hook 2 (HMGA2). The HMGA2 gene is involved in mesenchymal cell differentiation and tumor formation. Lentiviral-mediated re-expression of *let-7* resulted in reduced mammospheres formation, proliferation and a reduced number of undifferentiated stem-like cells *in vitro*. *let-7* expression also inhibited the tumor formation ability in NOD/SCID mice *in vivo* and *let-7* expressing tumors had less metastatic potential. Therefore, *let-7* is apparently responsible for the regulation of multiple stem cell-like properties of BCSCs, because reduced *let-7* expression inhibits differentiation and maintains proliferation [28]. As *let-7* is a common tumor suppressor and has anti-proliferative properties, it can regulate cell differentiation and apoptotic pathways. Its down-regulation has been reported in several cancers and reconstitution of regular *let-7* expression has been shown to inhibit cancer growth [29–31]. These findings suggest *let-7* as a potential molecular marker for BCSCs with a potential as therapeutical target in anti-cancer therapy [32]. In this context, the Lin28 protein is a RNA-binding protein which regulates *let-7* family members and expression of Lin28 blocks the biogenesis of *let-7* [33]. One recently published study indicates that suppression of *let-7* through Lin28 promotes tumorigenicity in breast cancer cells [34]. Inflammatory cytokines can lead to the induction of EMT. Guo and colleagues showed that inflammatory cytokines can trigger signal transducer and activator of transcription factor 3 (Stat3) which promotes Lin28 transcription. As a consequence, this process results in repression of *let-7* expression and up-regulation of the *let-7* target HMGA2. As HMGA2 is involved in EMT, this event leads to increased levels of mesenchymal markers. These findings suggest that the inflammation-induced and Stat3 mediated Lin28-let-7-HMGA2 signaling pathway might be involved in regulation of self-renewal and differentiation in CSCs [35].

2.1.2. *miR-200* Family

The *miR-200* family consists of five members of miRNAs: *miR-200a*, *miR-200b*, *miR-200c*, *miR-141* and *miR-429*. The family can be divided into genetically different subfamilies (gene clusters) according to their location at two different chromosomes: the *miR-200b*/*miR-200a*/*miR-429* gene cluster on chromosome 1 and the *miR-200c*/*miR-141* gene cluster on chromosome 12 [36]. Several

recent studies have associated *miR-200* family members and their target mRNAs with establishment, maintenance and regulation of the BCSC phenotype. One of the first studies that showed an involvement of the *miR-200* family in BCSCs came from Shimono *et al.* in 2009. Comparing the miRNA expression profile between fluorescence-activated cell sorted $CD44^{+}CD24^{-/low}$ lineage BCSCs and the remaining non-tumorigenic human breast cancer cells, they found 37 miRNAs differentially expressed between non-tumorigenic cancer cells and BCSCs in eleven human breast cancer samples. They showed that three clusters of miRNAs (*miRNA-200c-141*, *miR-200b-200a-429* and *miR-183-96-182*) were consistently down-regulated in BCSCs, in normal mammary stem cells and in embryonal carcinoma cells. This finding suggests that the down-regulation of these miRNAs may be relevant for stem cell functions in cancer cells, such as self-renewal or differentiation. Downregulation of *miR-200c* in BCSCs suppressed the expression of polycomb ring finger oncogene (B lymphoma Mo-MLV insertion region 1 homolog, Bmi-1), which is a regulator of stem cell self-renewal. *miR-200c* inhibited the clonal expansion of BCSCs *in vitro*. Interestingly, *miR-200c* repressed the ability of normal mammary stem cells to generate mammary ducts and also inhibits the tumor-formation capacity of BCSCs *in vivo*. These results indicate that down-regulation of *miR-200c* might be a molecular link between CSCs and normal stem cells [37]. Consistent to that the *miR-200* family was shown to be inhibited during BCSC formation in an inducible CSC model. One of the down-regulated *miR-200* family members in BCSCs was *miR-200b* and inhibition of *miR-200b* increased the formation of BCSCs. Down-regulation of *miR-200b* resulted in increased Suz12 expression (a subunit of a polycomb repressor complex, PRC2), which led to repression of E-cadherin [38]. This inhibition of E-cadherin through miRNA is sufficient to cause EMT [39]. Overexpression of *miR-200b* or inhibition of Suz12 significantly reduced BCSC growth. In tumors of breast cancer patients, *miR-200b* and Suz12 expression were inversely correlated. Apparently the *miR-200b*-Suz12-cadherin pathway is an important pathway to induce and sustain growth of BCSCs and the invasion and migration abilities of BCSCs [38]. Another recently published study confirmed the role of the *miR-200* family in BCSCs by showing that the spontaneous conversion of immortalized human mammary epithelial cells to a stem-like phenotype with mesenchymal and less differentiated properties was accompanied by loss of *miR-200* expression. In mammospheres, *miR-200a*, *miR-200b* and *miR-200c* were described as down-regulated. Expression of *miR-200* was shown to be epigenetically regulated by histone-modifications and DNA promoter methylation. Also, in samples of pleural or ascites effusions of breast cancer patients the *miR-200* family members were consistently down-regulated in $CD44^{+}CD24^{-/low}$ putative BCSCs. Re-expression of *miR-200* in these stem-like cells led to a partial reprogramming to a non-stem like phenotype, and the cells also did undergo mesenchymal to epithelial transition (MET). These data indicate that the *miR-200* family is functionally inducing the stem-like properties [40].

2.1.3. *miR-30* Family

Similar to the *let-7* family, Yu and colleagues also demonstrated the down-regulation of *miR-30*, particularly *miR-30e*, in tumor initiating BCSCs (in mammospheres SK-3rd as well as in primary BCSCs obtained from breast cancer patients). In accordance to the down-regulation of *miR-30e*, the protein levels of two direct target genes of *miR-30e*, ubiquitin-conjugating enzyme 9 (Ubc9) and integrin b3 (ITGB3), were significantly up-regulated. When *miR-30e* was constitutively expressed in

BCSCs, their self-renewal capacity was impaired. This inhibition occurred through decreased Ubc9 levels and induction of apoptosis via silencing of ITGB3. Blocking of *miR-30e* in differentiated breast cancer cells on the other hand led to regeneration of their self-renewal capacity. Overexpression of *miR-30e* in NOD/SCID mice reduced tumorigenesis and lung metastases, while blocking of *miR-30e* expression enhanced tumor formation and metastases. These results indicate that reduction of *miR-30* expression is responsible for maintaining the self-renewal and anti-apoptotic features of BCSCs. *miR-30* can therefore be considered as an important miRNA for modulation of the stem-like properties of BCSCs [41]. Down-regulation of *miR-30* family members in non-adherent mammospheres compared to breast cancer cells under adherent conditions was recently confirmed in an independent study. BCSCs growth under non-attachment conditions displayed a different miRNA expression pattern compared to adherent parental cells and members of the *miR-30* family were found to be the most consistently down-regulated miRNAs in putative BCSCs. Especially *miR-30a* was found to regulate the non-attachment growth of mammospheres. Overexpression of *miR-30a* significantly reduced the mammosphere formation ability, while inhibition of *miR-30a* dramatically increased the number of mammospheres in the human breast cancer cell line MCF-7.These results confirm the relevance of this miRNA in sustaining the growth of BCSCs under non-attachment conditions. Also down-regulation of potential *miR-30a* targets after overexpressing *miR-30a* was shown. Among the potential targets, the anti-apoptotic protein AVEN was one of the most significantly down-regulated genes after overexpression of *miR-30*.This study confirms that *miR-30* family members can control expression of genes involved in apoptosis and proliferation in BCSCs [42].

2.1.4. *miR-128*

The level of *miR-128* was shown to be significantly reduced in mammospheric BCSCs in two breast cancer cell lines (SK-3rd and MCF-7) and in BCSCs isolated from primary breast cancer patients. This reduction increased the protein levels of the polycomb ring finger oncogene Bmi-1 and ATP-binding cassette sub-family C member 5 (ABCC5), which are targets of *miR-128* [43]. Tumor initiating cells with stem cell-like features were shown to be more resistant to chemotherapeutical agents and radiotherapy than more differentiated tumor cells [9]. Ectopic expression of *miR-128* decreased Bmi-1 and ABCC5 levels in BCSCs and led to an enhanced pro-apoptotic and DNA-damaging effect when treated with the chemotherapeutical agent doxorubicin. This observation indicates a possible therapeutic potential of this miRNA. Furthermore, a reduction of *miR-128* in breast tumor tissues was linked with chemotherapeutic resistance and poor survival rates of breast cancer patients. Consequently the reduced levels of *miR-128* in BCSCs are likely to induce increased chemotherapeutic resistance [43]. In another study ectopic expression of *miR-128* led to a decreased number and size of mammospheres in an *in vitro* cell culture model, whereas *miR-128* depletion caused an increase of mammosphere growth. The *in vivo* tumor-initiating potential was also evaluated and it has been shown that overexpression of *miR-128-2* repressed the ability to form tumors in mice. Forced expression of *miR-128* reduced tumor growth *in vivo* and induced apoptosis [44].

2.1.5. *miR-34c*

MiR-34c has been identified as a putative tumor suppressor and has been reported to inhibit invasion, proliferation and to promote apoptosis. Reduced expression of *miR-34c* was revealed in two human breast cancer cell lines (MCF-7 and SK-3rd) enriched for BCSCs. Down-regulation of *miR-34c* apparently occurred via hypermethylation of the promoter region of BCSCs and resulted in increased self-renewal and epithelial-mesenchymal transition of these cells. Ectopic expression of this miRNA inhibited EMT and reduced mammosphere formation and self-renewal potential. It also led to silencing of its target gene Neurogenic Locus Notch Homolog Protein 4 (Notch4) and suppressed the migration of tumor cells. This study proposed *miR-34c* as a possible target for BCSCs, as this epigenetically regulated miRNA apparently controls self-renewal and EMT in these tumor initiating cells [45].

2.1.6. *miR-16*

A decreased level of *miR-16* has been shown in mammospheres derived from mammary tumors in mice compared to the whole tumor cell population. The oncogene wild-type p53-induced phosphatase 1 (Wip1) is apparently regulated by *miR-16* and protein levels of Wip1 were consequently increased in these mammospheres. Overexpression of *miR-16* in the human breast cancer cell line MCF-7 as well as inhibition of Wip1 decreased the number and size of mammospheres. When *miR-16* was overexpressed in the human breast cancer cell line MCF-7 it suppressed cell proliferation and led to an increased sensitivity to the chemotherapeutic drug doxorubicin. These findings indicate that *miR-16* is another miRNA that might be responsible for regulation of the proliferation and differentiation of mammary CSCs [46].

2.2. miRNAs Up-Regulated in BCSCs

2.2.1. *miR-181* Family

In three human breast cancer cell lines (BT474, MDA361 and MCF7) levels of *miR-181* family members were reported to be increased in tumor initiating mammospheres compared to non-tumorigenic parental cells. Transforming growth factor-β (TGF-β) seemed to induce sphere formation by up-regulation of *miR-181* at the post-transcriptional level. A potential target of *miR-181* is the serine/threonine kinase Ataxia telangiectasia mutated (ATM) which acts as a tumor suppressor. ATM was reduced in mammospheres and after treatment with TGF-β. This study suggests that the TGF-β pathway and the *miR-181* family interacts and plays a role in regulating the BCSC phenotype [47].

2.2.2. *miR-495*

Hwang-Verslues and colleagues isolated a novel highly tumorigenic subpopulation of BCSCs based on the surface markers $PROCR^+/ESA^+$ (PROCR is short for protein C receptor). In this BCSC subpopulation and also in the more commonly used $CD44^+CD24^{-/low}$ subpopulation, *miR-495* was highly up-regulated. As *miR-495* was found to be up-regulated in two distinct BCSC subpopulations, this mechanism might be important for maintaining stem cell-like features. This up-regulation of *miR-495* is regulated by the transcription factors E12 and E47. Overexpression of *miR-495 in vitro* promoted colony formation. *miR-495* overexpressing cells in mice led to significantly higher tumor

formation *in vivo*. These results indicate that ectopic expression of *miR-495* in human breast cancer cells increases tumorigenesis *in vivo* and enhances colony formation *in vitro*. E-cadherin expression, a marker considered surrogate for EMT, was down-regulated by overexpression of *miR-495*. Decreased E-cadherin expression was responsible for promoting cell invasion. *miR-495* also targets REDD1 (short for regulated in development and DNA damage responses) which is a factor for enhanced hypoxia resistant cell proliferation. Summarizing, these findings suggest an up-regulation of *miR-495* by E12 and E47 which in turn contributes to down-regulation of E-cadherin and REDD1, finally resulting in maintaining a stem cell-like phenotype in breast cancer [48].

3. Conclusions

Understanding the role of miRNAs in the biology of CSCs can provide promising advances for cancer treatment and might be helpful to improve cancer diagnosis. As miRNAs are post-transcriptional regulators of gene expression, they also play important roles in carcinogenesis. Several independent studies that are reviewed here have shown a dysregulation of several different miRNAs in BCSCs. Anticancer-therapy with miRNAs could eliminate the CSC self-renewal capacity and their anti-apoptotic features which can improve the development of resistance against current cancer treatment. For this reason, future research should address the therapeutic potential of miRNAs to prevent cancer progression, relapse and formation of metastases by eliminating CSCs.

Acknowledgements

This work was supported by funds of the Oesterreichische Nationalbank (Anniversary Fund, project number: 14869) and by the Start foundation of the Medical University of Graz (both to Martin Pichler).

Conflict of Interest

The authors declare no conflict of interest.

References

1. Jemal, A.; Bray, F.; Center, M.; Ferlay, J.; Ward, E.; Forman, D. Global cancer statistics. *CA: Cancer J. Clin.* **2011**, *61*, 69–90.
2. Siegel, R.; Naishadham, D.; Jemal, A. Cancer statistics, 2013. *CA: Cancer J. Clin.* **2013**, *63*, 11–30.
3. Liu, H. MicroRNAs in breast cancer initiation and progression. *Cell. Mol. Life Sci.* **2012**, *69*, 3587–3599.
4. Perou, C.M.; Sørlie, T.; Eisen, M.B.; van de Rijn, M.; Jeffrey, S.S.; Rees, C.A.; Pollack, J.R.; Ross, D.T.; Johnsen, H.; Akslen, L.A.; *et al*. Molecular portraits of human breast tumours. *Nature* **2000**, *406*, 747–752.
5. Magee, J.; Piskounova, E.; Morrison, S.J. Cancer stem cells: Impact, heterogeneity, and uncertainty. *Cancer Cell* **2012**, *21*, 283–296.
6. Visvader, J.E.; Lindeman, G. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat. Rev. Cancer.* **2008**, *8*, 755–768.

7. Vincent, A.; van Seuningen, I. On the epigenetic origin of cancer stem cells. *Biochim. Biophys. Acta* **2012**, *1826*, 83–88.
8. Wicha, M.; Liu, S.; Dontu, G. Cancer stem cells: An old idea—A paradigm shift. *Cancer Res.* **2006**, *66*, 1883–1890.
9. Al-Hajj, M. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3983–3988.
10. Ginestier, C.; Hur, M.; Charafe-Jauffret, E. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell* **2007**, *1*, 555–567.
11. Dontu, G.; Abdallah, W. *In vitro* propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev.* **2003**, *17*, 1253–1270.
12. Lehmann, C.; Jobs, G.; Thomas, M.; Burtscher, H.; Kubbies, M. Established breast cancer stem cell markers do not correlate with *in vivo* tumorigenicity of tumor-initiating cells. *Int. J. Oncol.* **2012**, *41*, 1932–1942.
13. Huang, S.-D.; Yuan, Y.; Tang, H.; Liu, X.-H.; Fu, C.-G.; Cheng, H.-Z.; Bi, J.-W.; Yu, Y.-W.; Gong, D.-J.; Zhang, W.; *et al*. Tumor cells positive and negative for the common cancer stem cell markers are capable of initiating tumor growth and generating both progenies. *PLoS One* **2013**, doi:10.1371/journal.pone.0054579.
14. Kiesslich, T.; Berr, F.; Alinger, B.; Kemmerling, R.; Pichler, M.; Ocker, M.; Neureiter, D. Current status of therapeutic targeting of developmental signalling pathways in oncology. *Curr. Pharm. Biotechnol.* **2012**, *13*, 2184–2220.
15. Thiery, J.P. Epithelial-mesenchymal transitions in tumour progression. *Nat. Rev. Cancer* **2002**, *2*, 442–454.
16. Mani, S.A.; Guo, W.; Liao, M.-J.; Eaton, E.N.; Ayyanan, A.; Zhou, A.Y.; Brooks, M.; Reinhard, F.; Zhang, C.C.; Shipitsin, M.; *et al*. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **2008**, *133*, 704–715.
17. Nana-Sinkam, S.P.; Croce, C.M. Clinical applications for microRNAs in cancer. *Clin. Pharmacol. Ther.* **2013**, *93*, 98–104.
18. Munker, R.; Calin, G.A. MicroRNA profiling in cancer. *Clin. Sci.* **2011**, *121*, 141–158.
19. Calin, G.A.; Croce, C.M. MicroRNA signatures in human cancers. *Nat. Rev. Cancer* **2006**, *6*, 857–866.
20. Al-Ali, B.M.; Ress, A.L.; Gerger, A.; Pichler, M. MicroRNAs in renal cell carcinoma: implications for pathogenesis, diagnosis, prognosis and therapy. *Anticancer Res.* **2012**, *32*, 3727–3732.
21. Pichler, M.; Winter, E.; Stotz, M.; Eberhard, K.; Samonigg, H.; Lax, S.; Hoefler, G. Down-regulation of KRAS-interacting miRNA-143 predicts poor prognosis but not response to EGFR-targeted agents in colorectal cancer. *Br. J. Cancer* **2012**, *106*, 1826–1832.
22. Bach, D.; Fuereder, J.; Karbiener, M.; Scheideler, M.; Ress, A.L.; Neureiter, D.; Kemmerling, R.; Dietze, O.; Wiederstein, M.; Berr, F.; *et al*. Comprehensive analysis of alterations in the miRNome in response to photodynamic treatment. *J. Photochem. Photobiol. B* **2013**, *120*, 74–81.
23. Van Roosbroeck, K.; Pollet, J.; Calin, G. miRNAs and long noncoding RNAs as biomarkers in human diseases. *Expert Rev. Mol. Diagn.* **2013**, *13*, 183–204.
24. Calin, G.A.; Konopleva, M. Small gene, big number, many effects. *Blood* **2012**, *120*, 240–241.
25. Bushati, N.; Cohen, S.M. microRNA functions. *Annu. Rev. Cell Dev. Biol.* **2007**, *23*, 175–205.

26. Iorio, M.V.; Croce, C.M. MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO Mol. Med.* **2012**, *4*, 143–159.
27. Ahmad, A. Pathways to breast cancer recurrence. *ISRN Oncol.* **2013**, *2013*, doi:10.1155/2013/290568.
28. Yu, F.; Yao, H.; Zhu, P.; Zhang, X.; Pan, Q.; Gong, C.; Huang, Y.; Hu, X.; Su, F.; Lieberman, J.; *et al*. *Let-7* regulates self renewal and tumorigenicity of breast cancer cells. *Cell* **2007**, *131*, 1109–1123.
29. Johnson, C.D.; Esquela-Kerscher, A.; Stefani, G.; Byrom, M.; Kelnar, K.; Ovcharenko, D.; Wilson, M.; Wang, X.; Shelton, J.; Shingara, J.; *et al*. The *let-7* microRNA represses cell proliferation pathways in human cells. *Cancer Res.* **2007**, *67*, 7713–7722.
30. Kumar, M.S.; Erkeland, S.J.; Pester, R.E.; Chen, C.Y.; Ebert, M.S.; Sharp, P.A.; Jacks, T. Suppression of non-small cell lung tumor development by the *let-7* microRNA family. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3903–3908.
31. Wang, Y.; Hu, X.; Greshock, J.; Shen, L.; Yang, X.; Shao, Z.; Liang, S.; Tanyi, J.L.; Sood, A.K.; Zhang, L. Genomic DNA copy-number alterations of the *let-7* family in human cancers. *PLoS One* **2012**, doi:10.1371/journal.pone.0044399.
32. Barh, D.; Malhotra, R.; Ravi, B.; Sindhurani, P. Micro rna *let-7*: An emerging next-generation cancer therapeutic. *Curr. Oncol.* **2010**, *17*, 70–80.
33. Newman, M.; Thomson, J.; Hammond, S. Lin-28 interaction with the *Let-7* precursor loop mediates regulated microRNA processing. *RNA* **2008**, *14*, 1539–1549.
34. Sakurai, M.; Miki, Y.; Masuda, M.; Hata, S.; Shibahara, Y.; Hirakawa, H.; Suzuki, T.; Sasano, H. LIN28: A regulator of tumor-suppressing activity of *let-7* microRNA in human breast cancer. *J. Steroid Biochem. Mol. Biol.* **2012**, *131*, 101–106.
35. Guo, L.; Chen, C.; Shi, M.; Wang, F.; Chen, X.; Diao, D.; Hu, M.; Yu, M.; Qian, L.; Guo, N. Stat3-coordinated Lin-28-*let-7*-HMGA2 and miR-200-ZEB1 circuits initiate and maintain oncostatin M-driven epithelial-mesenchymal transition. *Oncogene* **2013**, doi:10.1038/onc.2012.573.
36. Park, S.-M.; Gaur, A.B.; Lengyel, E.; Peter, M.E. The *miR-200* family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev.* **2008**, *22*, 894–907.
37. Shimono, Y.; Ugalde, M.Z.; Cho, R.W.; Lobo, N.; Dalerba, P.; Qian, D.; Diehn, M.; Liu, H.; Panula, S.P.; Chiao, E.; *et al*. Down-regulation of *miRNA-200c* links breast cancer stem cells with normal stem cells. *Cell* **2009**, *138*, 592–603.
38. Iliopoulos, D.; Lindahl-Allen, M. Loss of *miR-200* inhibition of Suz12 leads to polycomb-mediated repression required for the formation and maintenance of cancer stem cells. *Mol. Cell* **2010**, *39*, 761–772.
39. Onder, T.T.; Gupta, P.B.; Mani, S.A.; Yang, J.; Lander, E.S.; Weinberg, R.A. Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. *Cancer Res.* **2008**, *68*, 3645–3654.
40. Lim, Y.; Wright, J.; Attema, J.; Gregory, P.; Bert, A.; Smith, E.; Thomas, D.; Drew, P.; Khew-Goodall, Y.; Goodall, G. Epigenetic modulation of the *miR-200* family is associated with transition to a breast cancer stem cell-like state. *J. Cell Sci.* **2013**, doi:10.1242/jcs.122275.

41. Yu, F.; Deng, H.; Yao, H.; Liu, Q.; Su, F.; Song, E. *MiR-30* reduction maintains self-renewal and inhibits apoptosis in breast tumor-initiating cells. *Oncogene* **2010**, *29*, 4194–4204.
42. Ouzounova, M.; Vuong, T.; Ancey, P.-B.; Ferrand, M.; Durand, G.; Le-Calvez Kelm, F.; Croce, C.; Matar, C.; Herceg, Z.; Hernandez-Vargas, H. MicroRNA *miR-30* family regulates non-attachment growth of breast cancer cells. *BMC Genomics* **2013**, doi:10.1186/1471-2164-14-139.
43. Zhu, Y.; Yu, F.; Jiao, Y.; Feng, J.; Tang, W.; Yao, H.; Gong, C.; Chen, J.; Su, F.; Zhang, Y.; *et al*. Reduced *miR-128* in breast tumor-initiating cells induces chemotherapeutic resistance via Bmi-1 and ABCC5. *Clin. Cancer Res.* **2011**, *17*, 7105–7115.
44. Qian, P.; Banerjee, A.; Wu, Z.-S.; Zhang, X.; Wang, H.; Pandey, V.; Zhang, W.-J.; Lv, X.-F.; Tan, S.; Lobie, P.E.; *et al*. Loss of SNAIL regulated *miR-128*-2 on chromosome 3p22.3 targets multiple stem cell factors to promote transformation of mammary epithelial cells. *Cancer Res.* **2012**, *72*, 6036–6050.
45. Yu, F.; Jiao, Y.; Zhu, Y.; Wang, Y.; Zhu, J.; Cui, X.; Liu, Y.; He, Y.; Park, E.-Y.; Zhang, H.; *et al*. *MicroRNA 34c* gene down-regulation via DNA methylation promotes self-renewal and epithelial-mesenchymal transition in breast tumor-initiating cells. *J. Biol. Chem.* **2012**, *287*, 465–473.
46. Zhang, X.; Wan, G.; Mlotshwa, S.; Vance, V.; Berger, F.; Chen, H.; Lu, X. Oncogenic Wip1 phosphatase is inhibited by *miR-16* in the DNA damage signaling pathway. *Cancer Res.* **2010**, *70*, 7176–7186.
47. Wang, Y.; Yu, Y.; Tsuyada, A.; Ren, X.; Wu, X.; Stubblefield, K.; Rankin-Gee, E.; Wang, S. Transforming growth factor β regulates the sphere-initiating stem cell-like feature in breast cancer through miRNA-181 and ATM. *Oncogene* **2011**, *30*, 1470–1480.
48. Hwang-Verslues, W.W.; Chang, P.-H.; Wei, P.-C.; Yang, C.-Y.; Huang, C.-K.; Kuo, W.-H.; Shew, J.-Y.; Chang, K.-J.; Lee, E.Y.-H.P.; Lee, W.-H. *miR-495* is upregulated by E12/E47 in breast cancer stem cells, and promotes oncogenesis and hypoxia resistance via downregulation of E-cadherin and REDD1. *Oncogene* **2011**, *30*, 2463–2474.

Reprinted from *IJMS*. Cite as: Hannafon, B.N.; Ding, W.-Q. Intercellular Communication by Exosome-Derived microRNAs in Cancer. *Int. J. Mol. Sci.* **2013**, *14*, 17347-17377.

Review

Intercellular Communication by Exosome-Derived microRNAs in Cancer

Bethany N. Hannafon * and Wei-Qun Ding *

Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA; E-Mails: bethany-hannafon@ouhsc.edu (B.N.H.); weiqun-ding@ouhsc.edu (W-Q.D.); Tel.: +1-405-271-3828 (B.N.H.); +1-405-271-1605 (W.-Q.D.); Fax: +1-405-271-3910 (W.-Q.D.)

Received: 2 May 2013; in revised form: 14 June 2013 / Accepted: 17 June 2013 / Published: 9 July 2013

Abstract: The development of human cancers is a multistep process in which normal cells acquire characteristics that ultimately lead to their conversion into cancer cells. Many obstacles must be overcome for this process to occur; of these obstacles, is the ability to survive an inhospitable microenvironment. It is recognized that the intercommunication between tumor cells and their surrounding microenvironment is essential to overcoming this obstacle and for the tumor to progress, metastasize and establish itself at distant sites. Exosomes are membrane-derived vesicles that have recently been recognized as important mediators of intercellular communication, as they carry lipids, proteins, mRNAs and microRNAs that can be transferred to a recipient cell via fusion of the exosome with the target cell membrane. In the context of cancer cells, this process entails the transfer of cancer-promoting cellular contents to surrounding cells within the tumor microenvironment or into the circulation to act at distant sites, thereby enabling cancer progression. In this process, the transfer of exosomal microRNAs to a recipient cell where they can regulate target gene expression is of particular interest, both in understanding the basic biology of cancer progression and for the development of therapeutic approaches. This review discusses the exosome-mediated intercellular communication via microRNAs within the tumor microenvironment in human cancers, with a particular focus on breast cancer exosomes.

Keywords: exosomes; microvesicles; microRNAs; cancer; breast cancer; microenvironment

1. Introduction

1.1. Historical Perspective

Most cell types are known to continually release soluble factors and to exfoliate membrane derived vesicles into the extracellular space, including mast cells, dendritic cells, B-lymphocytes, platelets, neurons, adipocytes, endothelial cells and epithelial cells [1]. These membrane-derived vesicles are generally discriminated by size with two major classes; the larger class is called microvesicles (200–1000 nm) and the smaller class of nanometer size vesicles is called exosomes (30–200 nm). It is important to note that these are distinctly different from apoptotic bodies (0.5–3 μm), which are released from cells undergoing apoptosis or mechanical stress and are of a different cellular origin and molecular composition (see, for review, [2]). Exosomes were first observed three decades ago in differentiating reticulocytes. It was shown that during reticulocyte maturation, the transferrin receptor and many membrane-associated proteins were shed in small membrane vesicles via an unknown secretory process [3,4]. This process was considered as a way for cells to eliminate unwanted proteins and molecules, with exosomes functioning as cellular garbage disposals. However, in recent years, exosomes have emerged as important mediators of cellular communication that are involved in both normal physiological processes, such as lactation [4], immune response [5] and neuronal function [4], and also in the development and progression of diseases, such as liver disease [6], neurodegenerative diseases [7] and cancer. Exosomes have been identified in most bodily fluids, including urine and amniotic fluid [8], serum [9], saliva [10], breast-milk [5], cerebrospinal fluid [11], and nasal secretions [12]. Importantly, cancer cells have been shown to secrete exosomes in greater amounts than normal cells [13], indicating their potential use as biomarkers for diagnosis of disease.

1.2. Exosome Biogenesis and Secretion

Although the detailed mechanism for exosome biogenesis remains incompletely defined, current models suggest that exosomes are formed within the endocytic pathway and released from the plasma membrane via multivesicular bodies (MVBs) [14]. MVBs are formed during the maturation of early endosomes into late endosomes with the accumulation of intraluminal vesicles (ILVs), which correspond to exosomes [15]. Upon maturation, MVBs are either destined for fusion with the lysosome, where their contents will undergo lysosomal degradation, or with the plasma membrane, where their contents are released into the extracellular space. How these vesicles are sorted for either destination is not well understood. However, central players in this process are thought to be the endosomal sorting complexes required for transport (ESCRT). The ESCRT machinery is made up of five distinct complexes (ESCRT-0, -I, -II, -III and Vsp4; reviewed in [16]). This process is best characterized in yeast [17], where it was shown that the ESCRT machinery is responsible for generating MVB vesicles by initiating the budding of the endosome away from the cytoplasm and scission of the membrane to release of the mature MVB vesicles into the lumen of the lysosome (reviewed in [18]). ESCRT-0, -I and -II complexes recognize ubiquitinated proteins in the endosomal membrane [19], whereas the ESCRT-III complex may be responsible for membrane budding and vesicle scission [20]. However, there are additional pathways for MVB formation, sorting and exosome secretion. Most recently, an ESCRT-independent mechanism was described involving the

sphingolipid, ceramide. Ceramide is generated during cellular stress and apoptosis either by *de novo* synthesis or by sphingomyelinase, the enzyme that hydrolyzes sphingomyelin into ceramide. Ceramide contributes to cellular signaling by playing a role in membrane microdomain coalescence, receptor clustering, vesicle formation, membrane fusion/fission and vesicular trafficking [21]. Additionally, ceramide is enriched in exosome membranes (see, for review, [21,22]). Further validating the ESCRT-independent process, inhibition of neutral sphingomyelinase (nSMase) decreased exosome formation and release, whereas depletion of different ESCRT components did not reduce exosome secretion or the formation of MVBs [23]. Interestingly, exosomes produced by the ESCRT-independent/ sphingomyelinase pathway are enriched in tetraspanins, which are transmembrane proteins that may also be involved in endosomal sorting pathways [24]. Based on these observations, ESCRT-dependent sorting mechanisms may target proteins loaded into ILVs for lysosomal degradation, whereas ESCRT-independent sorting mechanisms may target ILVs for secretion. However, it is likely that this process is much more complex and may depend on the cell type, cargo or other stimulus. Furthermore, the signals that may control the switch between the two mechanisms remain unknown.

Exosome secretion is not considered a random event, but rather, a highly controlled process. Control of exosome secretion or "exocytosis", although largely still under investigation, is thought to be coordinated through the transport and fusion of MVBs with the plasma membrane by the microtubule and actin cytoskeleton, t- and v-SNAREs and Rab GTPases [25]. Rab GTPases are ubiquitously expressed proteins that are responsible for the coordination of various vesicle trafficking events [26]. For example, overexpression of Rab11 has been shown to stimulate exocytosis [27], and Rab27a and Rab27b control different steps of the exosome secretion pathway [28]. Exosome secretion can be initiated by many different mechanical, chemical and biological stimuli. For example, DNA damage due to γ-irradiation activates the p53 tumor suppressor gene and induces the release of exosomes [29]. When breast cancer cells were cultured under hypoxic conditions (1% to 0.1% O_2), their exosome secretion was significantly enhanced, whereas siRNA knockdown of HIF-1α prior to hypoxic exposure prevented this increase in exosome secretion [30]. Heparanase, an enzyme that cleaves heparan sulfate, which is upregulated in many cancers and is associated with enhanced tumor growth, was shown to dramatically increase exosome secretion in several human cancer cell lines [31]. Interestingly, heparanase could also alter the protein cargo carried by these exosomes, with increases in levels of syndecan-1, VEGF and HGF. Mechanical changes can also affect exosome secretion. For example, it was recently shown that detachment of adherent breast cancer cells from various surfaces could induce rapid exosome secretion [32]. Treatment with chemicals, such as calcium ionophores or statins to reduce membrane cholesterol levels and cholesterol biosynthesis, can also stimulate the release of exosomes in many cell types [23,33,34]. Finally, low pH level, a common hallmark of malignancy, in melanoma cells has been shown to increase exosome release and uptake, and pre-treatment with a proton pump inhibitor led to an inhibition of exosome uptake [35].

In order for secreted exosomes to exert any biological function, they must be absorbed by and deliver their contents to a recipient target cell. However, the specific targeting of exosomes to target cells and how this process unfolds in normal physiology or in the diseased state is not well understood. This process must critically depend on the specific adhesion molecules, integrins and antigenic factors expressed on the exosome, as well as the receptors or other docking molecules found on the surface of target cells. Presumably, any cell capable of endocytosis or phagocytosis may participate in the uptake

of exosomes. Many studies have documented the uptake of exosomes by target cells; however, to date, only a handful of examples have described specific exosome and target-cell interactions. For example, exosomes from T-, B- and dendritic immune cells were shown to communicate with antigen presenting cells by transferring their contents in a unidirectional manner and modulating gene expression in the recipient cell [36]. Uptake of ovarian cancer secreted exosomes by NK cells has been demonstrated and was found to require the surface expression of phosphatidylserine (PS) as an uptake signal [37]. Lastly, the expression of galectin-5 on the surface of rat reticulocyte exosomes was also found to modulate their uptake by macrophages [38]. These studies have demonstrated that there are distinct signals that mediate exosome and target-cell interactions; however, more work is required to fully understand the distinct mechanisms controlling this process.

1.3. Exosome Components

According to the current version of the exosome content database, Exocarta (Version 4; http://www.exocarta.org), 4,563 proteins, 194 lipids, 1,639 mRNAs and 764 microRNAs have been identified in exosomes of many different cell types and from multiple organisms [39,40], thus demonstrating their complexity. The most frequently identified proteins in exosomes (as compiled by Exocarta) include membrane transport and fusion proteins, such as tetraspanins (CD9, CD63, CD81), heat-shock proteins (Hspa8, Hsp90), GTPases (EEF1A1, EEF2), and MVB biogenesis proteins (Alix). Other identified proteins include cytoskeletal proteins (actin, syntenin, moesin), metabolic enzymes (GAPDH, LDHA, PGK1, aldolase, PKM), signal transduction proteins (annexin, 14-3-3ε, 14-3-3ξ) and the carrier protein, albumin. The specific protein composition will depend on the cell type or tissue source from which the exosome originates and may fluctuate according to physiological changes. In addition, many of these proteins may function as specific exosome markers, particularly the tetraspanins, CD63 and CD81.

Beyond proteins, exosomes are also enriched in lipids and may act as cell-to-cell lipid mediators. Exosomes predominantly contain lipids, such as cholesterol, diglycerides, sphingolipids (including sphingomyelin and ceramide), phospholipids, glycerophospholipids (including phosphatidylcholine (PC), phosphatidylserine (PS), phosphatidylethanolamine (PE) and phosphatidylinositol (PI)) and polyglycerophospholipids (*i.e.*, bisphosphate). The ratio is increased for certain exosomal lipids when compared to parental cell lipids; these include sphingomyelin, PS, PC, PI and cholesterol, which can be present at as much as four times greater amounts and may account for the increased membrane rigidity of exosomes. Exosomes have also been reported to contain bioactive lipids, such as prostaglandins and leukotrienes, and active enzymes of lipid metabolism that may generate these lipids [13,41]. The presence of certain lipids, such as PS, on the outer membrane of exosomes can function in exosome recognition and internalization [42]. In this way, exosomes function as lipid carriers, allowing the transport of the bioactive lipids they carry to a recipient cell. This process of exosomal trafficking, particularly in the context of the tumor microenvironment, could lead to an enrichment of certain tumor progressive/immunosuppressive lipids, such as prostaglandins [43]. On the other hand, it may also lead to a replacement of harmful exosome lipid contents with beneficial ones. For example, docosahexaenoic acid (DHA), an omega-3 polyunsaturated fatty acid with many health and anticancer benefits, could be supplied by exosomes throughout the tumor microenvironment

to affect cell-to-cell communication, reduce tumor cell growth and increase sensitivity to therapeutic interventions, particularly in breast tumors [44]. However, further studies are required to determine, which lipids participate in exosomal cell-to-cell communication and whether *ex vivo* manipulation is a plausible and/or effective therapeutic approach.

Exosomal transport of mRNAs and other non-coding RNAs, including microRNAs, was discovered only recently [45], and due to this exciting discovery, the interest in exosomes as carriers of genetic information is burgeoning, particularly in cancer research. Current reports have shown that the majority of the RNA present in exosomes is somewhat degraded and less than 200 nucleotides in length; however, recognizable proteins could be generated from *in vitro* translation of exosomal RNA extracts, thus demonstrating that full-length mRNAs are present [45]. Several studies have demonstrated that the RNA present in exosomes is very different from the parental cell RNA content, with the apparent lack of ribosomal RNA [36,45–47]. In contrast, the exosomal microRNA content is similar to that in the original tumor, thus peaking researchers' interests in the use of exosomal microRNA profiles for cancer diagnostics [9,48]. However, an abundance of certain microRNAs that are not present or present at very low levels in the parental cells has recently been observed [49,50]. These results suggest that certain microRNAs may be preferentially secreted. However, the mechanisms for selective packaging and release of exosomal microRNAs are currently unknown, and whether these microRNAs may serve as reliable markers of disease is yet to be determined.

The first in-depth screening study was recently conducted to examine the entire transcriptome, miRNome and proteome of exosomes derived from melanoma cells and normal melanocytes [51]. Thousands of mRNAs that are associated with melanoma progression and metastasis, as well as several microRNAs (miR-31, miR-185 and miR-34b) that are involved in melanoma invasion were identified. In addition, several differentially expressed proteins, such as HAPLN1, GRP78, syntenin-1, annexin A1 and annexinA2, were identified, which were specific to the melanoma exosomes and may be involved in the malignant conversion of melanocytes. This study demonstrates the need for more in-depth explorations of exosome contents, so that specific targets may be identified and translated into clinical applications for disease biomarkers or potential therapeutic targets for cancer patients.

1.4. Exosome Isolation and Examination

Because exosome membranes are enriched in cholesterol, sphingomyelin, ceramide and lipid raft-associated proteins [52,53], they are highly stable and can be collected from various bodily fluids or from cell culture mediums [54]. Due to their small size and low density, exosome isolation usually involves multiple centrifugation and ultracentrifugation steps with a rotational force up to 100,000 × *g* for sedimentation. Centrifugation is also sometimes combined with 0.1 μm to 0.22 μm filtration in order to separate the nano-sized particles and to exclude larger particles and cellular debris, (see, for review of methods, [54]). For reduction of protein aggregate contamination and for obtaining a purer exosome preparation, sucrose, iodixanol [55], deuterium oxide density gradients (also called cushions) or proprietary reagents, such as ExoQuick (System Biosciences), have also be utilized [56,57]. Immunoaffinity capture methods can be used to isolate exosomes from cancer cells or patient serum using beads coated with antibodies against presumably any exosome-specific surface marker, such as the tetraspanins, CD63 or CD82, as a way to forego any ultracentrifugation. Epithelial cell adhesion molecule (Ep-CAM)-positive exosomes have been collected from the serum from lung [38] and

ovarian cancer [48,58] patients, and HER2-positive exosomes have been isolated from HER2 overexpressing breast cancer cells [59] using this method.

Electron microscopy (EM) combined with negative staining is the standard method for visualization of whole-mount exosome preparations. Typical EM results show rounded vesicles with lipid bilayers, and sometimes, a "cup-shaped" morphology is observed, which may depend on the preparation process used. A review of these various isolation procedures and a comparison of the images obtained by EM analysis has been recently published [60].

Once the exosomes are collected, there are many downstream analysis options available. Exosomal proteins may be extracted utilizing standard cell lysis buffers or the TRIzol® reagent (Life Technologies-Invitrogen) [61]. Proteins may be detected and analyzed by immunoblot procedures [61] or mass spectrometry (see, for review, [62]). Because the RNA content of exosomes is mostly small RNAs, the selection of RNA isolation technique is an important consideration. Various RNA extraction techniques, including phenol-based techniques (TRIzol®), silica column (e.g., RNeasy® (Qiagen) or miRCURY™ (Exiqon)) and combined phenol and silica column approaches (e.g., TRIzol® followed by RNeasy (Qiagen), miRNeasy (Qiagen) or mirVana™ (Ambion)) have been utilized and compared [63–65]. The RNA yield can be determined by spectrophotometric analysis at 260 nm, and a profile of the exosomal RNAs can be determined using the Agilent 2100 Bioanalyzer Lab-on-a-Chip instrument system (Agilent Technologies). Typical profiles of RNA extracted from exosomes contain a size distribution of 25–2000 nucleotides and are characteristically absent of ribosomal RNAs [63]. Detection of specific small RNA or microRNA species can be determined by real-time reverse-transcription PCR assay and oligonucleotide microarray analysis [51], or more in-depth analysis next-generation RNA sequencing can be applied [47,51,66].

Due to their nanometer size, the process of quantifying exosomes is somewhat of a challenge. There are two primary approaches currently used to quantify the amount of exosomes isolated from a preparation: quantification of the amount of exosomal protein using enzyme-linked immunosorbent assays (ELISA) or by immunoblotting. However, new approaches to quantify exosome secretion have been demonstrated using cell lines stably expressing GFP tagged CD63 (a specific marker of exosomes), thus generating exosomes with a traceable marker that can be easily measured by fluorescent spectrometry [32]. New nanoparticle/exosome tracking analysis technologies have recently been developed by Nanosight Ltd. [67]. These systems are equipped with a blue laser and camera that can visualize and measure nanoparticles within the 30 nm to 1,000 nm range (as demonstrated in [30,68]). Because the field of exosome research is in a phase of rapid growth, the refinement of isolation, imaging and visualization methods are expected to improve, along with the identification of specific molecular markers for isolation and the development of new technological approaches.

2. Exosomes and Intercellular Communication in Tumor Progression

2.1. Exosome Mediation of Intercellular Communication

It is well recognized that tumor development and progression is dependent on the reciprocal relationship between cancerous cells and their surrounding microenvironment. While the cancerous cells, which harbor many pro-tumorigenic genetic mutations, are the main driving force of tumor

development, the surrounding stroma, which includes fibroblasts, endothelial and infiltrating immune cells, play a supportive and enabling role (reviewed in [69]). This reciprocal relationship requires not only a particular spatial interaction, but also the ability for the cancerous cells to communicate with the microenvironment by exchange of certain soluble proteins and genetic factors. Cancer cells are known to secrete factors that can promote the formation of new blood vessels, known as angiogenesis, to obtain oxygen to feed the tumor and to modify their adhesive properties in order to promote migration and invasion into the newly formed vasculature. Tumor cells of many different cancer types have been shown to secrete exosomes in greater amounts than normal cells [13], thus allowing the transfer of tumor-associated signaling molecules, including microRNAs, via fusion of the exosome with the target cell membrane [70].

Tumor-derived exosomes (TD-exosomes) are generally considered pro-tumorigenic. However, some anti-tumorigenic abilities have also been described. For example, exosome-like nanoparticles isolated from pancreatic cancer cells were shown to induce apoptosis in tumor cells [71]. Other studies have focused on the use of TD-exosomes as a source for tumor antigens for the development of exosome-based immunotherapies [72–74]. One report demonstrated that modified cell lines expressing interleukin-2 (IL-2) produced TD-exosomes containing IL-2 with increased antitumor effects [72]. A Chinese phase I clinical trial demonstrated that ascites-derived exosomes combined with granulocyte-macrophage colony stimulating factor could modulate the immune response and induce an antitumor cytotoxic T-lymphocyte (CTLs) response in 40 patients with colorectal cancer [73]. Another study showed that exosomes derived from IL-2 GPI-anchored renal cancer cells could induce CTLs and significant cytotoxic and antitumor effects *in vitro*, suggesting a novel strategy for an exosome-based vaccine for renal cell carcinoma [75]. Heat-stressed tumor cells were shown to produce exosomes that could attract and activate dendritic and T-cells, induce specific antitumor immune responses and inhibit tumor growth *in vivo* [76]. However, despite the above-described antitumor characteristics of exosomes, it is still unclear whether the immunomodulatory effects of exosomes secreted from tumor cells are either cancer-promoting or cancer-inhibiting, as no studies have demonstrated any immune stimulatory effects of TD-exosomes. Rather, many immune-evasion characteristics have been described. In this context, TD-exosomes from the ascites of ovarian cancer patients were recently shown to express the death ligands, FasL and TRAIL, which could trigger apoptosis in immune system cells, thereby inhibiting a tumor growth inhibitory immune response [77].

As described above, most initial studies on TD-exosomes were focused on their interaction with the immune system, while the effects of TD-exosomes on the tumor microenvironment has been less characterized. However, recent studies have shown that TD-exosomes have many pro-tumorigenic functions and are able to transfer their phenotypic traits (such as onco-proteins or onco-microRNAs) to a recipient cell and promote cancer stimulatory activities, such as proliferation, extracellular matrix remodeling, migration and invasion and angiogenesis, and contribute to the pre-metastatic niche formation for the promotion of metastasis. The following sections will discuss these topics in greater detail.

2.2. Exosome Modulation of Extracellular Matrix, Stromal Cells and Invasion

Remodeling of the extracellular matrix and alterations in cell-cell and cell-extracellular matrix interactions are the first barriers that must be overcome for a cancer cell to migrate and travel to distant sites. Several studies have shown that TD-exosomes can alter the extracellular matrix through

secretion of matrix metalloproteinases (MMPs) or activators of MMPs, such as heat shock proteins. MMPs are zinc-dependent plasma membrane endo-peptidases that can degrade extracellular matrix proteins, such as collagen, fibronectin, proteoglycans and laminins. In fibrosarcoma and melanoma cells, it was shown that MT1-MMP was secreted in exosomes and could activate pro-MMP-2 and degrade collagen and gelatin [78]. Other studies have demonstrated that heat shock proteins, such as hsp90, are also secreted via exosomes and can activate MMP-2 to enhance invasion of cancer cells [79]. The role of platelets in tumor progression has been recently investigated, where it was found that platelet-derived microvesicles contribute to cancer progression in lung cancer cell lines by stimulating proliferation, cyclin D2 expression, adhesion to endothelial cells, invasion and angiogenesis through activation of mitogen-activated protein kinase (MAPK) p42/44, MT1-MMP, MMP-9, IL-8 and VEGF and other factors controlling these processes [80]. The effects of TD-exosomes on mesenchymal stem cells (MSCs) have also been explored. For example, TD-exosomes from ovarian cancer cells can induce a tumor-associated myofibroblast-like phenotype on adipose-derived MSCs, suggesting that exosomes contribute to the generation of tumor-associated fibroblasts in the tumor stroma [81]. Recently, Bobrie *et al*. showed that inhibition of Rab27a in breast cancer cells decreases the secretion of exosomes and MMP-9, resulting in a decrease in primary tumor growth and lung dissemination *in vivo*, thus further demonstrating that exosome secretion promotes tumor formation and progression and that inhibition of exosome secretion may impede tumor growth [82].

2.3. Exosome Stimulation of Tumor Angiogenesis and Metastatic Niche Formation

In order for a tumor to sustain its growth and survival, they have to obtain greater amounts of oxygen and nutrients through the formation of new blood vessels or by metastasizing to more hospitable organ sites. TD-exosomes have been shown to transport oncogenic and pro-angiogenic factors to cells within the tumor microenvironment to induce neoangiogenic activity and to promote premetastatic niche formation. For example, a hypoxic tumor microenvironment has been shown to enhance the secretion and transport of exosomes and pro-angiogenic protein factors that could potentially modulate the microenvironment to facilitate angiogenesis and metastasis [83]. Recently, *in vitro* hypoxia experiments in glioma cells and patient samples showed an enrichment of hypoxia-regulated mRNAs and proteins, such as MMPs, IL-8, platelet derived growth factor (PDGF), caveolin-1 and lysyl oxidase, in secreted exosomes. In addition, it was demonstrated that these exosomes are also potent inducers of angiogenesis through possessing these growth factors and cytokines [84]. Colorectal cancer cell-derived microvesicles/exosomes are enriched in cell cycle-related mRNAs that could promote proliferation of vascular endothelial cells [85] and in several metastatic and signal transduction molecules [86]. Microvesicles released from human renal cancer stem cells were shown to stimulate angiogenesis and promote the formation of a premetastatic niche in the lungs *in vivo* [87]. As mentioned above, MSCs can also promote tumor growth; for example, bone marrow MSC-derived exosomes can enhance VEGF expression in tumor cells by activating the ERK1/2 pathway [88]. Likewise, platelet-derived exosomes have also been shown to stimulate mRNA expression of angiogenic factors (such as MMP-9), as described above [80]. Exosomes released by chronic myeloid leukemia cells were shown to promote angiogenesis in a Src-dependent fashion [89]. Melanoma-derived exosomes were found to prepare bone marrow progenitor cells for a pro-metastatic

phenotype through the receptor tyrosine kinase, MET [90]. A recent review proposes that the interaction and bidirectional exchange of genetic information, via secreted microvesicles, between macrophages and endothelial cells may work to promote vascular growth in the tumor microenvironment [91]. Overall, these studies clearly demonstrate that exosomes indeed function as pro-tumorigenic factors that can mediate intercellular communication in the tumor microenvironment and contribute to cancer progression.

3. Exosome-Derived microRNAs in Tumor Progression

3.1. Exosomes as Transporters of microRNAs

microRNAs are small (17–21 nt), non-coding RNAs that regulate gene expression at the post-transcriptional level through the RNA interference pathway. microRNAs are transcribed by RNA polymerase II as primary-microRNAs (pri-miRNAs) [92–94] and are processed in the nucleus by the enzyme, Drosha, into shorter hairpin structures of approximately 70 nucleotides in length, called pre-miRNAs. Pre-miRNAs are then transported from the nucleus, to the cytoplasm via Exportin 5 [95], where they are further processed into mature microRNA transcripts by the enzyme, Dicer [93]. The mature microRNA is then loaded into the ribonucleoprotein complex, known as the RNA induced silencing complex (RISC), and can bind in a sequence specific manner to the 3' untranslated region (UTR) of target mRNAs, resulting in either translational inhibition or mRNA degradation [96]. Presently, ~2000 microRNAs have been described in humans [97] and a single microRNA may regulate many mRNAs; likewise, a single mRNA may be targeted by many microRNAs, establishing microRNAs as the largest class of gene regulators [98]. Through this mechanism, microRNAs are an essential component to regulating most cellular and developmental processes, including developmental timing, organ development, differentiation, proliferation, apoptosis and immune regulation (see, for review, [99]). Therefore, it is of no surprise that microRNAs are involved in cancer development and progression, and depending upon their target gene and level of expression, microRNAs may function as either tumor suppressors or oncogenes and assist in the promotion or suppression of cancer growth and progression [100]. Aberrant microRNA expression has been described across many cancer types, with global downregulation of microRNA expression seen as a common trend [101,102]. The transport of mRNAs and microRNAs by exosomes was realized only recently, but has led to an explosion of interest in cancer research. The first study to demonstrate exchange of nucleic acids via exosomes examined secreted exosomes from mouse and human mast cell lines [45]. Using standard RNA and DNA extraction techniques of exosomes isolated by ultracentrifugation, Valadi *et al.* discovered the presence of small RNAs and mRNAs (but not DNA) from approximately 1300 genes present in exosomes that are not present in the parental cell and proposed that these RNAs be referred to as exosomal shuttle RNAs (esRNAs) to distinguish them from circulating microRNAs [45]. In addition, these esRNAs could be *in vitro* translated into functional proteins and transferred to other human and mouse mast cells, where new proteins were generated in the recipient cells [45]. This seminal study has generated much interest in the study of cell-cell communication via delivery of small RNAs by transfer through exosomes.

3.2. Secretion and Uptake of Exosome-Derived microRNAs

The exact mechanism of microRNA loading into MVBs in the endocytic pathway and secretion via exosomes is not well understood. Studies thus far have demonstrated that MVBs are associated with GW-bodies, also known as P-bodies, which are cytoplasmic foci, where post-transcriptional regulation of mRNAs occurs, and are enriched in GW182 and AGO2 proteins, two main components of the RISC [103]. In this study, it was shown that endosomes or MVBs are sites of microRNAs, microRNA-repressible mRNAs and RISC accumulation and action and that exosomes secreted via MVBs are enriched in GW182, suggesting a mechanism for microRNA loading. Depletion of some of the ESCRT components compromised microRNA-mediated gene silencing and led to an over-accumulation of GW182, thus suggesting that GW182 and microRNA-loaded AGO2 are sorted in to MVBs via ESCRT components [103]. However, as previously mentioned, a more recent study has demonstrated that microRNAs are released through a ceramide-dependent secretory mechanism [104]. Furthermore, a tumor-suppressive microRNA, secreted via this mechanism, was taken up by a recipient cell, where it exerted gene silencing and growth inhibition [104]. The GW182 protein may also be important for microRNA stability and secretion via exosomes [105]. In this study, it was shown that knockdown of GW182 by siRNA increased microRNA instability and reduced secretion via exosomes, whereas replenishment of GW182 restored microRNA stability, thereby demonstrating a role of GW182 in protecting AGO2-bound microRNA [105].

Other studies have demonstrated that certain microRNAs are selectively secreted in exosomes. For example, the let-7 microRNA family is selectively secreted via exosomes in metastatic gastric cancer cell lines. Since this family of microRNAs targets oncogenes, such as Ras and HMGA2, they are generally considered a tumor-suppressive group of microRNAs; however, whether their release via exosomes is to promote or inhibit oncogenesis remains unclear [106]. Breast cancer cell lines selectively release the majority of miR-451 and miR-1246 via exosomes as compared to their parental cell, whereas these specific microRNAs are retained in non-malignant mammary epithelial cells and normal fibroblast cells [49]. The authors note that the biogenesis of miR-451 is dicer-independent [107], thus raising the possibility that non-canonical processing of microRNAs may target them for selective exosome release. Overall, these results suggest that there is a specific selection mechanism for microRNA release. However, the exact mechanism remains to be elucidated.

While the exact mechanism of exosome-derived microRNA uptake and processing in recipient cells is largely unknown, many types of cells have been shown to absorb exosomal microRNAs, where they can induce post-translational repression of target mRNAs. For example, T-cells have been shown to transfer microRNAs-loaded exosomes in an antigen-driven unidirectional manner to antigen presenting cells in the immune synapsis, where they modulate gene expression [36]. Additionally, mouse dendritic cells (DCs) release exosomes containing different microRNAs, depending on their level of maturation. These microRNAs were absorbed by recipient DCs, where they were shown to repress target mRNA expression [108]. One of the first studies to show that this unique intercellular method of communication could contribute to the initiation and progression of cancer demonstrated that hepatocellular carcinoma cells produce exosomes that can be internalized by other cells. These exosomes were shown to transmit microRNAs that modulated the expression of transforming growth factor β activated kinase-1 (TAK1), whose loss is implicated in hepatocarcinogenesis [109]. Thus,

these studies have begun to shed light on the potential mechanisms of exosome uptake and the functional consequences of microRNA transfer.

3.3. microRNA Profiling of Tumor-Derived Exosomes in Clinical Samples

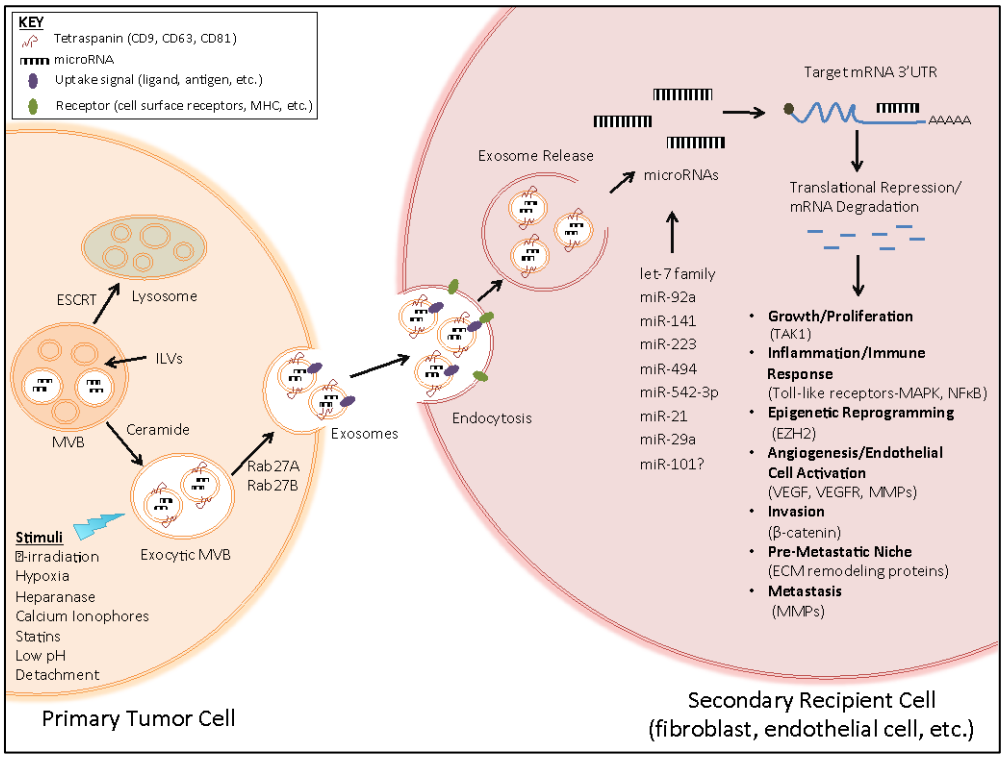
The discovery of the transport and exchange of microRNAs via exosomes has also generated much interest in the use of circulating TD-exosomes and their resident esRNAs as clinical diagnostic markers for cancer. To this end, a handful of studies have examined the microRNA profile from circulating TD-exosomes and compared the expression levels to the original tumor cells. For example, circulating TD-exosomes isolated from the serum of patients with ovarian cancer, age-matched controls and primary tumor cell cultures and matched sera were examined for microRNA expression changes. This study found that the expression of eight microRNAs, previously demonstrated to be diagnostic in ovarian cancer, were similar between cellular and exosomal microRNA preparations, thus suggesting that circulating TD-exosomes could be used as surrogate diagnostic markers for biopsy profiling, particularly in asymptomatic patient populations [9]. Another study from Taylor *et al*. evaluated the levels of microRNAs from the plasma of patients with lung adenocarcinoma, matched tumor samples and controls; they similarly observed no significant differences in exosome microRNA levels between microRNAs derived from circulating exosomes or from microRNAs from the primary tumors [48]. In another study, circulating exosomal/microvesicle-derived microRNAs were profiled from the plasma of prostate cancer patients with and without metastases [110], and a distinct set of 11 microRNAs was present at significantly greater amounts in patients with metastases compared to those without metastases. The association of two of these 11 microRNAs (miR-141 and miR-375) were confirmed in plasma exosomes from a separate patient cohort with recurrent or non-recurrent disease, thus demonstrating that changes in microRNA concentrations present in circulating exosomes from prostate cancer patients may be used for diagnosis and tumor staging [110]. In serum obtained from esophageal squamous cell carcinoma (ESCC) patients, microRNA expression profiling showed that miR-1246 was consistently elevated in patients *versus* controls and was an independent risk factor for poor survival [111]. The authors also indicate that miR-1246 was not upregulated in ESCC tissue samples; however, this observation is consistent with the previously mentioned report of preferential exosome secretion of miR-1246 from breast cancer cells [49]. Overall, these exosomal microRNA profiling studies, summarized in Table 1, have found that microRNA expression signatures are not significantly different between TD-exosomes and tumor cells, with the exception of miR-1246, suggesting that these circulating TD-exosome microRNAs could be utilized as a surrogate for biopsy microRNA profiling. In addition, a database called miRandola has been created to catalog all extracellular circulating microRNAs and currently contains 2312 entries with 581 unique mature microRNAs identified in circulation from 21 different types of samples [112].

Table 1. Summary of clinical microRNA profiling studies of circulating tumor exosomes/microvesicles.

Cancer type	Clinical samples	Exosome isolation method	Major findings	Potential diagnostic microRNAs	Reference
Ovarian cancer	Sera from patients with serous papillary adenocarcinoma (*n* = 50); sera from age-matched controls with benign ovarian adenoma (*n* = 10); primary ovarian adenocarcinoma tumor cell cultures and matched patient sera (*n* = 6).	Magnetic activated cell sorting using beads coupled with anti-EpCAM.	Exosomal microRNA profiles were similar in ovarian cancer patient samples and distinctly different from benign disease samples. microRNAs were elevated in exosomes *versus* tumor cells (31 out of 467).	miR-21, miR-141, miR-200a, miR-200c, miR-203, miR-205 and miR-214	[9]
Lung adenocarcinoma	Plasma from patients with lung adenocarcinoma (*n* = 27); plasma from controls (*n* = 9); matched plasma and lung tumor tissue (*n* = 4).	Size exclusion chromatography and magnetically activated cell sorting using beads coupled with anti-EpCAM.	No significant differences in exosome microRNA levels between microRNAs derived from circulating exosomes or from microRNAs from the primary tumor were observed.	miR-17-3p, miR-21, miR-106a, miR-146, miR-155, miR-191, miR-192, miR-203, miR-205, miR-210, miR-212, miR-214	[48]
Prostate cancer	Plasma from prostate cancer patients (*n* = 78); plasma from normal controls (*n* = 28); urine samples (*n* = 135); serum from patients with recurrent metastatic prostate cancer (*n* = 47) or non-recurrent disease (*n* = 72).	Filtration of plasma through a 1.2 μm filter, concentrated with a 150 kDa molecular weight cut-off.	The levels of 12 microRNAs were different between plasma exosomes of prostate cancer patients compared to control. Eleven microRNAs were present in significantly greater amounts in patients with metastases *versus* without. The association of exosomal miR-141 and miR-375 with metastases was confirmed in a second patient population.	miR-107, miR-130b, miR-141, miR-181a-2*, miR-2110, miR-301a, miR-326, miR-331-3p, miR-432, miR-438, miR-574-3p, miR-625*	[110]
Esophageal squamous cell carcinoma	Serum from ESCC patients (*n* = 101); Serum from healthy controls (*n* = 46).	Sequential centrifugation, 0.22 μm filtration and ultracentrifugation.	miR-1246 was markedly elevated in serum and exosomes from ESCC patients and was a strong independent risk factor for poor survival. miR-1246 expression was not increased in ESCC tissue samples.	miR-1246	[111]

Abbreviations: EpCAM = epithelial cell adhesion molecule; ESCC = esophageal squamous cell carcinoma.

Figure 1. Biogenesis, secretion and uptake of tumor-derived exosomes in the tumor microenvironment. Exosomes are formed by the inward budding of the multivesicular body (MVB) membrane in the form of intraluminal vesicles (ILVs). Exosome formation and cargo sorting into lysosomes involves the endosomal sorting complex required for transport (ESCRT), which recognizes ubiquitinated proteins. Exosome production and secretion also occurs through an ESCRT-independent process involving the sphingolipid, ceramide, and the enzyme neutral, sphingomyelinase (the enzyme that converts sphingomyelin to ceramide). Exosomes secretion can be stimulated by various chemical, environmental and mechanical stimuli, such as gamma-irradiation, hypoxia (low oxygen), low pH, matrix detachment, *etc.* Exosomes are secreted in exocytic MVBs following fusion of MVBs with the cell membrane, a process that depends on Rab GTPases (Rab27A, Rab27B). Exosomes released from a primary tumor cell will display similar membrane components as their cell of origin, such as receptor ligands or antigens. Endocytosis of exosomes may occur through activation of cell surface receptors or bioactive lipid ligands. Upon endocytosis by a secondary recipient cell, such as fibroblasts or vascular endothelial cells, exosomes can release their microRNA cargo. The transferred microRNAs are functionally active and can regulate gene expression in the recipient cell through post-translational regulation of target mRNA expression, leading to mRNA degradation or de-stabilization. microRNA-dependent gene regulation can activate various processes involved in tumor development and progression. Abbreviations: TAK1, transforming growth factor β activated kinase-1; MMPs, matrix metalloproteinases; MAPK, mitogen activated protein kinase; NFκB, nuclear factor kappa-light-chain-enhancer of activated B-cells; EZH2, enhancer of zeste homolog 2; VEGF/VEGF, vascular endothelial growth factor/receptor; ECM, extracellular matrix.



3.4. Pro-Tumorigenic Effects of Exosome-Derived microRNAs in Vitro

The pro-tumorigenic effects of exosome-derived microRNAs after uptake by a recipient cell have recently begun to emerge. Thus far, exosome-derived microRNAs, through target gene transcriptional repression, have the demonstrated ability to induce cell migration, inflammation, immune responses, angiogenesis (including endothelial cell migration and tube formation), invasion, pre-metastatic niche formation and metastasis; see Figure 1. Therefore, these studies have implicated cancer cell exosome-derived microRNAs in most aspects of tumor progression. For example, leukemia cell exosomes have been shown to communicate with human umbilical vein endothelial cells (HUVECs), leading to increased cell migration and tube formation [113]. In this study, K562 leukemia cells were transfected with a Cy3-labeled pre-miR-92a and co-cultured with HUVECs. The Cy3-labeled miR-92a, derived from the K562 cells, could be detected in the cytoplasm of the endothelial cells and was co-localized with the exosomal marker, CD63. In addition, the expression of integrin α5, a target of miR-92a, was also greatly reduced in the recipient cells, thus demonstrating that an exosome-derived microRNA can function as an endogenous microRNA in a recipient cell and that exosomal microRNAs play an important role in cancer-to-endothelial cell communication [113]. As previously mentioned, microvesicles released from human renal cancer stem cells containing pro-angiogenic mRNA, and microRNA were shown to greatly stimulate endothelial cell growth and vessel formation and enhance lung metastases after *in vivo* implantation in a severe combined immunodeficient (SCID) mouse model [87]. In this study, molecular characterization of microvesicles, derived from CD105-positive (a mesenchymal stem cell marker) renal cancer stem cells, was conducted and was found to contain a set of pro-angiogenic mRNAs and microRNAs that are implicated in tumor progression and metastases [87]. The previously mentioned study, which profiled the miRNome of melanoma exosomes, identified 228 microRNAs that were differentially regulated in melanoma exosomes *versus* normal melanocyte exosomes, 15 of which are known to be associated with melanoma invasion and metastasis [51].

Tumor-associated macrophages, which are known to promote invasion and metastasis, have been shown to secrete microvesicles containing microRNAs that could be taken up by breast cancer cells. In a co-culture system, it was demonstrated that uptake of IL-4 activated macrophage secreted exosomes could promote the invasion of breast cancer cells, due to uptake of miR-223 (a microRNA specific for IL-4 activated macrophages) and disruption of the Mef2c-β-catenin pathway [114].

Through pathogen recognition receptors, such as Toll-like receptors (TLRs), and their associated downstream signaling pathways, such as nuclear factor kappaB (NF-κB) and MAPK, exosomal microRNAs may also play a large role in the regulation and homeostasis of the innate immune response by fine-tuning the mechanisms responsible for the production and release of cytokines/chemokines, adhesion and co-stimulatory molecules in epithelial cells (see, for review, [115]). These mechanisms, in the context of cancer, could be disrupted, thereby promoting an immune-evasion response and cancer promotion. For example, an interesting recent study has demonstrated that secreted microRNAs may act as ligands by binding to TLRs on recipient cells. Specifically, miR-21 and miR-29a secreted in exosomes from lung cancer cell lines were shown to bind to murine TLR7 and human TLR8 and triggered a TLR-mediated pro-metastatic inflammatory response that could lead to tumor growth and metastasis [116].

The process of malignant transformation may also alter the specific species of microRNAs that are secreted in exosomes or retained in cells. For instance, selective release of certain microRNA

populations has been demonstrated in malignant breast cancer cells. Specifically, the microRNAs, miR-451 and miR-1246, produced by malignant breast epithelial cells are released, whereas the majority of these microRNAs are retained in non-malignant mammary epithelial cells [49]. A follow-up study from this same group demonstrated that these selectively exported microRNAs are packaged in exosomes that are larger than conventional exosomes and are enriched in CD44, a protein relevant to breast cancer metastasis. In contrast, they showed that normal cells release microRNAs in a homogenous type of vesicle, suggesting that the process of malignant transformation may alter the pathways by which microRNAs are exported from cells, thus leading to differences in exosome content and morphology [117].

TD-exosomes may also modulate pre-metastatic niche formation via long-range transfer of microRNAs. One study has demonstrated that exosomes from metastatic rat adenocarcinoma cells are preferentially taken up by lymph node stroma cells and lung fibroblasts [118]. The transferred microRNAs significantly affected mRNA translation of many genes, including proteases, adhesion molecules, chemokine ligands, cell cycle- and angiogenesis-promoting genes and oxidative stress response. In particular, miR-494 and miR-542-3p modulated the expression of cadherin-17 with concomitant upregulation of matrix metalloproteinases. Together, these findings demonstrate that TD-exosomes may target non-transformed cells in pre-metastatic tissues, leading to modulation of gene expression in these cells specifically through transfer of microRNA and priming distant tissues for tumor cell hosting [118].

Exosomal transfer of microRNAs could also induce permanent changes in recipient cell phenotypes via transfer of microRNAs that are known to regulate genes involved in epigenetic reprogramming (*i.e.*, miR-101 regulation of the histone methyltransferase EZH2). For example, it has been demonstrated that microvesicles derived from one cell type can deliver mRNAs that could mediate gene expression and alter cell fate in a secondary recipient cell type (reviewed in [119]). While we are unaware of any studies demonstrating changes in epigenetic programming or cell fate through the exchange of exosomal microRNAs, it is certainly a highly plausible occurrence. In conclusion, it is clear from these studies, summarized in Table 2, that transfer of microRNAs via exosomes is indeed a mechanism of intercellular communication that can initiate and promote tumor progression via transfer of genetic information at local and distant cells and tissues.

Table 2. Summary of *in vitro* studies of microRNAs derived from cancer cell exosomes.

Cell Line model	Major findings	Predominant microRNAs	Target genes or pathways	Reference
Human and mouse mast cells	Identified small RNAs, including 121 microRNAs and 1,300 specific mRNAs. Detected mouse exosomal RNA and new mouse proteins in human mast cells after treatment with mouse mast cell exosomes. Coined the term "exosomal shuttle RNA (esRNA)".	let-7, miR-1, miR-15, miR-16, miR-181, miR-375.	None tested	[45]
Metastatic gastric cell line	Profiled microRNA expression by microarray in exosomes isolated from gastric cancer cells. let-7 microRNA family was enriched in exosomes.	let-7 family	None tested	[106]
Co-culture of IL-4-activated macrophages and breast cancer cells	miRNAs can be transferred from macrophages to breast cancer cells. miR-223 released by macrophages was found in MCF7 and MDA-MB-231 cells and promoted invasion.	miR-223	Mef2c-β-catenin pathway	[114]
Mouse dendritic cells	Exosomal microRNA from dendritic cells can be transferred to a recipient dendritic cell and repress microRNA target mRNAs in the acceptor cell.	miR-148a, miR-451	Luciferase reporter containing tandem microRNA target sequences	[108]
Leukemia cells and endothelial cells	Leukemia cells released microRNAs from the miR-17-92 cluster and were taken up by human umbilical vein endothelial cells (HUVECs) and repressed a target mRNA. Did not affect the growth of HUVEC cells, but did enhance cell migration and tube formation.	miR-92a	Integrin α5	[113]
Hepatocellular carcinoma cells	Transmission of exosome microRNAs from hepatocellular carcinoma cells could contribute to the initiation and progression of hepatocellular carcinoma by targeting a tumor suppressor frequently lost in hepatocarcinogenesis.	miR-584, miR-517c, miR-378, miR-520f, miR-142-5p, miR-451, miR-518d, miR-215, miR-376a*, miR-133b, miR-367	Transforming growth factor β activated kinase-1 (TAK1) pathway	[109]
Renal cancer stem cells	Microvesicles were secreted from human renal cell carcinoma that could trigger angiogenesis and premetastatic niche formation in a severe combined immunodeficient (SCID) mouse model.	miR-92, miR-141, miR-29a, miR-650, miR-151, miR-19b, miR-29c	Increase in VEGFR1 and MMP-9 expression	[87]

Table 2. *Cont.*

Cell Line model	Major findings	Predominant microRNAs	Target genes or pathways	Reference
Breast cancer cells	Selective release of certain microRNA populations was demonstrated in malignant breast cancer cells that are retained in non-malignant mammary epithelial cells.	miR-451, miR-1246	None tested	[49]
Metastatic rat adenocarcinoma cells	Exosomes were preferentially taken up by lymph node stroma cells and lung fibroblasts. The transferred microRNAs affected mRNA translation of many genes.	miR-494, miR-542-3p	Cadherin-17 and many proteases, adhesion molecules, chemokine ligands, cell cycle- and angiogenesis-promoting and oxidative stress response genes.	[118]
Lung cancer cell lines	miR-21 and miR-29a were secreted in exosomes and could bind to murine TLR7 and human TLR8 and trigger a Toll-like receptor (TLR)-mediated prometastatic inflammatory response that could lead to tumor growth and metastasis.	miR-21 and miR-29a	Toll-like receptor (TLR) 8 and 9	[116]
Melanoma and normal melanocyte cells	The first in-depth screening to examine the entire exosome transcriptome, miRNome and proteome. Thousands of mRNAs and 15 microRNAs that are associated with melanoma progression and metastasis were identified.	let-7 family, miR-138, miR-125b, miR-130a, miR-34a, miR-196a, miR-199b-3p, miR-25, miR-27a, miR-200b, miR-23b, miR-146a, miR-613, miR-205, miR-149	None tested	[51]

4. Exosomes and Breast Cancer

4.1. Exosomes in Normal Mammary Epithelium

The human mammary gland is comprised of two main compartments: the branching epithelial ductal-lobular system and the supporting stroma. The epithelial component is comprised of two epithelial cell types: the luminal cells, whose function is to maintain the apical-basal polarity within the lumen and to produce and secrete milk into the ducts, and the myoepithelial cells, whose function is to maintain the organization of the mammary gland and to contract and eject the milk in response to hormonal signals. The mammary gland is one of the few organs to undergo significant developmental changes after birth, including growth at the onset of puberty and pregnancy, lactation and regression upon cessation of lactation [120]. Breast milk, which is a complex and nutrient rich liquid that contains proteins, lipids, carbohydrates and trace elements, is known to be the optimal nutrition and an important source of immunoprotective components for infants during their first months of life [121]. Exosomes have been identified in both colostrum and mature human breast milk that have the capacity to potentially influence the immune response in infants [5]. Specifically, breast milk exosomes were found to inhibit anti-CD3-induced IL-2 and interferon (IFN)-γ production and could increase the number of T-regulatory cells from peripheral blood mononuclear cells when incubated with milk vesicle preparations [3]. Furthermore, breast milk exosomes are also known to contain microRNAs [122]. Over 600 unique microRNAs, originating from ~450 microRNA precursors, have been identified in human breast milk exosomes using deep sequencing technology, and greater than 65% of the known immune-related microRNAs were enriched in these exosomes [122]. These results suggest that microRNAs can be transferred from the mother's milk to the infant via the digestive tract, where they could play an important role in the development of the infant's immune system, although further work must be done to confirm these speculations. At this time, however, it is unknown if exosomes are secreted in breast luminal epithelial cells during the growth phase of puberty or pregnancy, during regression or during the mammary resting state. It is also unknown what role the myoepithelial cells play in exosome secretion or in crosstalk between the two epithelial lineages and the breast stroma in the normal mammary gland. Furthermore, the aberrant secretion of exosomes and their specific contribution to breast cancer development and early progression are also unknown at this time.

4.2. Exosomes in Breast Cancer Development and Progression

Breast cancer is one of the most frequently observed cancers in industrialized countries and is the second leading cause of cancer death among women in the US. **Every year, ~200,000 new cases of invasive breast cancer will be diagnosed, and around 40,000 women are expected to die from breast cancer [123].** The general view of the tumorigenic process involves cells that have acquired critical genetic and epigenetic abnormalities that inhibit their responsiveness to normal growth and regulatory signals. In the mammary gland, the majority of these critical genetic changes occurs in the luminal epithelium at the transition from normal, to hyperplastic, to pre-invasive lesions and contains greater measures of genetic and morphological abnormalities, including aneuploidy [124–126], oncogene amplification [127,128] or allele imbalance [129,130]. It is thought that fewer abnormalities exist in precursor lesions and more are acquired as the cancer progresses. In addition, the loss of the

normal myoepithelium cell layer and apical-basal polarity are early signs of tumorigenesis. It has also been demonstrated that the microenvironment also undergoes changes and has a dramatic influence on the tumor, even at the pre-invasive stage [131]. It is suggested that a reciprocal relationship between breast cancer cells and their surrounding microenvironment predominantly influences the energetics and growth of the cancer. However, the role of exosomes in mediating this reciprocal relationship is only beginning to emerge. Early studies in breast cancer exosomes demonstrated that exosomes might play a role in the control of tumor growth. For example, mice pretreated with exosomes derived from murine mammary carcinomas had increased rates of tumor growth, due to inhibition of natural killer cells [132]. In co-culture experiments, tumor-associated macrophages were shown to transfer miR-223, a microRNA specific for IL-4 activated macrophages, into breast cancer cells, where it promoted invasion through activation of the Mef2c-β-catenin pathway [114]. Uptake of the epidermal growth factor receptor (EGFR) ligand, amphiregulin, carried by exosomes to breast cancer cells, increased their invasiveness compared to exosomes carrying other EGFR ligands, suggesting a role for exosomes in the cancer "field effect" and metastatic niche priming [133]. Another study examined the effects of exosomes derived from a triple-negative breast cancer (TNBC) cell line (Hs578T) *versus* its more invasive variant (Hs578T(i)$_8$) on three recipient TNBC cell lines. It was shown that exosomes from the more invasive variant increase proliferation, migration, and invasion and stimulate significantly more endothelial tubule formation in all recipient cell lines [134]. The intercellular communication between fibroblasts and breast cancer cells was recently examined, which showed that fibroblast-secreted exosomes could promote breast cancer cell protrusive and motile activity through Wnt-planar cell polarity (PCP) signaling and that co-injection of fibroblasts with breast cancer cells in an orthotopic mouse model could dramatically increase metastasis that was dependent on the PCP pathway [135]. A recent study utilized GFP-tagged CD63 expressing breast cancer cells to examine the fate of cancer-cell derived exosomes in a nude mouse model of breast cancer. They demonstrated that breast cancer cells could transfer their exosomes to other cancer cells and normal lung tissue *in vitro* and into the tumor microenvironment and the circulation of mice with breast cancer metastases *in vivo* [136].

While very few studies thus far have profiled or examined the microRNAs present in breast cancer exosomes, a previously mentioned study has identified a specific set microRNAs that are secreted in exosomes or retained in cells that differ between non-malignant and malignant breast cancer cells [49]. Further studies demonstrated that the selectively exported microRNAs from breast cancer cells are packaged in exosomes that differ from conventional exosomes [117]. However, the mechanism of this selective microRNA transport and altered exosome formation in malignant breast cancer cells is unknown at this time.

In summary, these studies suggest that breast cancer derived-exosomes may contribute significantly to breast tumor growth and development, promotion of angiogenesis, invasion and formation of a pre-metastatic niche to promote tumor growth and metastasis. Moreover, further elucidation of the mechanisms of how exosomes and their residual components mediate the intercellular communication in the breast tumor microenvironment and how this process unfolds during early malignant transformation and cancer progression are of great research interest.

4.3. Potential for Diagnostic, Prognostic and Therapeutic Interference

Much of the excitement surrounding TD-exosome research is due to its high clinical relevance. In particular, because breast cancer exosomes can be easily isolated through minimally invasive procedures, such as from the blood or ductal lavage of breast cancer patients, they have great potential in breast cancer diagnosis. Because exosomal microRNA profiles of circulating tumor exosomes tend to be unique from those in normal controls, breast cancer-specific exosomal-microRNA signatures may also be developed to predict tumor development. It is accepted that breast cancer is a highly heterogeneous disease with phenotypically diverse tumors, which have been categorized by their gene expression profiles as the intrinsic molecular subtypes of breast cancer. These are generally defined as: basal-like breast cancer, which generally corresponds to estrogen receptor (ER) negative, progesterone receptor (PR) and HER-2 negative (*i.e.*, triple negative), luminal A (ER positive, low grade), luminal B (ER positive, high grade) and HER-2 (or ErbB2) positive [137]. These subtypes have distinctly different gene and microRNA expression profiles [138,139]. Therefore, their secreted exosomes may also have distinctly different RNA profiles that could correspond to the molecular subtype of their host tumors. Exosomes may also be important players in chemotherapy and chemoresistance. For example, release of exosomes from HER2-overexpressing breast cancer cell lines (BT474 and SKBR3) or from HER2-positive breast cancer patient serum could bind to Trastuzumab (a monoclonal antibody therapy that interferes with the HER2 receptor) and lead to an inhibition of the anti-proliferative effects of Trastuzumab on SKBR3 cells. These results suggest that HER2-positive exosomes may interfere with anticancer therapy and may promote HER-2 driven tumor aggressiveness [140]. These findings have led to the development of a novel therapeutic strategy for exosome removal as an adjuvant to chemotherapy. In particular, Aethlon Medical, Inc. (San Diego, CA, USA) has introduced the HER2osome^{TM} as a new therapeutic strategy to maximize the effects of anti-HER2 therapies in combating breast cancer [141]. Furthermore, exosomes may also function to shuttle chemotherapies, such as cisplatin, out of the tumor cell, thus reducing their effectiveness [142]. For example, acquired resistance to cisplatin is associated with abnormalities of protein trafficking and secretion. Cisplatin-resistant ovarian cancer cells release more protein via exosomes, including greater levels of the lysosome-associated protein 1 (LAMP1), the putative cisplatin export transporter, MRP2, and the copper export proteins, ATP7A and ATP7B, than those released by cisplatin-sensitive cells [142]. This implicates the exosome secretion pathway in the resistance of breast cancer to cancer therapy. Lastly, because exosomes are naturally produced, cell-derived nucleic acid carriers, they also hold the potential to function as biological therapeutic delivery systems. For example, exosomes engineered to express the transmembrane domain of the PDGF receptor fused to the GE11 peptide were shown to successfully deliver the let-7a microRNA to EGFR-expressing breast cancer cells *in vitro* and EGFR-expressing xenograft breast cancer tissue *in vivo*, thus demonstrating the potential of exosomes to be used therapeutically to specifically target cancer cells with nucleic acid drug targets [143].

5. Future Directions

Current knowledge of TD-exosomes suggests that they can play an important role in the development and progression of cancer through modulation of intercellular communication within the tumor

microenvironment by the transfer of protein, lipid and RNA cargo. A further exploration of exosome secretion in the normal physiological state and during cancer development and progression, as well as the specific content exosomes transport under these conditions, will increase our understanding of their role in intercellular communication and tumorigenesis. Identification and modification of cancer cell-derived exosome contents may allow for the development of novel diagnostic, preventive and therapeutic approaches, with potentially minimally invasive procedures. Utilization of exosomes for therapeutic delivery may also prove to be the answer for the field of RNA therapeutics, whose main roadblock has been the development of an effective RNA delivery system. Furthermore, the concept of exploiting these extracellular vesicles or the creation of synthetic exosomes, called "exosome mimetics" for drug delivery may also allow for specific targeting of cancer cells by developing exosomes that harbor cell-specific targeting factors [144]. Although this field is in its infancy, it is easy to imagine all the future possibilities these natural nanoparticles hold.

Acknowledgments

We acknowledge grant support from the American Cancer Society (CNE-117557), the Susan G. Komen for the Cure Foundation (KG081083), the NIH OK-INBRE program (3P20RR016478-09S2), and the Oklahoma Center for the Advancement of Science and Technology (HR09-025).

Conflict of Interest

The authors declare no conflict of interest.

References

49. Thery, C.; Zitvogel, L.; Amigorena, S. Exosomes: Composition, biogenesis and function. *Nat. Rev. Immunol.* **2002**, *2*, 569–579.
50. Wickman, G.; Julian, L.; Olson, M.F. How apoptotic cells aid in the removal of their own cold dead bodies. *Cell Death Differ.* **2012**, *19*, 735–742.
51. Pan, B.T.; Teng, K.; Wu, C.; Adam, M.; Johnstone, R.M. Electron microscopic evidence for externalization of the transferrin receptor in vesicular form in sheep reticulocytes. *J. Cell Biol.* **1985**, *101*, 942–948.
52. Harding, C.; Heuser, J.; Stahl, P. Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes. *J. Cell Biol.* **1983**, *97*, 329–339.
53. Admyre, C.; Johansson, S.M.; Qazi, K.R.; Filen, J.J.; Lahesmaa, R.; Norman, M.; Neve, E.P.; Scheynius, A.; Gabrielsson, S. Exosomes with immune modulatory features are present in human breast milk. *J. Immunol.* **2007**, *179*, 1969–1978.
54. Masyuk, A.I.; Masyuk, T.V.; Larusso, N.F. Exosomes in the pathogenesis, diagnostics and therapeutics of liver diseases. *J. Hepatol.* **2013**, doi:10.1016/j.jhep.2013.03.028.
55. Vella, L.J.; Sharples, R.A.; Nisbet, R.M.; Cappai, R.; Hill, A.F. The role of exosomes in the processing of proteins associated with neurodegenerative diseases. *Eur. Biophys. J.* **2008**, *37*, 323–332.

56. Keller, S.; Rupp, C.; Stoeck, A.; Runz, S.; Fogel, M.; Lugert, S.; Hager, H.D.; Abdel-Bakky, M.S.; Gutwein, P.; Altevogt, P. CD24 is a marker of exosomes secreted into urine and amniotic fluid. *Kidney Int.* **2007**, *72*, 1095–1102.
57. Taylor, D.D.; Gercel-Taylor, C. MicroRNA signatures of tumor-derived exosomes as diagnostic biomarkers of ovarian cancer. *Gynecol. Oncol.* **2008**, *110*, 13–21.
58. Gallo, A.; Tandon, M.; Alevizos, I.; Illei, G.G. The majority of microRNAs detectable in serum and saliva is concentrated in exosomes. *PLoS One* **2012**, *7*, e30679.
59. Saman, S.; Kim, W.; Raya, M.; Visnick, Y.; Miro, S.; Saman, S.; Jackson, B.; McKee, A.C.; Alvarez, V.E.; Lee, N.C.; *et al.* Exosome-associated tau is secreted in tauopathy models and is selectively phosphorylated in cerebrospinal fluid in early Alzheimer disease. *J. Biol. Chem.* **2012**, *287*, 3842–3849.
60. Qiu, S.; Duan, X.; Geng, X.; Xie, J.; Gao, H. Antigen-specific activities of CD8+ T cells in the nasal mucosa of patients with nasal allergy. *Asian Pac. J. Allergy Immunol. Launched Allergy Immunol. Soc. Thail.* **2012**, *30*, 107–113.
61. Record, M. Exosomal Lipids in Cell–Cell Communication. In *Emerging Concepts of Tumor Exosome–Mediated Cell-Cell Communication*; Zhang, H.-G., Ed.; Springer: New York, NY, USA, 2013; pp. 47–68.
62. Bobrie, A.; Colombo, M.; Raposo, G.; Thery, C. Exosome secretion: Molecular mechanisms and roles in immune responses. *Traffic* **2011**, *12*, 1659–1668.
63. Hanson, P.I.; Cashikar, A. Multivesicular body morphogenesis. *Annu. Rev. Cell Dev. Biol.* **2012**, *28*, 337–362.
64. Schmidt, O.; Teis, D. The ESCRT machinery. *Curr. Biol.* **2012**, *22*, R116–R120.
65. Raymond, C.K.; Howald-Stevenson, I.; Vater, C.A.; Stevens, T.H. Morphological classification of the yeast vacuolar protein sorting mutants: Evidence for a prevacuolar compartment in class E vps mutants. *Mol. Biol. Cell* **1992**, *3*, 1389–1402.
66. Raiborg, C.; Stenmark, H. The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins. *Nature* **2009**, *458*, 445–452.
67. Katzmann, D.J.; Babst, M.; Emr, S.D. Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. *Cell* **2001**, *106*, 145–155.
68. Wollert, T.; Wunder, C.; Lippincott-Schwartz, J.; Hurley, J.H. Membrane scission by the ESCRT-III complex. *Nature* **2009**, *458*, 172–177.
69. Van Blitterswijk, W.J.; van der Luit, A.H.; Veldman, R.J.; Verheij, M.; Borst, J. Ceramide: Second messenger or modulator of membrane structure and dynamics? *Biochem. J.* **2003**, *369*, 199–211.
70. Subra, C.; Laulagnier, K.; Perret, B.; Record, M. Exosome lipidomics unravels lipid sorting at the level of multivesicular bodies. *Biochimie* **2007**, *89*, 205–212.
71. Trajkovic, K.; Hsu, C.; Chiantia, S.; Rajendran, L.; Wenzel, D.; Wieland, F.; Schwille, P.; Brugger, B.; Simons, M. Ceramide triggers budding of exosome vesicles into multivesicular endosomes. *Science* **2008**, *319*, 1244–1247.

72. Van Niel, G.; Charrin, S.; Simoes, S.; Romao, M.; Rochin, L.; Saftig, P.; Marks, M.S.; Rubinstein, E.; Raposo, G. The tetraspanin CD63 regulates ESCRT-independent and -dependent endosomal sorting during melanogenesis. *Dev. Cell* **2011**, *21*, 708–721.
73. Rodriguez-Boulan, E.; Kreitzer, G.; Musch, A. Organization of vesicular trafficking in epithelia. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 233–247.
74. Hutagalung, A.H.; Novick, P.J. Role of Rab GTPases in membrane traffic and cell physiology. *Physiol. Rev.* **2011**, *91*, 119–149.
75. Savina, A.; Vidal, M.; Colombo, M.I. The exosome pathway in K562 cells is regulated by Rab11. *J. Cell Sci.* **2002**, *115*, 2505–2515.
76. Ostrowski, M.; Carmo, N.B.; Krumeich, S.; Fanget, I.; Raposo, G.; Savina, A.; Moita, C.F.; Schauer, K.; Hume, A.N.; Freitas, R.P.; *et al.* Rab27a and Rab27b control different steps of the exosome secretion pathway. *Nat. Cell Biol.* **2010**, *12*, 19–30.
77. Yu, X.; Harris, S.L.; Levine, A.J. The regulation of exosome secretion: A novel function of the p53 protein. *Cancer Res.* **2006**, *66*, 4795–4801.
78. King, H.W.; Michael, M.Z.; Gleadle, J.M. Hypoxic enhancement of exosome release by breast cancer cells. *BMC Cancer* **2012**, *12*, 421.
79. Thompson, C.A.; Purushothaman, A.; Ramani, V.C.; Vlodavsky, I.; Sanderson, R.D. Heparanase Regulates Secretion, Composition, and Function of Tumor Cell-derived Exosomes. *J. Biol. Chem.* **2013**, *288*, 10093–10099.
80. Koumangoye, R.B.; Sakwe, A.M.; Goodwin, J.S.; Patel, T.; Ochieng, J. Detachment of breast tumor cells induces rapid secretion of exosomes which subsequently mediate cellular adhesion and spreading. *PLoS One* **2011**, *6*, e24234.
81. Llorente, A.; van Deurs, B.; Sandvig, K. Cholesterol regulates prostasome release from secretory lysosomes in PC-3 human prostate cancer cells. *Eur. J. Cell Biol.* **2007**, *86*, 405–415.
82. Faure, J.; Lachenal, G.; Court, M.; Hirrlinger, J.; Chatellard-Causse, C.; Blot, B.; Grange, J.; Schoehn, G.; Goldberg, Y.; Boyer, V.; *et al.* Exosomes are released by cultured cortical neurones. *Mol. Cell. Neurosci.* **2006**, *31*, 642–648.
83. Parolini, I.; Federici, C.; Raggi, C.; Lugini, L.; Palleschi, S.; de Milito, A.; Coscia, C.; Iessi, E.; Logozzi, M.; Molinari, A.; *et al.* Microenvironmental pH is a key factor for exosome traffic in tumor cells. *J. Biol. Chem.* **2009**, *284*, 34211–34222.
84. Mittelbrunn, M.; Gutierrez-Vazquez, C.; Villarroya-Beltri, C.; Gonzalez, S.; Sanchez-Cabo, F.; Gonzalez, M.A.; Bernad, A.; Sanchez-Madrid, F. Unidirectional transfer of microRNA-loaded exosomes from T cells to antigen-presenting cells. *Nat. Commun.* **2011**, *2*, 282.
85. Keller, S.; Konig, A.K.; Marme, F.; Runz, S.; Wolterink, S.; Koensgen, D.; Mustea, A.; Sehouli, J.; Altevogt, P. Systemic presence and tumor-growth promoting effect of ovarian carcinoma released exosomes. *Cancer Lett.* **2009**, *278*, 73–81.
86. Barres, C.; Blanc, L.; Bette-Bobillo, P.; Andre, S.; Mamoun, R.; Gabius, H.J.; Vidal, M. Galectin-5 is bound onto the surface of rat reticulocyte exosomes and modulates vesicle uptake by macrophages. *Blood* **2010**, *115*, 696–705.
87. Mathivanan, S.; Fahner, C.J.; Reid, G.E.; Simpson, R.J. ExoCarta 2012: Database of exosomal proteins, RNA and lipids. *Nucleic Acids Res.* **2012**, *40*, D1241–D1244.

88. Mathivanan, S.; Simpson, R.J. ExoCarta: A compendium of exosomal proteins and RNA. *Proteomics* **2009**, *9*, 4997–5000.
89. Subra, C.; Grand, D.; Laulagnier, K.; Stella, A.; Lambeau, G.; Paillasse, M.; de Medina, P.; Monsarrat, B.; Perret, B.; Silvente-Poirot, S.; *et al.* Exosomes account for vesicle-mediated transcellular transport of activatable phospholipases and prostaglandins. *J. Lipid Res.* **2010**, *51*, 2105–2120.
90. Zakharova, L.; Svetlova, M.; Fomina, A.F. T cell exosomes induce cholesterol accumulation in human monocytes via phosphatidylserine receptor. *J. Cell. Physiol.* **2007**, *212*, 174–181.
91. Xiang, X.; Poliakov, A.; Liu, C.; Liu, Y.; Deng, Z.B.; Wang, J.; Cheng, Z.; Shah, S.V.; Wang, G.J.; Zhang, L.; *et al.* Induction of myeloid-derived suppressor cells by tumor exosomes. *Int. J. Cancer* **2009**, *124*, 2621–2633.
92. Cao, W.; Ma, Z.; Rasenick, M.M.; Yeh, S.; Yu, J. N-3 poly-unsaturated fatty acids shift estrogen signaling to inhibit human breast cancer cell growth. *PLoS One* **2012**, *7*, e52838.
93. Valadi, H.; Ekstrom, K.; Bossios, A.; Sjostrand, M.; Lee, J.J.; Lotvall, J.O. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **2007**, *9*, 654–659.
94. Hessvik, N.P.; Phuyal, S.; Brech, A.; Sandvig, K.; Llorente, A. Profiling of microRNAs in exosomes released from PC-3 prostate cancer cells. *Biochim. Biophys. Acta* **2012**, *1819*, 1154–1163.
95. Bellingham, S.A.; Coleman, B.M.; Hill, A.F. Small RNA deep sequencing reveals a distinct miRNA signature released in exosomes from prion-infected neuronal cells. *Nucleic Acids Res.* **2012**, *40*, 10937–10949.
96. Rabinowits, G.; Gercel-Taylor, C.; Day, J.M.; Taylor, D.D.; Kloecker, G.H. Exosomal microRNA: A diagnostic marker for lung cancer. *Clin. Lung Cancer* **2009**, *10*, 42–46.
97. Pigati, L.; Yaddanapudi, S.C.; Iyengar, R.; Kim, D.J.; Hearn, S.A.; Danforth, D.; Hastings, M.L.; Duelli, D.M. Selective release of microRNA species from normal and malignant mammary epithelial cells. *PLoS One* **2010**, *5*, e13515.
98. Jaiswal, R.; Luk, F.; Gong, J.; Mathys, J.M.; Grau, G.E.; Bebawy, M. Microparticle conferred microRNA profiles—Implications in the transfer and dominance of cancer traits. *Mol. Cancer* **2012**, *11*, 37.
99. Xiao, D.; Ohlendorf, J.; Chen, Y.; Taylor, D.D.; Rai, S.N.; Waigel, S.; Zacharias, W.; Hao, H.; McMasters, K.M. Identifying mRNA, microRNA and protein profiles of melanoma exosomes. *PLoS One* **2012**, *7*, e46874.
100. Simpson, R.J.; Jensen, S.S.; Lim, J.W. Proteomic profiling of exosomes: Current perspectives. *Proteomics* **2008**, *8*, 4083–4099.
101. de Gassart, A.; Geminard, C.; Fevrier, B.; Raposo, G.; Vidal, M. Lipid raft-associated protein sorting in exosomes. *Blood* **2003**, *102*, 4336–4344.
102. Thery, C.; Amigorena, S.; Raposo, G.; Clayton, A. Isolation and characterization of exosomes from cell culture supernatants and biological fluids. *Curr. Protoc. Cell Biol.* **2006**, doi:10.1002/0471143030.cb0322s30.

103. Cantin, R.; Diou, J.; Belanger, D.; Tremblay, A.M.; Gilbert, C. Discrimination between exosomes and HIV-1: Purification of both vesicles from cell-free supernatants. *J. Immunol. Methods* **2008**, *338*, 21–30.
104. Lamparski, H.G.; Metha-Damani, A.; Yao, J.Y.; Patel, S.; Hsu, D.H.; Ruegg, C.; Le Pecq, J.B. Production and characterization of clinical grade exosomes derived from dendritic cells. *J. Immunol. Methods* **2002**, *270*, 211–226.
105. System Biosciences. Available online: http://www.systembio.com/exoquick (accessed on 15 April 2013).
106. Taylor, D.D.; Atay, S.; Metzinger, D.S.; Gercel-Taylor, C. Characterization of humoral responses of ovarian cancer patients: Antibody subclasses and antigenic components. *Gynecol. Oncol.* **2010**, *116*, 213–221.
107. Koga, K.; Matsumoto, K.; Akiyoshi, T.; Kubo, M.; Yamanaka, N.; Tasaki, A.; Nakashima, H.; Nakamura, M.; Kuroki, S.; Tanaka, M.; Katano, M. Purification, characterization and biological significance of tumor-derived exosomes. *Anticancer Res.* **2005**, *25*, 3703–3707.
108. Tauro, B.J.; Greening, D.W.; Mathias, R.A.; Ji, H.; Mathivanan, S.; Scott, A.M.; Simpson, R.J. Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes. *Methods* **2012**, *56*, 293–304.
109. Taylor, D.D.; Zacharias, W.; Gercel-Taylor, C. Exosome isolation for proteomic analyses and RNA profiling. *Methods Mol. Biol.* **2011**, *728*, 235–246.
110. Simona, F.; Laura, S.; Simona, T.; Riccardo, A. Contribution of proteomics to understanding the role of tumor-derived exosomes in cancer progression: State of the art and new perspectives. *Proteomics* **2013**, *13*, 1581–1594.
111. Eldh, M.; Lotvall, J.; Malmhall, C.; Ekstrom, K. Importance of RNA isolation methods for analysis of exosomal RNA: Evaluation of different methods. *Mol. Immunol.* **2012**, *50*, 278–286.
112. Lasser, C.; Eldh, M.; Lotvall, J. Isolation and characterization of RNA-containing exosomes. *J. Vis. Exp.* **2012**, e3037.
113. Lasser, C. Identification and analysis of circulating exosomal microRNA in human body fluids. *Methods Mol. Biol.* **2013**, *1024*, 109–128.
114. Huang, X.; Yuan, T.; Tschannen, M.; Sun, Z.; Jacob, H.; Du, M.; Liang, M.; Dittmar, R.L.; Liu, Y.; Liang, M.; *et al.* Characterization of human plasma-derived exosomal RNAs by deep sequencing. *BMC Genomics* **2013**, *14*, 319.
115. Nanosight. Available online: http://www.nanosight.com (accessed on 15 April 2013).
116. Zheng, Y.; Campbell, E.C.; Lucocq, J.; Riches, A.; Powis, S.J. Monitoring the Rab27 associated exosome pathway using nanoparticle tracking analysis. *Exp. Cell Res.* **2012**, doi:10.1016/j.yexcr.2012.10.006.
117. Hanahan, D.; Coussens, L.M. Accessories to the crime: Functions of cells recruited to the tumor microenvironment. *Cancer Cell* **2012**, *21*, 309–322.
118. Tian, T.; Wang, Y.; Wang, H.; Zhu, Z.; Xiao, Z. Visualizing of the cellular uptake and intracellular trafficking of exosomes by live-cell microscopy. *J. Cell. Biochem.* **2010**, *111*, 488–496.

119. Ristorcelli, E.; Beraud, E.; Verrando, P.; Villard, C.; Lafitte, D.; Sbarra, V.; Lombardo, D.; Verine, A. Human tumor nanoparticles induce apoptosis of pancreatic cancer cells. *FASEB J.* **2008**, *22*, 3358–3369.
120. Yang, Y.; Xiu, F.; Cai, Z.; Wang, J.; Wang, Q.; Fu, Y.; Cao, X. Increased induction of antitumor response by exosomes derived from interleukin-2 gene-modified tumor cells. *J. Cancer Res. Clin. Oncol.* **2007**, *133*, 389–399.
121. Dai, S.; Wei, D.; Wu, Z.; Zhou, X.; Wei, X.; Huang, H.; Li, G. Phase I clinical trial of autologous ascites-derived exosomes combined with GM-CSF for colorectal cancer. *Mol. Ther.* **2008**, *16*, 782–790.
122. Zhang, Y.; Wu, X.H.; Luo, C.L.; Zhang, J.M.; He, B.C.; Chen, G. Interleukin-12-anchored exosomes increase cytotoxicity of T lymphocytes by reversing the JAK/STAT pathway impaired by tumor-derived exosomes. *Int. J. Mol. Med.* **2010**, *25*, 695–700.
123. Zhang, Y.; Luo, C.L.; He, B.C.; Zhang, J.M.; Cheng, G.; Wu, X.H. Exosomes derived from IL-12-anchored renal cancer cells increase induction of specific antitumor response *in vitro*: A novel vaccine for renal cell carcinoma. *Int. J. Oncol.* **2010**, *36*, 133–140.
124. Chen, T.; Guo, J.; Yang, M.; Zhu, X.; Cao, X. Chemokine-containing exosomes are released from heat-stressed tumor cells via lipid raft-dependent pathway and act as efficient tumor vaccine. *J. Immunol.* **2011**, *186*, 2219–2228.
125. Peng, P.; Yan, Y.; Keng, S. Exosomes in the ascites of ovarian cancer patients: Origin and effects on anti-tumor immunity. *Oncol. Rep.* **2011**, *25*, 749–762.
126. Hakulinen, J.; Sankkila, L.; Sugiyama, N.; Lehti, K.; Keski-Oja, J. Secretion of active membrane type 1 matrix metalloproteinase (MMP-14) into extracellular space in microvesicular exosomes. *J. Cell. Biochem.* **2008**, *105*, 1211–1218.
127. McCready, J.; Sims, J.D.; Chan, D.; Jay, D.G. Secretion of extracellular hsp90alpha via exosomes increases cancer cell motility: A role for plasminogen activation. *BMC Cancer* **2010**, *10*, 294.
128. Janowska-Wieczorek, A.; Wysoczynski, M.; Kijowski, J.; Marquez-Curtis, L.; Machalinski, B.; Ratajczak, J.; Ratajczak, M.Z. Microvesicles derived from activated platelets induce metastasis and angiogenesis in lung cancer. *Int. J. Cancer* **2005**, *113*, 752–760.
129. Cho, J.A.; Park, H.; Lim, E.H.; Kim, K.H.; Choi, J.S.; Lee, J.H.; Shin, J.W.; Lee, K.W. Exosomes from ovarian cancer cells induce adipose tissue-derived mesenchymal stem cells to acquire the physical and functional characteristics of tumor-supporting myofibroblasts. *Gynecol. Oncol.* **2011**, *123*, 379–386.
130. Bobrie, A.; Krumeich, S.; Reyal, F.; Recchi, C.; Moita, L.F.; Seabra, M.C.; Ostrowski, M.; Thery, C. Rab27a supports exosome-dependent and -independent mechanisms that modify the tumor microenvironment and can promote tumor progression. *Cancer Res.* **2012**, *72*, 4920–4930.
131. Park, J.E.; Tan, H.S.; Datta, A.; Lai, R.C.; Zhang, H.; Meng, W.; Lim, S.K.; Sze, S.K. Hypoxic tumor cell modulates its microenvironment to enhance angiogenic and metastatic potential by secretion of proteins and exosomes. *Mol. Cell. Proteomics* **2010**, *9*, 1085–1099.

132. Kucharzewska, P.; Christianson, H.C.; Welch, J.E.; Svensson, K.J.; Fredlund, E.; Ringner, M.; Morgelin, M.; Bourseau-Guilmain, E.; Bengzon, J.; Belting, M. Exosomes reflect the hypoxic status of glioma cells and mediate hypoxia-dependent activation of vascular cells during tumor development. *Proc. Natl. Acad. Sci. USA* **2013**, doi:10.1073/pnas.1220998110.
133. Hong, B.S.; Cho, J.H.; Kim, H.; Choi, E.J.; Rho, S.; Kim, J.; Kim, J.H.; Choi, D.S.; Kim, Y.K.; Hwang, D.; *et al.* Colorectal cancer cell-derived microvesicles are enriched in cell cycle-related mRNAs that promote proliferation of endothelial cells. *BMC Genomics* **2009**, *10*, 556.
134. Ji, H.; Greening, D.W.; Barnes, T.W.; Lim, J.W.; Tauro, B.J.; Rai, A.; Xu, R.; Adda, C.; Mathivanan, S.; Zhao, W.; *et al.* Proteome profiling of exosomes derived from human primary and metastatic colorectal cells reveal differential expression of key metastatic factors and signal transduction components. *Proteomics* **2013**, *13*, 1672–1686.
135. Grange, C.; Tapparo, M.; Collino, F.; Vitillo, L.; Damasco, C.; Deregibus, M.C.; Tetta, C.; Bussolati, B.; Camussi, G. Microvesicles released from human renal cancer stem cells stimulate angiogenesis and formation of lung premetastatic niche. *Cancer Res.* **2011**, *71*, 5346–5356.
136. Zhu, W.; Huang, L.; Li, Y.; Zhang, X.; Gu, J.; Yan, Y.; Xu, X.; Wang, M.; Qian, H.; Xu, W. Exosomes derived from human bone marrow mesenchymal stem cells promote tumor growth *in vivo*. *Cancer Lett.* **2012**, *315*, 28–37.
137. Mineo, M.; Garfield, S.H.; Taverna, S.; Flugy, A.; De Leo, G.; Alessandro, R.; Kohn, E.C. Exosomes released by K562 chronic myeloid leukemia cells promote angiogenesis in a Src-dependent fashion. *Angiogenesis* **2012**, *15*, 33–45.
138. Peinado, H.; Aleckovic, M.; Lavotshkin, S.; Matei, I.; Costa-Silva, B.; Moreno-Bueno, G.; Hergueta-Redondo, M.; Williams, C.; Garcia-Santos, G.; Ghajar, C.; *et al.* Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nat. Med.* **2012**, *18*, 883–891.
139. Baer, C.; Squadrito, M.L.; Iruela-Arispe, M.L.; De Palma, M. Reciprocal interactions between endothelial cells and macrophages in angiogenic vascular niches. *Exp. Cell Res.* **2013**, doi:10.1016/j.yexcr.2013.03.026.
140. Cai, X.; Hagedorn, C.H.; Cullen, B.R. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* **2004**, *10*, 1957–1966.
141. Lee, Y.; Jeon, K.; Lee, J.T.; Kim, S.; Kim, V.N. MicroRNA maturation: Stepwise processing and subcellular localization. *EMBO J.* **2002**, *21*, 4663–4670.
142. Lee, Y.; Kim, M.; Han, J.; Yeom, K.H.; Lee, S.; Baek, S.H.; Kim, V.N. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **2004**, *23*, 4051–4060.
143. Zeng, Y.; Cullen, B.R. Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic Acids Res.* **2004**, *32*, 4776–4785.
144. Bartel, D.P. MicroRNAs: Target recognition and regulatory functions. *Cell* **2009**, *136*, 215–233.
145. Griffiths-Jones, S. The microRNA registry. *Nucleic Acids Res.* **2004**, *32*, D109–D111.
146. Lim, L.P.; Lau, N.C.; Garrett-Engele, P.; Grimson, A.; Schelter, J.M.; Castle, J.; Bartel, D.P.; Linsley, P.S.; Johnson, J.M. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **2005**, *433*, 769–773.
147. Taft, R.J.; Pang, K.C.; Mercer, T.R.; Dinger, M.; Mattick, J.S. Non-coding RNAs: Regulators of disease. *J. Pathol.* **2010**, *220*, 126–139.

148. Melo, S.A.; Esteller, M. Dysregulation of microRNAs in cancer: Playing with fire. *FEBS Lett.* **2011**, *585*, 2087–2099.
149. Lu, J.; Getz, G.; Miska, E.A.; Alvarez-Saavedra, E.; Lamb, J.; Peck, D.; Sweet-Cordero, A.; Ebert, B.L.; Mak, R.H.; Ferrando, A.A.; *et al.* microRNA expression profiles classify human cancers. *Nature* **2005**, *435*, 834–838.
150. Guo, Y.; Chen, Z.; Zhang, L.; Zhou, F.; Shi, S.; Feng, X.; Li, B.; Meng, X.; Ma, X.; Luo, M.; *et al.* Distinctive microRNA profiles relating to patient survival in esophageal squamous cell carcinoma. *Cancer Res.* **2008**, *68*, 26–33.
151. Gibbings, D.J.; Ciaudo, C.; Erhardt, M.; Voinnet, O. Multivesicular bodies associate with components of miRNA effector complexes and modulate miRNA activity. *Nat. Cell Biol.* **2009**, *11*, 1143–1149.
152. Kosaka, N.; Iguchi, H.; Yoshioka, Y.; Takeshita, F.; Matsuki, Y.; Ochiya, T. Secretory mechanisms and intercellular transfer of microRNAs in living cells. *J. Biol. Chem.* **2010**, *285*, 17442–17452.
153. Yao, B.; La, L.B.; Chen, Y.C.; Chang, L.J.; Chan, E.K. Defining a new role of GW182 in maintaining miRNA stability. *EMBO Rep.* **2012**, *13*, 1102–1108.
154. Ohshima, K.; Inoue, K.; Fujiwara, A.; Hatakeyama, K.; Kanto, K.; Watanabe, Y.; Muramatsu, K.; Fukuda, Y.; Ogura, S.; Yamaguchi, K.; *et al.* Let-7 microRNA family is selectively secreted into the extracellular environment via exosomes in a metastatic gastric cancer cell line. *PLoS One* **2010**, *5*, e13247.
155. Cifuentes, D.; Xue, H.; Taylor, D.W.; Patnode, H.; Mishima, Y.; Cheloufi, S.; Ma, E.; Mane, S.; Hannon, G.J.; Lawson, N.D.; *et al.* A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* **2010**, *328*, 1694–1698.
156. Montecalvo, A.; Larregina, A.T.; Shufesky, W.J.; Stolz, D.B.; Sullivan, M.L.; Karlsson, J.M.; Baty, C.J.; Gibson, G.A.; Erdos, G.; Wang, Z.; *et al.* Mechanism of transfer of functional microRNAs between mouse dendritic cells via exosomes. *Blood* **2012**, *119*, 756–766.
157. Kogure, T.; Lin, W.L.; Yan, I.K.; Braconi, C.; Patel, T. Intercellular nanovesicle-mediated microRNA transfer: A mechanism of environmental modulation of hepatocellular cancer cell growth. *Hepatology* **2011**, *54*, 1237–1248.
158. Bryant, R.J.; Pawlowski, T.; Catto, J.W.; Marsden, G.; Vessella, R.L.; Rhees, B.; Kuslich, C.; Visakorpi, T.; Hamdy, F.C. Changes in circulating microRNA levels associated with prostate cancer. *Br. J. Cancer* **2012**, *106*, 768–774.
159. Takeshita, N.; Hoshino, I.; Mori, M.; Akutsu, Y.; Hanari, N.; Yoneyama, Y.; Ikeda, N.; Isozaki, Y.; Maruyama, T.; Akanuma, N.; *et al.* Serum microRNA expression profile: miR-1246 as a novel diagnostic and prognostic biomarker for oesophageal squamous cell carcinoma. *Br. J. Cancer* **2013**, *108*, 644–652.
160. Russo, F.; di Bella, S.; Nigita, G.; Macca, V.; Lagana, A.; Giugno, R.; Pulvirenti, A.; Ferro, A. miRandola: Extracellular circulating microRNAs database. *PLoS One* **2012**, *7*, e47786.
161. Umezu, T.; Ohyashiki, K.; Kuroda, M.; Ohyashiki, J.H. Leukemia cell to endothelial cell communication via exosomal miRNAs. *Oncogene* **2012**, *32*, 2747–2755.
162. Yang, M.; Chen, J.; Su, F.; Yu, B.; Su, F.; Lin, L.; Liu, Y.; Huang, J.D.; Song, E. Microvesicles secreted by macrophages shuttle invasion-potentiating microRNAs into breast cancer cells. *Mol. Cancer* **2011**, *10*, 117.

163. Zhou, R.; O'Hara, S.P.; Chen, X.M. MicroRNA regulation of innate immune responses in epithelial cells. *Cell. Mol. Immunol.* **2011**, *8*, 371–379.
164. Fabbri, M.; Paone, A.; Calore, F.; Galli, R.; Gaudio, E.; Santhanam, R.; Lovat, F.; Fadda, P.; Mao, C.; Nuovo, G.J.; *et al.* MicroRNAs bind to Toll-like receptors to induce prometastatic inflammatory response. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E2110–E2116.
165. Palma, J.; Yaddanapudi, S.C.; Pigati, L.; Havens, M.A.; Jeong, S.; Weiner, G.A.; Weimer, K.M.; Stern, B.; Hastings, M.L.; Duelli, D.M. MicroRNAs are exported from malignant cells in customized particles. *Nucleic Acids Res.* **2012**, *40*, 9125–9138.
166. Rana, S.; Malinowska, K.; Zoller, M. Exosomal tumor microRNA modulates premetastatic organ cells. *Neoplasia* **2013**, *15*, 281–295.
167. Quesenberry, P.J.; Aliotta, J.M. Cellular phenotype switching and microvesicles. *Adv. Drug Deliv. Rev.* **2010**, *62*, 1141–1148.
168. Weigelt, B.; Bissell, M.J. Unraveling the microenvironmental influences on the normal mammary gland and breast cancer. *Semin. Cancer Biol.* **2008**, *18*, 311–321.
169. Section on Breastfeeding. Breastfeeding and the use of human milk. *Pediatrics* **2012**, *129*, e827–e841.
170. Zhou, Q.; Li, M.; Wang, X.; Li, Q.; Wang, T.; Zhu, Q.; Zhou, X.; Wang, X.; Gao, X.; Li, X. Immune-related microRNAs are abundant in breast milk exosomes. *Int. J. Biol. Sci.* **2012**, *8*, 118–123.
171. Society, A.C. Breast Cancer Facts & Figures 2011–2012. Available online: http://www.cancer.org/research/cancerfactsfigures/breastcancerfactsfigures/breast-cancer-facts-and-figures-2011-2012 (accessed on 15 April 2013).
172. Locker, A.; Horrocks, C.; Gilmour, A.; Ellis, I.; Dowle, C.; Elston, C.; Blamey, R. Flow cytometric and histological analysis of ductal carcinoma in situ of the breast. *Br. J. Surg.* **1990**, *77*, 564–567.
173. Ottesen, G.; Christensen, I.; Larsen, J.; Christiansen, J.; Hansen, B.; Andersen, J. DNA analysis of in situ ductal carcinoma of the breast via flow cytometry. *Cytometry* **1995**, *22*, 168–176.
174. Leal, C.; Schmitt, F.; Bento, M.; Maia, N.; Lopes, C. Ductal carcinoma in situ of the breast. Histologic categorization and its relationship to ploidy and immunohistochemical expression of hormone receptors, p53, and c-erbB-2 protein. *Cancer* **1995**, *75*, 2123–2131.
175. Maguire, H.; Hellman, M.; Greene, M.; Yeh, I. Expression of c-erbB-2 in in situ and in adjacent invasive ductal adenocarcinomas of the female breast. *Pathobiology* **1992**, *60*, 117–121.
176. Vos, C.; Ter Haar, N.; Peterse, J.; Cornelisse, C.; van de Vijver, M. Cyclin D1 gene amplification and overexpression are present in ductal carcinoma in situ of the breast. *J. Pathol.* **1999**, *187*, 279–284.
177. Radford, D.; Fair, K.; Phillips, N.; Ritter, J.; Steinbrueck, T.; Holt, M.; Donis-Keller, H. Allelotyping of ductal carcinoma in situ of the breast: Deletion of loci on 8p, 13q, 16q, 17p and 17q. *Cancer Res.* **1995**, *55*, 3399–3405.
178. Radford, D.; Phillips, N.; Fair, K.; Ritter, J.; Holt, M.; Donis-Keller, H. Allelic loss and the progression of breast cancer. *Cancer Res.* **1995**, *55*, 5180–5183.
179. Ma, X.J.; Dahiya, S.; Richardson, E.; Erlander, M.; Sgroi, D.C. Gene expression profiling of the tumor microenvironment during breast cancer progression. *Breast Cancer Res.* **2009**, *11*, R7.

180. Liu, C.; Yu, S.; Zinn, K.; Wang, J.; Zhang, L.; Jia, Y.; Kappes, J.C.; Barnes, S.; Kimberly, R.P.; Grizzle, W.E.; *et al.* Murine mammary carcinoma exosomes promote tumor growth by suppression of NK cell function. *J. Immunol.* **2006**, *176*, 1375–1385.
181. Higginbotham, J.N.; Demory Beckler, M.; Gephart, J.D.; Franklin, J.L.; Bogatcheva, G.; Kremers, G.J.; Piston, D.W.; Ayers, G.D.; McConnell, R.E.; Tyska, M.J.; *et al.* Amphiregulin exosomes increase cancer cell invasion. *Curr. Biol.* **2011**, *21*, 779–786.
182. O'Brien, K.; Rani, S.; Corcoran, C.; Wallace, R.; Hughes, L.; Friel, A.M.; McDonnell, S.; Crown, J.; Radomski, M.W.; O'Driscoll, L. Exosomes from triple-negative breast cancer cells can transfer phenotypic traits representing their cells of origin to secondary cells. *Eur. J. Cancer* **2013**, *49*, 1845–1859.
183. Luga, V.; Zhang, L.; Viloria-Petit, A.M.; Ogunjimi, A.A.; Inanlou, M.R.; Chiu, E.; Buchanan, M.; Hosein, A.N.; Basik, M.; Wrana, J.L. Exosomes mediate stromal mobilization of autocrine Wnt-PCP signaling in breast cancer cell migration. *Cell* **2012**, *151*, 1542–1556.
184. Suetsugu, A.; Honma, K.; Saji, S.; Moriwaki, H.; Ochiya, T.; Hoffman, R.M. Imaging exosome transfer from breast cancer cells to stroma at metastatic sites in orthotopic nude-mouse models. *Adv. Drug Deliv. Rev.* **2013**, *65*, 383–390.
185. Sotiriou, C.; Pusztai, L. Gene-expression signatures in breast cancer. *N. Engl. J. Med.* **2009**, *360*, 790–800.
186. Blenkiron, C.; Goldstein, L.D.; Thorne, N.P.; Spiteri, I.; Chin, S.F.; Dunning, M.J.; Barbosa-Morais, N.L.; Teschendorff, A.E.; Green, A.R.; Ellis, I.O.; *et al.* MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. *Genome Biol.* **2007**, *8*, R214.
187. Perou, C.M.; Sorlie, T.; Eisen, M.B.; van de Rijn, M.; Jeffrey, S.S.; Rees, C.A.; Pollack, J.R.; Ross, D.T.; Johnsen, H.; Akslen, L.A.; *et al.* Molecular portraits of human breast tumours. *Nature* **2000**, *406*, 747–752.
188. Ciravolo, V.; Huber, V.; Ghedini, G.C.; Venturelli, E.; Bianchi, F.; Campiglio, M.; Morelli, D.; Villa, A.; Della Mina, P.; Menard, S.; *et al.* Potential role of HER2-overexpressing exosomes in countering trastuzumab-based therapy. *J. Cell. Physiol.* **2012**, *227*, 658–667.
189. Marleau, A.M.; Chen, C.S.; Joyce, J.A.; Tullis, R.H. Exosome removal as a therapeutic adjuvant in cancer. *J. Transl. Med.* **2012**, *10*, 134.
190. Safaei, R.; Larson, B.J.; Cheng, T.C.; Gibson, M.A.; Otani, S.; Naerdemann, W.; Howell, S.B. Abnormal lysosomal trafficking and enhanced exosomal export of cisplatin in drug-resistant human ovarian carcinoma cells. *Mol. Cancer Ther.* **2005**, *4*, 1595–1604.
191. Ohno, S.; Takanashi, M.; Sudo, K.; Ueda, S.; Ishikawa, A.; Matsuyama, N.; Fujita, K.; Mizutani, T.; Ohgi, T.; Ochiya, T.; *et al.* Systemically injected exosomes targeted to EGFR deliver antitumor microRNA to breast cancer cells. *Mol. Ther.* **2013**, *21*, 185–191.
192. Kooijmans, S.A.; Vader, P.; van Dommelen, S.M.; van Solinge, W.W.; Schiffelers, R.M. Exosome mimetics: A novel class of drug delivery systems. *Int. J. Nanomedicine* **2012**, *7*, 1525–1541.

Reprinted from *IJMS*. Cite as: Lee, C.-H.; Kuo, W.-H.; Lin, C-C.; Oyang, Y.-J.; Huang, H.-C.; Juan, H.-F. MicroRNA-Regulated Protein-Protein Interaction Networks and Their Functions in Breast Cancer. *Int. J. Mol. Sci.* **2013**, *14*, 11560-11606.

Article

MicroRNA-Regulated Protein-Protein Interaction Networks and Their Functions in Breast Cancer

Chia-Hsien Lee [1,†], Wen-Hong Kuo [2,†], Chen-Ching Lin [1], Yen-Jen Oyang [1], Hsuan-Cheng Huang [3,*] and Hsueh-Fen Juan [1,4,*]

[1] Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei 106, Taiwan; E-Mails: tfglove@gmail.com (C.-H.L.); kdragongo@gmail.com (C.-C.L.); yjoyang@csie.ntu.edu.tw (Y.-J.O.)

[2] Department of Physiology, College of Medicine, National Taiwan University, Taipei 100, Taiwan; E-Mail: brcancer@gmail.com

[3] Institute of Biomedical Informatics and Center for Systems and Synthetic Biology, National Yang-Ming University, Taipei 112, Taiwan

[4] Institute of Molecular and Cellular Biology and Department of Life Science, National Taiwan University, Taipei 106, Taiwan

† These authors contributed equally to this work.

* Authors to whom correspondence should be addressed; E-Mails: hsuancheng@ym.edu.tw (H.-C.H.); yukijuan@ntu.edu.tw (H.-F.J.); Tel.: +886-2-2826-7357 (H.-C.H.); +886-2-3366-4536 (H.-F.J.); Fax: +886-2-2820-2508 (H.-C.H.); +886-2-2367-3374 (H.-F.J.).

Received: 17 April 2013; in revised form: 21 May 2013 / Accepted: 22 May 2013 / Published: 30 May 2013

Abstract: MicroRNAs, which are small endogenous RNA regulators, have been associated with various types of cancer. Breast cancer is a major health threat for women worldwide. Many miRNAs were reported to be associated with the progression and carcinogenesis of breast cancer. In this study, we aimed to discover novel breast cancer-related miRNAs and to elucidate their functions. First, we identified confident miRNA-target pairs by combining data from miRNA target prediction databases and expression profiles of miRNA and mRNA. Then, miRNA-regulated protein interaction networks (PINs) were constructed with confident pairs and known interaction data in the human protein reference database (HPRD). Finally, the functions of miRNA-regulated PINs were elucidated by functional enrichment analysis. From the results, we identified some previously reported breast cancer-related miRNAs and functions of the PINs, e.g., miR-125b, miR-125a,

miR-21, and miR-497. Some novel miRNAs without known association to breast cancer were also found, and the putative functions of their PINs were also elucidated. These include miR-139 and miR-383. Furthermore, we validated our results by receiver operating characteristic (ROC) curve analysis using our miRNA expression profile data, gene expression-based outcome for breast cancer online (GOBO) survival analysis, and a literature search. Our results may provide new insights for research in breast cancer-associated miRNAs.

Keywords: miRNA; breast cancer; protein interaction network; functional analysis

1. Introduction

Breast cancer is a global health threat for women. According to a 2008 survey [1], breast cancer was the leading cause of cancer deaths in women. Our knowledge of possible risk factors has led to developments in diagnostic methods, drugs, and surgery procedures for treatment [2,3]; however, the details of breast carcinoma progression, and perhaps most importantly, how to cure breast cancer, remain elusive.

Previous research has identified a number of risk factors for breast cancer. Early menarche, late menopause, obesity, late first full pregnancy, and hormone replacement therapy were considered as high risk factors for breast cancer [2]. Breast cancer risk has also been reported to be related to fat intake in diets rich in red meats and high-fat dairy foods [3].

MicroRNAs (miRNAs) [4], are short endogenous non-coding RNAs which are able to regulate gene expression. After miRNA precursors are transcribed from the genome or generated from spliceosomes, they are exported to the cytoplasm and further processed by the Dicer complex [5]. The mature miRNA is then bound to Argonaute protein, forming a miRNA-protein complex known as the RNA-induced silencing complex (RISC) [6], miRNP, or RNAi (RNA interference) enzyme complex [7,8]. The RISC has been reported to down-regulate target genes by translational repression [9] or mRNA cleavage [10].

Like other protein-based regulators, miRNAs have been associated with cancer. Calin *et al.*, reported that miR-15 and miR-16 were deleted in leukemia [11], which was believed to be one of the earliest reports associating miRNAs with cancer [12]. After this report, many miRNAs were found to act as tumor suppressors or oncogenes (also known as oncomirs). For example, miR-21 was identified as an oncomir in hepatocellular cancer [13], breast cancer [14], and kidney cancer [15]. On the other hand, let-7c was found to be a tumor suppressor in prostate cancer [16], and miR-181a was reported as a tumor suppressor in glioma [17]. Further, miR-125b [18], and miR-145 [19] were identified as tumor suppressors in breast cancer, and miR-125a was found to repress tumor growth in breast cancer [20]. Thus, it is highly likely that miRNAs play an important role in breast cancer.

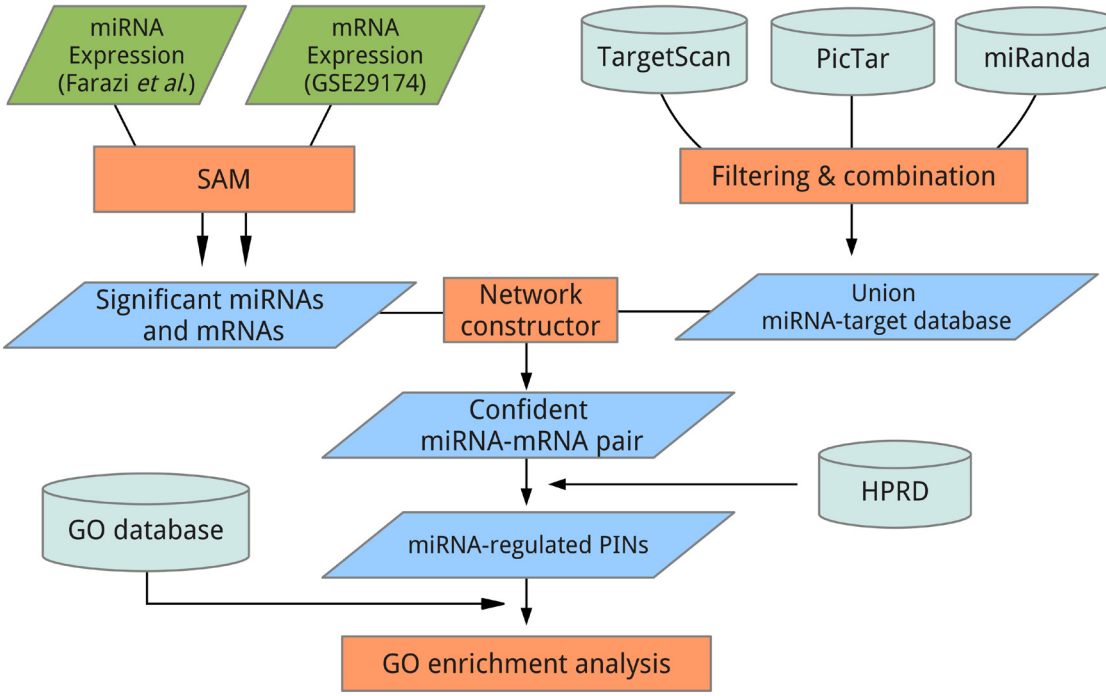
Since miRNA functions by regulating its target genes, we may deduce the effects of miRNAs by analyzing their regulated networks. To use such a method to elucidate miRNA functions, targets of miRNAs should be deduced. Currently, predictions in most target prediction database are based on sequence and statistical methods [21]. For example, in TargetScan, seed base pairing, target site context, conservation of target site and miRNA, and site accessibility are considered in the prediction process [22].

Another method to elucidate miRNA targets is to integrate expression profiles of miRNA and mRNA. In the work of Huang *et al.* [23], a Bayesian-based algorithm, GenMiR++, was developed to predict possible targets of 104 miRNAs in humans. They also verified their results by RT-PCR and microarray experiments. However, the power of other sequence-based target prediction algorithms was not utilized in their work.

It is also possible to combine sequence-based target prediction and expression-based target prediction methods. By integrating expression data into sequence-based predictions, possible false positives can be reduced. Previously, miRNA-mRNA interactions were explored with splitting-averaging Bayesian networks [24]. In that work, expression profiles of miRNA and mRNA from public databases, miRNA target prediction databases, and miRNA sequence information were integrated together to discover miRNA-mRNA interaction networks.

Here, we combined expression profiles of miRNA and mRNA, and three target prediction databases, TargetScan, PicTar and miRanda, to obtain confident miRNA-mRNA relationships and construct miRNA-regulated protein-protein networks for breast cancer. Furthermore, we explored the functions of the miRNAs by inspecting the underlying protein interaction networks (PINs) of the miRNAs with functional enrichment analysis. This method, as described in Figure 1, was used to elucidate the functions of gastric cancer-related miRNAs in our previous work [25]. In that study, a gastric cancer-associated miRNA, miR-148a, was identified and validated as being involved in tumor proliferation, invasion, migration, and the survival rate of the patients. By using a similar method, we aimed to elucidate breast cancer-related miRNA-regulated PINs and their functions.

Figure 1. Analysis flow chart used in this work. After expression profiles and target prediction databases were fetched and preprocessed, they were subjected to the analysis process described here and in the "Experimental Section".



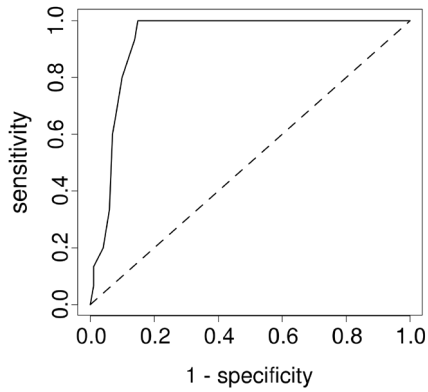
2. Results and Discussion

To construct miRNA-regulated PINs, differentially expressed miRNAs and genes from the dataset from Farazi *et al.* [26] were extracted following proper processing of the expression profiles. From our selected public miRNA dataset, we found 89 down-regulated miRNAs (93 prior to fold-change filtering) and only 1 up-regulated miRNA (Table S1). In gene expression dataset GSE29174, we found a total of 1268 down-regulated genes and 587 up-regulated genes before applying the fold change filter. There were 726 down-regulated genes (Table S2) and 437 up-regulated genes (Table S3) after significantly and differentially expressed genes were filtered by fold change (fold change >2).

From the results of SAM analysis, we identified some well-known breast cancer-related miRNAs (Table S1). For example, miR-214-3p [27] and miR-335-5p [28] have been previously reported to be down-regulated in breast cancer. Let-7c was found to be down-regulated in this work, while let-7a, another member of the let-7 family, was found to be down-regulated in another work [29]. MicroRNAs of the let-7 family were also reportedly down-regulated in several types of cancer [30]. We also found that miR-21-5p, the sole up-regulated miRNA in our list, was also previously found to be up-regulated [14,31]. However, changes in the expression of most of the miRNAs in our down-regulated list have not been reported in the literature. Therefore, we could not rule out the possibility that these miRNAs were novel breast cancer-related miRNAs. There are also some well-known miRNAs not presented in our list (for such a list, one may see [32–34]). The reason that some known miRNAs, for example, miR-19a, miR-155 and miR-205, did not show up in our result might be that we used a very stringent threshold (described in Experimental Section) when selecting differentially expressed miRNAs for PIN construction.

Since the miRNAs of the miRNA-regulated PINs were differentially expressed between normal and tumor tissues, and we identified some cancer-related functions in our functional enrichment analysis, the miRNAs may potentially be useful diagnostic markers for breast cancer. To verify this, we applied ROC curve analysis on the miRNA expression profile that was not used in constructing the miRNA-regulated PINs. Notably, our results (Figures 2 and S1 and Table S4) showed that let-7c (Figure 2), miR-497-5p, miR-125b-5p, and some other miRNAs of miRNA-regulated PINs, performed well when used as breast cancer diagnostic markers.

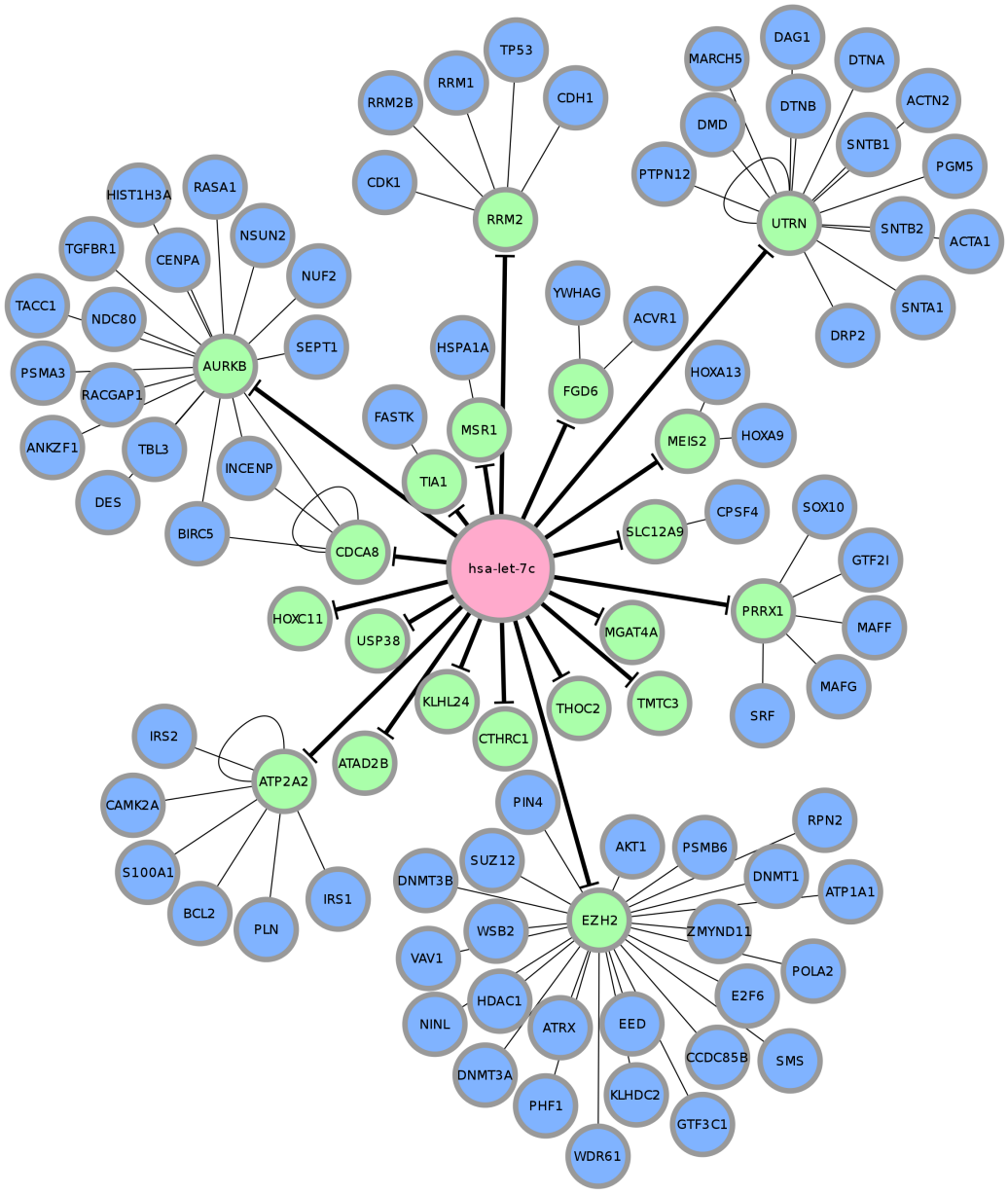
Figure 2. Receiver operating characteristic (ROC) curve of let-7c from our miRNA array dataset. For ROC curves of other miRNAs, see Figure S1.



Following elucidation of differentially expressed miRNAs and genes, miRNA-regulated PINs could then be constructed. We identified and constructed partial networks, containing the miRNA and its

direct target, with the differentially expressed miRNAs and genes as described in the "Experimental Section". We then extended the network by appending known interactions from the HPRD database. Finally, 18 miRNA-regulated PINs were constructed by the steps described above (Figure 3, Figures S2–S13, and Table S5).

Figure 3. The let-7c-regulated protein interaction network (PIN). This is one of the 18 miRNA-regulated PINs constructed in this work. Figures of all other miRNA-regulated PINs are displayed in Supplementary Figures S2–S13.



After construction of the 18 PINs was completed, we observed that the sizes of the PINs were not similar: some of the miRNAs seemed to regulate larger sized PINs, while other miRNAs affected only a small number of genes. Small miRNA-regulated PINs may be caused by the strict *q*-value threshold set during SAM analysis, the processing steps performed on the target prediction databases discussed previously, and possibly by lack of protein-protein interaction data for some proteins in the HPRD.

Although the HPRD may be considered the most comprehensive source of protein-protein interaction data [35], some proteins may not have been considered and researched by other investigators, and therefore, interaction data for those proteins would not be included in the HPRD. However, it may be true that some of the miRNA-regulated PINs were small in breast cancer, since the construction of the PINs were based on differentially expressed miRNAs and genes between normal tissues and tumor samples, and those miRNAs with small PINs may not be as important as others with larger PINs.

To elucidate the functions of a miRNA with a regulated PIN, GO enrichment analysis was applied to the miRNA-regulated PINs. We did not consider the miRNAs with ≤5 genes in their regulated PINs, and some of the miRNA-regulated PINs had no enriched functions using the defined threshold, FDR < 0.0001. To exclude GO terms that describe a broad range of concepts, we only included high level GO terms, *i.e.*, larger than 5.

The results of the GO enrichment analysis for let-7c related to cancer are listed in Table 1. (Results of all miRNA-regulated PINs were in Table S6). We defined a GO term as cancer-related if a GO term contained "cell proliferation", "cell death", "apoptosis", "signaling", "microtubule", and "actin". We noted that 7 miRNAs had enriched GO terms related to apoptosis, cell death, and cell proliferation, *i.e.*, miR-520d-3p, miR-497-5p, miR-125b-5p, miR-21-5p, miR-31-5p, let-7c, and miR-125-5p. Further, some miRNA-regulated PINs may have functions other than cell survival. For example, the nerve growth factor receptor pathway was enriched in miR-regulated PINs of miR-520d-3p, miR-497-5p, miR-125a-5p, miR-125b-5p, and miR-31-5p, and the epidermal growth factor receptor pathway was enriched in miR-regulated PINs of miR-520d-3p, miR-21-5p, and miR-497-5p. Most of the miRNAs had been previously described and were known to be implicated in breast cancer. Let-7c was not only likely to be down-regulated in breast cancer [29], but was also found to be a tumor suppressor in prostate cancer [16]. Another reported tumor suppressor was miR-125b-5p, which was found to be down-regulated in breast tumor tissue [18], and this finding was consistent with our functional enrichment results. The miR-125a-5p-regulated PIN was found to be able to inhibit apoptosis and regulate epithelial cell proliferation, and has been reported to repress cell growth [36].

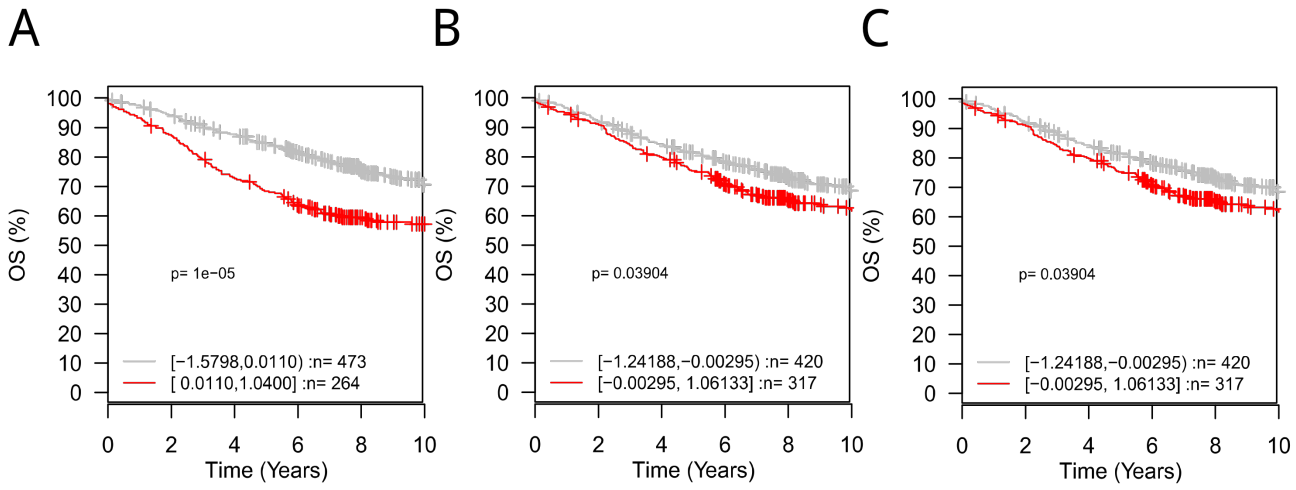
Table 1. Selected enriched functions of let-7c. Member genes of the let-7c-regulated network annotated with corresponding enriched functions are listed.

MIMAT0000064(hsa-let-7c)		
GO term	**Genes**	**Adj. *p*-value**
GO:0043067, Regulation of programmed cell death	RRM2B, ACVR1, TP53, RASA1, TGFBR1, PSMA3, BIRC5, ACTN2, HOXA13, IRS2, FASTK, VAV1, PSMB6, BCL2, CDK1, HDAC1, SOX10, TIA1, AKT1, AURKB	3.43×10^{-8}
GO:0043069, Negative regulation of programmed cell death	RRM2B, ACVR1, TP53, RASA1, TGFBR1, BIRC5, IRS2, BCL2, CDK1, HDAC1, SOX10, AKT1, AURKB	6.58×10^{-6}
GO:0060548, Negative regulation of cell death	RRM2B, ACVR1, TP53, RASA1, TGFBR1, BIRC5, IRS2, BCL2, CDK1, HDAC1, SOX10, AKT1, AURKB	8.92×10^{-6}
GO:0015630, Microtubule cytoskeleton	INCENP, SNTB2, SEPT1, TACC1, BIRC5, RACGAP1, PIN4, CDCA8, CDK1, PHF1, AKT1, AURKB, NINL, CCDC85B	5.15×10^{-5}

We also found that some breast cancer-specific functions were enriched in our results. For example, in the miR-497-5p-regulated PIN, the term "androgen receptor signaling pathway" was enriched. Although it is not clear whether androgens are related to breast cancer, androgen receptors are known to be up-regulated in breast cancer and related to node invasiveness [37].

To further verify the results of the enriched cancer-related functions, we used GOBO for survival analysis. Our hypothesis is that expression of genes annotated with enriched cancer-related terms may be related to survival outcome of patients. With exception to some cell death/proliferation-related terms, it is already known that some pathways or functions are also related to clinical outcomes. For example, cell proliferation-related GO terms have a high probability of affecting survival of cancer tissues, and patient outcome may worsen if cancer tissue survives. In addition, some signaling pathways were known to enhance invasiveness, migration abilities, or were associated with reduced patient survival. For example, the BMP signaling pathway is known to confer various tumor cells with enhanced migration and invasion abilities [38], and nerve growth factor receptor (NGFR) was found to be associated with overall survival of breast cancer [38]. Furthermore, Toll-like receptor 4 has been reported to promote adhesion and invasive migration in breast cancer [39]. Finally, the cytoskeleton plays an important role in regulating cell motility in all cells. Actin filaments are known to participate in the invasive migration of cancer cells [40]. Since some of these functions were present in our enriched terms, we wished to test if the expression of a gene set annotated to the cancer-related enriched terms in the PIN would be related to clinical outcome of patients.

Figure 4. Validation of let-7c result. Gene expression-based outcome for breast cancer online (GOBO) survival analysis of let-7c-regulated PIN members marked with the following functions: (**A**) Microtubule cytoskeleton; (**B**) Negative regulation of programmed cell death; and (**C**) negative regulation of cell death. **Red**: samples with high expression of selected gene set (PIN members); **grey**: samples with low expression of selected gene set (PIN members).



As shown in Figure 4, Figure S14 and Table S7, only some of the enriched terms were significantly associated with clinical outcome. This may be because the changes in these key genes occurred at the protein level, such as in protein expression or even post-translational modification; therefore, mRNA expression-centric tools like GOBO cannot explore association of such genes to clinical outcomes.

Alternatively, it is possible the miRNA did not regulate the whole pathway, or the miRNA did not target the key part of the pathway directly, and thus the clinical outcome of gene sets of the enriched GO terms in this condition cannot be determined. However, some functions associated with clinical outcomes were observed. For example, proteins annotated with the terms "microtubule cytoskeleton", "negative regulation of programmed cell death", and "negative regulation of cell death" in the let-7c-regulated PIN were related to 10 year survival rate of patients, as reported by GOBO (Figure 4). Also, the enriched term "regulation of epithelial cell proliferation" for both miR-125a-5p and miR-125b-5p were found to be associated with the 10-year survival rate of patients. Therefore, these results further supported the GO enrichment analysis discussed previously.

3. Experimental Section

3.1. miRNA Microarray Experiments

We performed a miRNA microarray to obtain the expression profiles for receiver operating characteristic (ROC) curve analysis. This dataset was deposited in Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/, Series Accession: GSE45666). In total, there were 15 normal samples and 101 tumor samples in the expression profile. Detailed pathophysiogical characteistics of these samples were in Table S8 and GSE45666. All human tissue samples collected from breast cancer patients were approved and human subject confidentiality was protected by the Institute Review Board of National Taiwan University Hospital (IRB, 20071211R).

Total RNA was extracted from tissues collected from the patients using Trizol® Reagent (Invitrogen, Carlsbad, CA, USA, USA) according to the manufacturer's protocol. Purified RNA was quantified at OD260nm by using a ND-1000 spectrophotometer (Nanodrop Technology, Wilmington, DE, USA) and qualitated by using a Bioanalyzer 2100 with the RNA 6000 Nano LabChip kit (Agilent Technologies, Santa Clara, CA, USA).

After RNA extraction, 100 ng of total RNA was dephosphorylated and labeled with pCp-Cy3 using the Agilent miRNA Complete Labeling and Hyb Kit in conjunction with the microRNA Spike-In kit (Agilent Technologies, Santa Clara, CA, USA). Briefly, 2X hybridization buffer (Agilent Technologies) was added to the labeled mixture to a final volume of 45 μL. The mixture was heated for 5 min at 100 °C and immediately cooled to 0 °C. Each 45 μL sample was hybridized onto an Agilent human miRNA Microarray Release 12.0, 8 × 15 K (Agilent Technologies) at 55 °C for 20 h. After hybridization, slides were washed for 5 min in Gene Expression Wash Buffer 1 at room temperature, then for 5 min in Gene Expression Wash Buffer 2 at 37 °C. Slides were scanned on an Agilent microarray scanner (Agilent Technologies, model G2505C) at 100% and 5% sensitivity settings. Feature Extraction (Agilent Technologies) software version 10.7.3.1 (Agilent Technologies, Santa Clara, CA, USA) was used for image analysis.

3.2. mRNA Expression Profiles

For miRNA-regulated protein-protein interaction network construction, the mRNA expression profile was fetched from GEO (Series Accession: GSE29174). This dataset was produced by Farazi *et al*. [26], which was the only dataset publicly available with large size of tumor samples and

reasonably sized normal samples. In total, 161 clinical samples were collected from breast cancer patients by biopsy: 110 invasive ductal carcinoma (IDC), 11 normal, 17 ductal carcinoma *in situ* (DCIS), 1 mucinous A, 8 atypical medullary, 4 apocrine, 8 metaplastic, and 2 adenoid, as classified by Farazi *et al.* The110 IDC samples were classified as the tumor group and the 11 normal samples were classified as the normal group in this study.

3.3. miRNA Expression Profiles

The miRNA expression profile used in this work was fetched from Table S4A of the work of Farazi *et al*. [26]. There were 189 samples in this dataset, with 6 of them from cell lines, and another 183 samples from patient tissues. In the 183 clinical samples collected from breast cancer patients by biopsy in this dataset, there were 128 IDC, 11 normal, 18 DCIS, 1 mucinous A, 8 atypical medullary, 4 apocrine, 9 metaplastic, and 2 adenoid cases, as classified by Farazi *et al.* The 128 IDC samples were classified as the tumor group and the 11 normal samples as were classified as the normal group.

3.4. Data Analysis

The overall workflow design was similar to the previous work of Tseng *et al*. [25] (see Figure 1). However, we applied the workflow on breast cancer expression profiles instead of gastric cancer as in the work of Tseng *et al.* Additionally, we used 3 miRNA target prediction databases, TargetScan (v6.0) [22,41,42], PicTar [43,44], and miRanda (release August 2010) [45] here, while only TargetScan was used in the previous study.

To construct the networks, we first elucidated differentially expressed miRNAs and mRNAs from published datasets. Both of miRNA (generated with miRNA-Seq technique) and mRNA array expression data were produced by Farazi *et al.* [26] from the same batch of clinical tumor samples. To obtain a list of differentially expressed genes and miRNAs between normal and tumor groups, we used significance analysis of microarrays (SAM) [46] implemented in R package samr (version 2.0; Stanford University, Stanford, CA, USA). We set the false discovery rate (FDR) as ≤0.0001%, fold change as ≥2.5 for miRNA, and fold change as ≥1.9 for genes as thresholds to reduce false positives. If a gene/miRNA was expressed higher in the normal group compared to the tumor group, we defined that gene/miRNA as a down-regulated gene/miRNA, and *vice versa*.

Following this, we paired the miRNAs and mRNAs with different expression trends. For example, an up-regulated miRNA would be paired with a down-regulated mRNA. If such pairs could be found in 2 (or more) of the 3 miRNA target prediction databases, they were then added to their corresponding miRNA-regulated network. To further extend the coverage of our network, we incorporated the human protein reference database (HPRD) [47], which contains experimentally verified interaction data, into our miRNA-regulated PINs.

Finally, we used gene ontology (GO) enrichment analysis to explore the function of the miRNA-regulated PINs. Hypergeometric tests were used to determine if a GO term was enriched in a PIN. In this section, we excluded PINs with less than 5 proteins. We also excluded GO terms with levels of less than 5 to avoid non-specific GO terms. Since we tested multiple GO terms on each miRNA-regulated PIN, we adjusted the significance of the test with the FDR method developed by Benjamini *et al.* [48]. We used the adjusted p-value < 0.0001 as our threshold. A GO term would

be excluded if its *p*-value was larger than 0.0001. Therefore, for each network, we ran the GO enrichment analysis, collected the calculated *p*-values, and adjusted these values using the methods described above.

3.5. ROC and GOBO Survival Analysis

After the PINs were constructed, we attempted to verify our results by literature search, ROC, and GOBO survival analysis [49]. To determine if the miRNAs we found could serve as classification markers for discriminating between normal and tumor samples, we applied ROC analysis on our miRNA array data (described in Section 2.1). ROC analysis is usually used to evaluate the efficiency of a classifier or a biological marker. R package ROCR [50] (version 1.0.4) was used to plot the ROC curve and calculate the area under curve (AUC). The standard error of AUC was then calculated as described in the work of Hanlye and McNeil [51]. The *p*-value of AUC was thus calculated with standard error obtained in the previous step. To further validate if the PINs we found were related to cancer, we used survival analysis implemented in GOBO [49] (available at http://co.bmc.lu.se/gobo), which provides a large amount of breast cancer gene expression profiles collected from public databases with clinical outcome data. In both ROC analysis and GOBO survival analysis, we considered our results significant when the *p*-value was smaller than 0.05.

4. Conclusions

Using integrative analysis of miRNA and mRNA expression profiles, we have identified not only breast cancer-related miRNAs and genes, but also putative roles for miRNAs in cancer as elucidated from miRNA-regulated PINs constructed in this work. Here, some previously known functions of miRNAs were again presented in our results, e.g., the relationship between the miRNAs, let-7c, miR-125a-5p, miR-125b-5p, and miR-21-5p, and breast cancer were demonstrated in this research. Furthermore, we have identified additional miRNAs and their related functions that have not been previously reported or discussed, providing valuable resources for further research in breast cancer.

Acknowledgments

This work was supported by the National Science Council of Taiwan (NSC 97-2311-B-002-010-MY3 and NSC 99-2621-B-002-005-MY3), National Taiwan University Cutting-Edge Steering Research Project (NTU-CESRP-102R7602C3) and the National Health Research Institutes, Taiwan (NHRI-EX98-9819PI).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Information

Table S1. Significantly differentially expressed miRNAs found in miRNA dataset in Farazi *et al*. [26]. There are 89 down-regulated miRNAs and 1 up-regulated miRNA in this list. *Q*-values reported by SAM were 0 for all miRNAs in this list.

miRBase Accession	miRNA Name	Fold Change	miRBase Accession	miRNA Name	Fold Change
MIMAT0004761	hsa-miR-483-5p	0.01	**MIMAT0000077**	hsa-miR-22-3p	0.21
MIMAT0004552	hsa-miR-139-3p	0.01	**MIMAT0000089**	hsa-miR-31-5p	0.21
MIMAT0000738	hsa-miR-383	0.02	**MIMAT0004612**	hsa-miR-186-3p	0.21
MIMAT0002856	hsa-miR-520d-3p	0.02	**MIMAT0004592**	hsa-miR-125b-1-3p	0.22
MIMAT0002811	hsa-miR-202-3p	0.03	**MIMAT0001639**	hsa-miR-409-3p	0.22
MIMAT0002177	hsa-miR-486-5p	0.04	**MIMAT0015032**	hsa-miR-3158-3p	0.22
MIMAT0022721	hsa-miR-1247-3p	0.05	**MIMAT0004496**	hsa-miR-23a-5p	0.22
MIMAT0002175	hsa-miR-485-5p	0.06	**MIMAT0000690**	hsa-miR-296-5p	0.22
MIMAT0000265	hsa-miR-204-5p	0.07	**MIMAT0000731**	hsa-miR-378a-5p	0.23
MIMAT0000752	hsa-miR-328	0.07	**MIMAT0000448**	hsa-miR-136-5p	0.23
MIMAT0000421	hsa-miR-122-5p	0.07	**MIMAT0004796**	hsa-miR-576-3p	0.23
MIMAT0000447	hsa-miR-134	0.08	**MIMAT0010133**	hsa-miR-2110	0.23
MIMAT0000722	hsa-miR-370	0.09	**MIMAT0004951**	hsa-miR-887	0.23
MIMAT0004513	hsa-miR-101-5p	0.09	**MIMAT0003239**	hsa-miR-574-3p	0.25
MIMAT0000446	hsa-miR-127-3p	0.10	**MIMAT0005901**	hsa-miR-1249	0.25
MIMAT0000097	hsa-miR-99a-5p	0.10	**MIMAT0000510**	hsa-miR-320a	0.26
MIMAT0004566	hsa-miR-218-2-3p	0.10	**MIMAT0002172**	hsa-miR-376b	0.26
MIMAT0000729	hsa-miR-376a-3p	0.11	**MIMAT0000250**	hsa-miR-139-5p	0.27
MIMAT0009197	hsa-miR-205-3p	0.11	**MIMAT0005825**	hsa-miR-1180	0.27
MIMAT0004615	hsa-miR-195-3p	0.11	**MIMAT0000437**	hsa-miR-145-5p	0.28
MIMAT0005899	hsa-miR-1247-5p	0.11	**MIMAT0004601**	hsa-miR-145-3p	0.28
MIMAT0000720	hsa-miR-376c	0.12	**MIMAT0003322**	hsa-miR-652-3p	0.28
MIMAT0000762	hsa-miR-324-3p	0.12	**MIMAT0000756**	hsa-miR-326	0.28
MIMAT0004679	hsa-miR-296-3p	0.12	**MIMAT0000098**	hsa-miR-100-5p	0.29
MIMAT0004614	hsa-miR-193a-5p	0.12	**MIMAT0003296**	hsa-miR-627	0.29
MIMAT0003880	hsa-miR-671-5p	0.12	**MIMAT0002820**	hsa-miR-497-5p	0.31
MIMAT0004795	hsa-miR-574-5p	0.12	**MIMAT0004507**	hsa-miR-92a-1-5p	0.31
MIMAT0004599	hsa-miR-143-5p	0.13	**MIMAT0000271**	hsa-miR-214-3p	0.32
MIMAT0000423	hsa-miR-125b-5p	0.13	**MIMAT0004702**	hsa-miR-339-3p	0.33
MIMAT0004957	hsa-miR-760	0.13	**MIMAT0004611**	hsa-miR-185-3p	0.33
MIMAT0004911	hsa-miR-874	0.14	**MIMAT0000064**	hsa-let-7c	0.34
MIMAT0004603	hsa-miR-125b-2-3p	0.15	**MIMAT0004673**	hsa-miR-29c-5p	0.35
MIMAT0004952	hsa-miR-665	0.15	**MIMAT0000733**	hsa-miR-379-5p	0.35
MIMAT0018205	hsa-miR-3928	0.15	**MIMAT0004594**	hsa-miR-132-5p	0.35
MIMAT0004767	hsa-miR-193b-5p	0.15	**MIMAT0000765**	hsa-miR-335-5p	0.35
MIMAT0002861	hsa-miR-518e-3p	0.15	**MIMAT0002819**	hsa-miR-193b-3p	0.36
MIMAT0004604	hsa-miR-127-5p	0.16	**MIMAT0000088**	hsa-miR-30a-3p	0.36
MIMAT0002807	hsa-miR-491-5p	0.16	**MIMAT0005951**	hsa-miR-1307-3p	0.36
MIMAT0004689	hsa-miR-377-5p	0.16	**MIMAT0004597**	hsa-miR-140-3p	0.37
MIMAT0004762	hsa-miR-486-3p	0.16	**MIMAT0004556**	hsa-miR-10b-3p	0.37
MIMAT0000732	hsa-miR-378a-3p	0.17	**MIMAT0000272**	hsa-miR-215	0.37
MIMAT0017981	hsa-miR-3605-5p	0.18	**MIMAT0004511**	hsa-miR-99a-3p	0.37
MIMAT0004605	hsa-miR-129-2-3p	0.19	**MIMAT0000443**	hsa-miR-125a-5p	0.38
MIMAT0006789	hsa-miR-1468	0.20	**MIMAT0004482**	hsa-let-7b-3p	0.38
MIMAT0000737	hsa-miR-382-5p	0.21	**MIMAT0000076**	hsa-miR-21-5p	6.58

Table S2. Down-regulated genes found in dataset GSE29174. There are 726 down-regulated genes in this list. *Q*-values reported by SAM were 0 for all genes in this list.

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
2949	GSTM5	0.06	**619373**	MBOAT4	0.17
10894	LYVE1	0.06	**130399**	ACVR1C	0.17
5950	RBP4	0.07	**1646**	AKR1C2	0.17
762	CA4	0.09	**80763**	C12orf39	0.17
54997	TESC	0.09	**2159**	F10	0.18
3489	IGFBP6	0.09	**84889**	SLC7A3	0.18
3952	LEP	0.09	**1308**	COL17A1	0.18
213	ALB	0.09	**83699**	SH3BGRL2	0.18
3131	HLF	0.10	**84417**	C2orf40	0.18
4023	LPL	0.10	**4081**	MAB21L1	0.18
10633	RASL10A	0.11	**3484**	IGFBP1	0.18
364	AQP7	0.11	**5239**	PGM5	0.19
1908	EDN3	0.11	**4969**	OGN	0.19
1811	SLC26A3	0.11	**2719**	GPC3	0.19
91851	CHRDL1	0.11	**116362**	RBP7	0.19
729359	PLIN4	0.13	**948**	CD36	0.19
1149	CIDEA	0.13	**5764**	PTN	0.19
5959	RDH5	0.13	**3043**	HBB	0.19
5348	FXYD1	0.14	**56920**	SEMA3G	0.20
5346	PLIN1	0.14	**94274**	PPP1R14A	0.20
10249	GLYAT	0.14	**57447**	NDRG2	0.20
158800	RHOXF1	0.14	**84795**	PYROXD2	0.20
221476	PI16	0.14	**84649**	DGAT2	0.20
3040	HBA2	0.14	**2690**	GHR	0.20
6939	TCF15	0.14	**22802**	CLCA4	0.20
79645	EFCAB1	0.14	**5179**	PENK	0.20
80343	SEL1L2	0.14	**6663**	SOX10	0.20
9413	FAM189A2	0.15	**6649**	SOD3	0.21
26289	AK5	0.15	**54922**	RASIP1	0.21
25891	PAMR1	0.15	**8406**	SRPX	0.21
3679	ITGA7	0.15	**1446**	CSN1S1	0.21
1264	CNN1	0.15	**7123**	CLEC3B	0.22
92304	SCGB3A1	0.15	**9647**	PPM1F	0.22
2167	FABP4	0.15	**1842**	ECM2	0.22
23285	KIAA1107	0.15	**3909**	LAMA3	0.22
7145	TNS1	0.16	**8639**	AOC3	0.23
4881	NPR1	0.16	**2934**	GSN	0.23
1028	CDKN1C	0.16	**9370**	ADIPOQ	0.23
1036	CDO1	0.16	**3202**	HOXA5	0.23
130271	PLEKHH2	0.16	**9452**	ITM2A	0.23
8736	MYOM1	0.16	**6290**	SAA3P	0.23
8908	GYG2	0.16	**4604**	MYBPC1	0.23

Table S2. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
79785	RERGL	0.16	**1128**	CHRM1	0.23
221091	LRRN4CL	0.17	**83878**	USHBP1	0.24
3991	LIPE	0.17	**63970**	TP53AIP1	0.24
27175	TUBG2	0.24	**79192**	IRX1	0.28
1346	COX7A1	0.24	**3400**	ID4	0.28
6376	CX3CL1	0.24	**57519**	STARD9	0.29
50486	G0S2	0.24	**57666**	FBRSL1	0.29
6285	S100B	0.24	**3590**	IL11RA	0.29
443	ASPA	0.24	**57664**	PLEKHA4	0.29
947	CD34	0.25	**197257**	LDHD	0.29
84632	AFAP1L2	0.25	**66036**	MTMR9	0.29
3866	KRT15	0.25	**2321**	FLT1	0.29
147463	ANKRD29	0.25	**126**	ADH1C	0.29
2878	GPX3	0.25	**1363**	CPE	0.29
7079	TIMP4	0.25	**56131**	PCDHB4	0.29
54345	SOX18	0.25	**22915**	MMRN1	0.29
51277	DNAJC27	0.25	**7069**	THRSP	0.29
84870	RSPO3	0.25	**57161**	PELI2	0.30
55323	LARP6	0.25	**770**	CA11	0.30
6387	CXCL12	0.25	**53342**	IL17D	0.30
137835	TMEM71	0.25	**79987**	SVEP1	0.30
5212	VIT	0.25	**857**	CAV1	0.30
26577	PCOLCE2	0.25	**222166**	C7orf41	0.30
845	CASQ2	0.25	**27190**	IL17B	0.30
6422	SFRP1	0.25	**116159**	CYYR1	0.30
10351	ABCA8	0.26	**4487**	MSX1	0.30
10840	ALDH1L1	0.26	**9068**	ANGPTL1	0.30
65983	GRAMD3	0.26	**10411**	RAPGEF3	0.30
84327	ZBED3	0.26	**3199**	HOXA2	0.30
57124	CD248	0.26	**2944**	GSTM1	0.30
3235	HOXD9	0.26	**2920**	CXCL2	0.30
2192	FBLN1	0.26	**201134**	CEP112	0.31
91653	BOC	0.26	**220001**	VWCE	0.31
4147	MATN2	0.26	**83888**	FGFBP2	0.31
126669	SHE	0.27	**6366**	CCL21	0.31
2788	GNG7	0.27	**6711**	SPTBN1	0.31
129804	FBLN7	0.27	**85378**	TUBGCP6	0.31
270	AMPD1	0.27	**26040**	SETBP1	0.31
79656	BEND5	0.27	**4692**	NDN	0.31
58503	PROL1	0.27	**25890**	ABI3BP	0.31
3316	HSPB2	0.27	**23531**	MMD	0.31
729440	CCDC61	0.27	**30846**	EHD2	0.31
54438	GFOD1	0.27	**6196**	RPS6KA2	0.31
5243	ABCB1	0.27	**2009**	EML1	0.31

Table S2. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
810	CALML3	0.27	**6289**	SAA2	0.31
6898	TAT	0.27	**345275**	HSD17B13	0.31
5648	MASP1	0.28	**2701**	GJA4	0.32
25999	CLIP3	0.28	**112609**	MRAP2	0.32
125875	CLDND2	0.28	**727**	C5	0.32
7102	TSPAN7	0.28	**477**	ATP1A2	0.32
1879	EBF1	0.28	**9627**	SNCAIP	0.32
23252	OTUD3	0.28	**4435**	CITED1	0.32
5493	PPL	0.28	**10974**	C10orf116	0.32
83987	CCDC8	0.28	**11005**	SPINK5	0.32
9073	CLDN8	0.28	**80325**	ABTB1	0.33
221981	THSD7A	0.28	**221395**	GPR116	0.33
64102	TNMD	0.28	**10014**	HDAC5	0.33
137872	ADHFE1	0.33	**1489**	CTF1	0.37
27151	CPAMD8	0.33	**35**	ACADS	0.37
387923	SERP2	0.33	**3749**	KCNC4	0.37
145581	LRFN5	0.33	**140738**	TMEM37	0.37
6263	RYR3	0.33	**2791**	GNG11	0.37
2354	FOSB	0.33	**23604**	DAPK2	0.37
51302	CYP39A1	0.33	**10217**	CTDSPL	0.37
4128	MAOA	0.34	**23550**	PSD4	0.37
117248	GALNTL2	0.34	**4306**	NR3C2	0.37
10268	RAMP3	0.34	**119587**	CPXM2	0.37
7730	ZNF177	0.34	**7942**	TFEB	0.37
10873	ME3	0.34	**3815**	KIT	0.37
7461	CLIP2	0.34	**1805**	DPT	0.37
7049	TGFBR3	0.34	**23242**	COBL	0.37
79901	CYBRD1	0.34	**4313**	MMP2	0.37
5152	PDE9A	0.34	**4139**	MARK1	0.37
50805	IRX4	0.34	**9104**	RGN	0.37
8644	AKR1C3	0.34	**2329**	FMO4	0.37
5915	RARB	0.34	**25802**	LMOD1	0.38
2770	GNAI1	0.34	**4239**	MFAP4	0.38
54996	2-Mar	0.35	**10392**	NOD1	0.38
79791	FBXO31	0.35	**6794**	STK11	0.38
54776	PPP1R12C	0.35	**85458**	DIXDC1	0.38
9079	LDB2	0.35	**4123**	MAN2C1	0.38
57104	PNPLA2	0.35	**54476**	RNF216	0.38
30008	EFEMP2	0.35	**9920**	KBTBD11	0.38
91461	PKDCC	0.35	**6329**	SCN4A	0.38
23368	PPP1R13B	0.35	**10253**	SPRY2	0.38
23461	ABCA5	0.35	**1910**	EDNRB	0.38
9572	NR1D1	0.35	**9249**	DHRS3	0.38
23338	PHF15	0.35	**22869**	ZNF510	0.38

Table S2. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
114800	CCDC85A	0.35	**3384**	ICAM2	0.38
2550	GABBR1	0.35	**8613**	PPAP2B	0.38
4638	MYLK	0.35	**1950**	EGF	0.38
2327	FMO2	0.35	**55273**	TMEM100	0.38
139411	PTCHD1	0.35	**6297**	SALL2	0.38
10391	CORO2B	0.35	**9365**	KL	0.38
25854	FAM149A	0.35	**8863**	PER3	0.38
55701	ARHGEF40	0.36	**8404**	SPARCL1	0.38
1759	DNM1	0.36	**2202**	EFEMP1	0.38
22849	CPEB3	0.36	**8369**	HIST1H4G	0.38
57716	PRX	0.36	**5187**	PER1	0.39
1628	DBP	0.36	**30815**	ST6GALNAC6	0.39
80031	SEMA6D	0.36	**256364**	EML3	0.39
259217	HSPA12A	0.36	**57381**	RHOJ	0.39
6909	TBX2	0.36	**761**	CA3	0.39
1511	CTSG	0.36	**83989**	FAM172A	0.39
79971	WLS	0.36	**1408**	CRY2	0.39
90865	IL33	0.36	**2281**	FKBP1B	0.39
11343	MGLL	0.36	**51222**	ZNF219	0.39
55800	SCN3B	0.36	**54540**	FAM193B	0.39
1949	EFNB3	0.36	**4053**	LTBP2	0.39
284217	LAMA1	0.36	**55184**	DZANK1	0.39
22927	HABP4	0.37	**5740**	PTGIS	0.39
23645	PPP1R15A	0.39	**84814**	PPAPDC3	0.42
342574	KRT27	0.39	**79365**	BHLHE41	0.42
83543	AIF1L	0.39	**316**	AOX1	0.42
624	BDKRB2	0.39	**23380**	SRGAP2	0.42
347	APOD	0.39	**84033**	OBSCN	0.42
84935	C13orf33	0.39	**90353**	CTU1	0.42
858	CAV2	0.39	**9013**	TAF1C	0.42
5138	PDE2A	0.40	**474344**	GIMAP6	0.42
114928	GPRASP2	0.40	**84883**	AIFM2	0.42
58190	CTDSP1	0.40	**58480**	RHOU	0.42
513	ATP5D	0.40	**65982**	ZSCAN18	0.42
57684	ZBTB26	0.40	**666**	BOK	0.42
7041	TGFB1I1	0.40	**79762**	C1orf115	0.42
5787	PTPRB	0.40	**525**	ATP6V1B1	0.42
7294	TXK	0.40	**4675**	NAP1L3	0.42
56301	SLC7A10	0.40	**3257**	HPS1	0.43
55937	APOM	0.40	**55781**	RIOK2	0.43
6368	CCL23	0.40	**63947**	DMRTC1	0.43
55020	TTC38	0.40	**1969**	EPHA2	0.43
134265	AFAP1L1	0.40	**25927**	CNRIP1	0.43
4485	MST1	0.40	**57685**	CACHD1	0.43

Table S2. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
51559	NT5DC3	0.40	**29997**	GLTSCR2	0.43
7169	TPM2	0.40	**26051**	PPP1R16B	0.43
51705	EMCN	0.40	**83604**	TMEM47	0.43
8938	BAIAP3	0.40	**2308**	FOXO1	0.43
10365	KLF2	0.40	**55225**	RAVER2	0.43
59	ACTA2	0.40	**54839**	LRRC49	0.43
80309	SPHKAP	0.40	**122953**	JDP2	0.43
3779	KCNMB1	0.41	**29775**	CARD10	0.43
10826	C5orf4	0.41	**166**	AES	0.43
219654	ZCCHC24	0.41	**25924**	MYRIP	0.43
92162	TMEM88	0.41	**2852**	GPER	0.43
7450	VWF	0.41	**51421**	AMOTL2	0.43
10266	RAMP2	0.41	**124936**	CYB5D2	0.43
25875	LETMD1	0.41	**1294**	COL7A1	0.43
1938	EEF2	0.41	**127435**	PODN	0.43
121551	BTBD11	0.41	**84952**	CGNL1	0.43
2119	ETV5	0.41	**83483**	PLVAP	0.43
9696	CROCC	0.41	**1958**	EGR1	0.43
1031	CDKN2C	0.41	**230**	ALDOC	0.43
9037	SEMA5A	0.41	**65987**	KCTD14	0.43
3397	ID1	0.41	**4804**	NGFR	0.44
84707	BEX2	0.41	**64852**	TUT1	0.44
57616	TSHZ3	0.41	**84253**	GARNL3	0.44
1471	CST3	0.41	**5866**	RAB3IL1	0.44
55214	LEPREL1	0.41	**10608**	MXD4	0.44
3914	LAMB3	0.41	**4211**	MEIS1	0.44
57478	USP31	0.41	**83547**	RILP	0.44
3783	KCNN4	0.41	**9172**	MYOM2	0.44
8839	WISP2	0.41	**57192**	MCOLN1	0.44
1583	CYP11A1	0.42	**255877**	BCL6B	0.44
10124	ARL4A	0.42	**56904**	SH3GLB2	0.44
738	C11orf2	0.42	**51285**	RASL12	0.44
29800	ZDHHC1	0.42	**3425**	IDUA	0.44
23135	KDM6B	0.44	**402117**	VWC2L	0.46
171024	SYNPO2	0.44	**81490**	PTDSS2	0.46
10350	ABCA9	0.44	**283748**	PLA2G4D	0.46
3691	ITGB4	0.44	**23523**	CABIN1	0.46
2348	FOLR1	0.44	**6146**	RPL22	0.46
11145	PLA2G16	0.44	**85360**	SYDE1	0.46
554	AVPR2	0.45	**60468**	BACH2	0.46
64072	CDH23	0.45	**57451**	ODZ2	0.46
80177	MYCT1	0.45	**4013**	VWA5A	0.46
5957	RCVRN	0.45	**339768**	ESPNL	0.46
408	ARRB1	0.45	**3860**	KRT13	0.46

Table S2. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
144699	FBXL14	0.45	**7094**	TLN1	0.46
83719	YPEL3	0.45	**4232**	MEST	0.46
22841	RAB11FIP2	0.45	**1410**	CRYAB	0.46
283927	NUDT7	0.45	**57452**	GALNTL1	0.47
293	SLC25A6	0.45	**63935**	PCIF1	0.47
90507	SCRN2	0.45	**25873**	RPL36	0.47
37	ACADVL	0.45	**9812**	KIAA0141	0.47
112744	IL17F	0.45	**51665**	ASB1	0.47
6709	SPTAN1	0.45	**64123**	ELTD1	0.47
8086	AAAS	0.45	**6122**	RPL3	0.47
7423	VEGFB	0.45	**222962**	SLC29A4	0.47
64221	ROBO3	0.45	**23102**	TBC1D2B	0.47
7273	TTN	0.45	**3476**	IGBP1	0.47
2657	GDF1	0.45	**93408**	MYL10	0.47
59271	C21orf63	0.45	**5310**	PKD1	0.47
132160	PPM1M	0.45	**4628**	MYH10	0.47
27244	SESN1	0.45	**221935**	SDK1	0.47
51310	SLC22A17	0.45	**23328**	SASH1	0.47
4828	NMB	0.45	**8522**	GAS7	0.47
54360	CYTL1	0.45	**10023**	FRAT1	0.47
203245	NAIF1	0.45	**7301**	TYRO3	0.47
23166	STAB1	0.45	**2767**	GNA11	0.47
2121	EVC	0.45	**9457**	FHL5	0.47
116496	FAM129A	0.45	**4094**	MAF	0.47
23239	PHLPP1	0.45	**65268**	WNK2	0.47
51673	TPPP3	0.45	**54585**	LZTFL1	0.47
64094	SMOC2	0.45	**375449**	MAST4	0.47
6383	SDC2	0.45	**138311**	FAM69B	0.47
2180	ACSL1	0.45	**160622**	GRASP	0.47
23770	FKBP8	0.45	**22837**	COBLL1	0.47
55901	THSD1	0.46	**51435**	SCARA3	0.47
25895	METTL21B	0.46	**217**	ALDH2	0.47
23731	C9orf5	0.46	**6236**	RRAD	0.47
126393	HSPB6	0.46	**8322**	FZD4	0.47
4056	LTC4S	0.46	**653275**	CFC1B	0.47
79825	CCDC48	0.46	**10908**	PNPLA6	0.47
10810	WASF3	0.46	**57526**	PCDH19	0.47
29911	HOOK2	0.46	**8424**	BBOX1	0.47
583	BBS2	0.46	**9905**	SGSM2	0.48
28984	C13orf15	0.46	**10435**	CDC42EP2	0.48
1465	CSRP1	0.46	**23087**	TRIM35	0.48
55258	THNSL2	0.46	**60314**	C12orf10	0.48
161198	CLEC14A	0.46	**1073**	CFL2	0.48
3699	ITIH3	0.48	**5256**	PHKA2	0.49

Table S2. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
92922	CCDC102A	0.48	**6237**	RRAS	0.49
65057	ACD	0.48	**5288**	PIK3C2G	0.49
9095	TBX19	0.48	**10252**	SPRY1	0.49
6441	SFTPD	0.48	**79026**	AHNAK	0.49
22846	VASH1	0.48	**9693**	RAPGEF2	0.49
51066	C3orf32	0.48	**51226**	COPZ2	0.49
23179	RGL1	0.48	**158326**	FREM1	0.49
4664	NAB1	0.48	**1956**	EGFR	0.49
50511	SYCP3	0.48	**5360**	PLTP	0.49
6430	SRSF5	0.48	**290**	ANPEP	0.49
11078	TRIOBP	0.48	**1756**	DMD	0.49
78991	PCYOX1L	0.48	**5118**	PCOLCE	0.49
6623	SNCG	0.48	**56654**	NPDC1	0.49
23384	SPECC1L	0.48	**9254**	CACNA2D2	0.49
53826	FXYD6	0.48	**55536**	CDCA7L	0.49
9397	NMT2	0.48	**124975**	GGT6	0.49
6041	RNASEL	0.48	**1906**	EDN1	0.49
113510	HELQ	0.48	**81029**	WNT5B	0.49
64788	LMF1	0.48	**2646**	GCKR	0.49
2217	FCGRT	0.48	**9811**	CTIF	0.50
79720	VPS37B	0.48	**145376**	PPP1R36	0.50
6764	ST5	0.48	**222865**	TMEM130	0.50
252969	NEIL2	0.48	**92999**	ZBTB47	0.50
8987	STBD1	0.48	**168002**	DACT2	0.50
41	ACCN2	0.48	**6829**	SUPT5H	0.50
7905	REEP5	0.48	**9992**	KCNE2	0.50
5919	RARRES2	0.48	**58509**	C19orf29	0.50
10544	PROCR	0.48	**79706**	PRKRIP1	0.50
6876	TAGLN	0.48	**1153**	CIRBP	0.50
8436	SDPR	0.49	**9639**	ARHGEF10	0.50
23500	DAAM2	0.49	**4054**	LTBP3	0.50
130132	RFTN2	0.49	**1120**	CHKB	0.50
80310	PDGFD	0.49	**286046**	XKR6	0.50
4215	MAP3K3	0.49	**9590**	AKAP12	0.50
282775	OR5J2	0.49	**64115**	C10orf54	0.50
51161	C3orf18	0.49	**2067**	ERCC1	0.50
29098	RANGRF	0.49	**7507**	XPA	0.50
53336	CPXCR1	0.49	**22897**	CEP164	0.50
9081	PRY	0.49	**652**	BMP4	0.50
9459	ARHGEF6	0.49	**55702**	CCDC94	0.50
2995	GYPC	0.49	**57613**	KIAA1467	0.50
23057	NMNAT2	0.49	**28514**	DLL1	0.50
4669	NAGLU	0.49	**169270**	ZNF596	0.50
6452	SH3BP2	0.49	**83982**	IFI27L2	0.50

Table S2. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
51458	RHCG	0.49	**2247**	FGF2	0.50
1112	FOXN3	0.49	**26248**	OR2K2	0.50
29954	POMT2	0.49	**84303**	CHCHD6	0.50
9612	NCOR2	0.49	**3615**	IMPDH2	0.50
3198	HOXA1	0.49	**1813**	DRD2	0.50
5311	PKD2	0.49	**80148**	PQLC1	0.50
2946	GSTM2	0.49	**390081**	OR52E4	0.50
2109	ETFB	0.49	**352954**	GATS	0.50
56062	KLHL4	0.49	**90871**	C9orf123	0.50
6915	TBXA2R	0.50	**50945**	TBX22	0.52
64288	ZNF323	0.50	**5204**	PFDN5	0.52
5195	PEX14	0.50	**5338**	PLD2	0.52
84557	MAP1LC3A	0.50	**94**	ACVRL1	0.52
6164	RPL34	0.50	**54039**	PCBP3	0.52
8835	SOCS2	0.50	**7691**	ZNF132	0.52
2735	GLI1	0.50	**338**	APOB	0.52
26022	TMEM98	0.50	**84658**	EMR3	0.52
3908	LAMA2	0.50	**283232**	TMEM80	0.52
1825	DSC3	0.50	**5430**	POLR2A	0.52
5730	PTGDS	0.50	**54623**	PAF1	0.52
162515	SLC16A11	0.51	**11070**	TMEM115	0.52
274	BIN1	0.51	**10395**	DLC1	0.52
79654	HECTD3	0.51	**57140**	RNPEPL1	0.52
22863	ATG14	0.51	**79781**	IQCA1	0.52
25949	SYF2	0.51	**1838**	DTNB	0.52
84872	ZC3H10	0.51	**51386**	EIF3L	0.52
23187	PHLDB1	0.51	**56919**	DHX33	0.52
5434	POLR2E	0.51	**57542**	KLHDC5	0.52
6181	RPLP2	0.51	**3628**	INPP1	0.52
6141	RPL18	0.51	**4520**	MTF1	0.52
84747	UNC119B	0.51	**8547**	FCN3	0.52
23399	CTDNEP1	0.51	**60401**	EDA2R	0.52
599	BCL2L2	0.51	**8082**	SSPN	0.52
197258	FUK	0.51	**80755**	AARSD1	0.52
5207	PFKFB1	0.51	**710**	SERPING1	0.52
8131	NPRL3	0.51	**56246**	MRAP	0.52
25839	COG4	0.51	**10555**	AGPAT2	0.52
10816	SPINT3	0.51	**949**	SCARB1	0.52
60485	SAV1	0.51	**23743**	BHMT2	0.52
5681	PSKH1	0.51	**3910**	LAMA4	0.52
80318	GKAP1	0.51	**60370**	AVPI1	0.52
57088	PLSCR4	0.51	**5021**	OXTR	0.52
93129	ORAI3	0.51	**55997**	CFC1	0.52
5829	PXN	0.51	**23144**	ZC3H3	0.52

Table S2. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
56776	FMN2	0.51	**150709**	ANKAR	0.52
85456	TNKS1BP1	0.51	**6591**	SNAI2	0.52
283	ANG	0.51	**10129**	FRY	0.52
7035	TFPI	0.51	**5166**	PDK4	0.52
51232	CRIM1	0.51	**146433**	IL34	0.52
112616	CMTM7	0.51	**118812**	MORN4	0.53
22981	NINL	0.51	**10516**	FBLN5	0.53
8727	CTNNAL1	0.51	**9463**	PICK1	0.53
9902	MRC2	0.51	**127495**	LRRC39	0.53
10900	RUNDC3A	0.51	**7753**	ZNF202	0.53
51299	NRN1	0.51	**79827**	CLMP	0.53
79632	FAM184A	0.52	**203260**	CCDC107	0.53
80820	EEPD1	0.52	**83657**	DYNLRB2	0.53

Table S1. Up-regulated genes found in dataset GSE29174. There are 437 up-regulated genes in this list. *Q*-values reported by SAM were 0 for all genes in this list.

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
1300	COL10A1	42.74	**54443**	ANLN	5.79
3007	HIST1H1D	29.72	**6710**	SPTB	5.71
8366	HIST1H4B	25.58	**7272**	TTK	5.64
6286	S100P	25.19	**10635**	RAD51AP1	5.49
1301	COL11A1	24.72	**4069**	LYZ	5.37
3627	CXCL10	17.83	**55183**	RIF1	5.34
4283	CXCL9	15.88	**891**	CCNB1	5.34
1387	CREBBP	12.83	**91543**	RSAD2	5.31
27299	ADAMDEC1	12.78	**81610**	FAM83D	5.24
54986	ULK4	12.46	**64581**	CLEC7A	5.10
55771	PRR11	12.02	**10051**	SMC4	5.02
54790	TET2	11.25	**4085**	MAD2L1	4.96
6241	RRM2	10.60	**55872**	PBK	4.83
3433	IFIT2	10.49	**991**	CDC20	4.82
6999	TDO2	9.73	**9221**	NOLC1	4.74
1656	DDX6	9.72	**2124**	EVI2B	4.66
55088	C10orf118	9.37	**375248**	ANKRD36	4.66
9648	GCC2	9.24	**1164**	CKS2	4.64
6696	SPP1	8.92	**1230**	CCR1	4.62
2803	GOLGA4	8.57	**890**	CCNA2	4.56
83540	NUF2	7.73	**127933**	UHMK1	4.49
10112	KIF20A	7.66	**10274**	STAG1	4.45
9833	MELK	7.59	**597**	BCL2A1	4.43
55165	CEP55	7.50	**55355**	HJURP	4.41
10142	AKAP9	7.44	**54210**	TREM1	4.36
9447	AIM2	7.42	**253558**	LCLAT1	4.26

Table S3. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
2706	GJB2	7.33	**1033**	CDKN3	4.24
6498	SKIL	7.13	**79801**	SHCBP1	4.23
219285	SAMD9L	7.06	**126731**	C1orf96	4.21
10261	IGSF6	7.01	**6772**	STAT1	4.20
2335	FN1	6.95	**55729**	ATF7IP	4.14
699	BUB1	6.75	**6713**	SQLE	4.14
1058	CENPA	6.75	**157570**	ESCO2	4.10
332	BIRC5	6.73	**79871**	RPAP2	4.09
51203	NUSAP1	6.59	**9493**	KIF23	4.09
259266	ASPM	6.54	**4751**	NEK2	4.05
1063	CENPF	6.49	**10631**	POSTN	4.03
165918	RNF168	6.44	**23515**	MORC3	4.02
9232	PTTG1	6.34	**7153**	TOP2A	4.02
5996	RGS1	6.07	**10403**	NDC80	4.00
29089	UBE2T	5.96	**10915**	TCERG1	3.99
22974	TPX2	5.94	**57650**	KIAA1524	3.99
4321	MMP12	5.91	**23049**	SMG1	3.93
983	CDK1	5.89	**80231**	CXorf21	3.87
85444	LRRCC1	5.87	**5111**	PCNA	3.86
29121	CLEC2D	3.83	**79682**	MLF1IP	3.11
4090	SMAD5	3.80	**29123**	ANKRD11	3.09
2123	EVI2A	3.80	**5429**	POLH	3.09
57695	USP37	3.79	**701**	BUB1B	3.07
133418	EMB	3.76	**200030**	NBPF11	3.06
4131	MAP1B	3.76	**55677**	IWS1	3.06
9787	DLGAP5	3.75	**160418**	TMTC3	3.04
9768	KIAA0101	3.74	**9147**	NEMF	3.04
54625	PARP14	3.73	**11320**	MGAT4A	3.04
2215	FCGR3B	3.71	**5238**	PGM3	3.03
9134	CCNE2	3.70	**2820**	GPD2	3.02
3117	HLA-DQA1	3.68	**388886**	FAM211B	3.01
10380	BPNT1	3.67	**7852**	CXCR4	3.00
79056	PRRG4	3.63	**57082**	CASC5	2.99
10673	TNFSF13B	3.63	**22926**	ATF6	2.98
8467	SMARCA5	3.61	**7594**	ZNF43	2.98
115908	CTHRC1	3.61	**968**	CD68	2.97
3428	IFI16	3.61	**7171**	TPM4	2.96
1520	CTSS	3.61	**11004**	KIF2C	2.96
10797	MTHFD2	3.57	**10808**	HSPH1	2.95
55681	SCYL2	3.57	**84909**	C9orf3	2.94
9749	PHACTR2	3.57	**1894**	ECT2	2.93
94240	EPSTI1	3.56	**1629**	DBT	2.92
64151	NCAPG	3.51	**116969**	ART5	2.90
25879	DCAF13	3.51	**3227**	HOXC11	2.88

Table S3. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
116064	LRRC58	3.47	**3149**	HMGB3	2.87
29899	GPSM2	3.47	**10437**	IFI30	2.87
135114	HINT3	3.45	**57489**	ODF2L	2.87
27333	GOLIM4	3.43	**2151**	F2RL2	2.86
55839	CENPN	3.43	**23215**	PRRC2C	2.85
23213	SULF1	3.41	**128710**	C20orf94	2.85
81671	VMP1	3.39	**23594**	ORC6	2.84
9889	ZBED4	3.36	**5205**	ATP8B1	2.83
3092	HIP1	3.34	**51430**	C1orf9	2.80
51512	GTSE1	3.34	**57405**	SPC25	2.80
92797	HELB	3.34	**112401**	BIRC8	2.80
51426	POLK	3.30	**3606**	IL18	2.80
5611	DNAJC3	3.30	**115362**	GBP5	2.80
6596	HLTF	3.28	**50515**	CHST11	2.79
9910	RABGAP1L	3.25	**83461**	CDCA3	2.79
528	ATP6V1C1	3.23	**10744**	PTTG2	2.78
3833	KIFC1	3.23	**51765**	MST4	2.77
197131	UBR1	3.20	**10926**	DBF4	2.76
29923	HILPDA	3.20	**27125**	AFF4	2.75
28998	MRPL13	3.19	**10615**	SPAG5	2.75
58527	C6orf115	3.19	**55143**	CDCA8	2.74
79000	C1orf135	3.19	**51602**	NOP58	2.74
9857	CEP350	3.18	**51478**	HSD17B7	2.73
84296	GINS4	3.18	**2209**	FCGR1A	2.73
81034	SLC25A32	3.15	**9958**	USP15	2.72
55723	ASF1B	3.14	**5469**	MED1	2.72
7110	TMF1	3.14	**8813**	DPM1	2.70
84081	NSRP1	3.14	**6731**	SRP72	2.70
23075	SWAP70	3.12	**9991**	PTBP3	2.70
6726	SRP9	2.69	**79866**	BORA	2.41
55215	FANCI	2.68	**7072**	TIA1	2.40
57590	WDFY1	2.67	**55632**	G2E3	2.40
55142	HAUS2	2.66	**2213**	FCGR2B	2.40
23047	PDS5B	2.66	**3987**	LIMS1	2.39
5373	PMM2	2.66	**829**	CAPZA1	2.39
11065	UBE2C	2.66	**26973**	CHORDC1	2.38
23085	ERC1	2.66	**435**	ASL	2.38
389197	C4orf50	2.65	**29979**	UBQLN1	2.38
11260	XPOT	2.65	**8548**	BLZF1	2.37
29980	DONSON	2.65	**9694**	TTC35	2.37
64399	HHIP	2.64	**55055**	ZWILCH	2.36
6453	ITSN1	2.63	**4481**	MSR1	2.36
29108	PYCARD	2.63	**10213**	PSMD14	2.35
9877	ZC3H11A	2.62	**9966**	TNFSF15	2.35

Table S3. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
81624	DIAPH3	2.62	**51582**	AZIN1	2.35
79723	SUV39H2	2.61	**54843**	SYTL2	2.34
55789	DEPDC1B	2.61	**9039**	UBA3	2.33
10097	ACTR2	2.59	**933**	CD22	2.33
23036	ZNF292	2.58	**5685**	PSMA4	2.33
22936	ELL2	2.57	**9885**	OSBPL2	2.33
8477	GPR65	2.57	**9262**	STK17B	2.33
23397	NCAPH	2.57	**56942**	C16orf61	2.32
3015	H2AFZ	2.54	**10767**	HBS1L	2.32
55749	CCAR1	2.53	**87178**	PNPT1	2.32
25937	WWTR1	2.52	**6303**	SAT1	2.32
360023	ZBTB41	2.51	**7316**	UBC	2.32
5080	PAX6	2.51	**4205**	MEF2A	2.32
4193	MDM2	2.51	**85465**	EPT1	2.31
24137	KIF4A	2.51	**84640**	USP38	2.31
9212	AURKB	2.51	**5810**	RAD1	2.30
168850	ZNF800	2.50	**64397**	ZFP106	2.29
55109	AGGF1	2.49	**5706**	PSMC6	2.29
23185	LARP4B	2.49	**22948**	CCT5	2.29
51571	FAM49B	2.49	**10672**	GNA13	2.29
51077	FCF1	2.49	**339344**	MYPOP	2.28
23167	EFR3A	2.49	**7292**	TNFSF4	2.28
23468	CBX5	2.48	**57103**	C12orf5	2.28
5396	PRRX1	2.48	**388403**	YPEL2	2.28
10096	ACTR3	2.47	**54876**	DCAF16	2.27
10308	ZNF267	2.47	**113235**	SLC46A1	2.27
6782	HSPA13	2.47	**11177**	BAZ1A	2.27
3832	KIF11	2.47	**339175**	METTL2A	2.26
917	CD3G	2.47	**26586**	CKAP2	2.26
80821	DDHD1	2.46	**55785**	FGD6	2.26
52	ACP1	2.46	**24145**	PANX1	2.25
4179	CD46	2.46	**253461**	ZBTB38	2.25
10499	NCOA2	2.44	**23232**	TBC1D12	2.25
60558	GUF1	2.44	**995**	CDC25C	2.25
55676	SLC30A6	2.43	**55974**	SLC50A1	2.25
6646	SOAT1	2.43	**472**	ATM	2.25
5440	POLR2K	2.43	**23008**	KLHDC10	2.24
84955	NUDCD1	2.42	**10024**	TROAP	2.24
54739	XAF1	2.42	**9521**	EEF1E1	2.24
84295	PHF6	2.23	**7402**	UTRN	2.09
7295	TXN	2.23	**55589**	BMP2K	2.08
2710	GK	2.23	**158747**	MOSPD2	2.08
10905	MAN1A2	2.22	**56886**	UGGT1	2.07
6780	STAU1	2.22	**203100**	HTRA4	2.07

Table S3. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
10282	BET1	2.22	**55279**	ZNF654	2.07
134430	WDR36	2.21	**54499**	TMCO1	2.07
4299	AFF1	2.21	**81930**	KIF18A	2.07
6747	SSR3	2.21	**142686**	ASB14	2.06
7334	UBE2N	2.21	**55209**	SETD5	2.06
5965	RECQL	2.21	**9736**	USP34	2.04
4605	MYBL2	2.2	**116285**	ACSM1	2.04
6093	ROCK1	2.19	**2201**	FBN2	2.04
161725	OTUD7A	2.19	**963**	CD53	2.04
23518	R3HDM1	2.18	**55159**	RFWD3	2.03
2239	GPC4	2.18	**9871**	SEC24D	2.03
28977	MRPL42	2.18	**9887**	SMG7	2.02
64859	OBFC2A	2.18	**23376**	UFL1	2.02
3845	KRAS	2.18	**79646**	PANK3	2.01
51388	NIP7	2.18	**50613**	UBQLN3	2.00
7586	ZKSCAN1	2.18	**201595**	STT3B	2.00
10762	NUP50	2.17	**59345**	GNB4	1.99
7328	UBE2H	2.17	**5876**	RABGGTB	1.99
10730	YME1L1	2.17	**79820**	CATSPERB	1.99
23093	TTLL5	2.17	**6637**	SNRPG	1.99
6790	AURKA	2.17	**51330**	TNFRSF12A	1.99
22889	KIAA0907	2.17	**9928**	KIF14	1.99
10875	FGL2	2.17	**286097**	EFHA2	1.98
23161	SNX13	2.17	**9131**	AIFM1	1.98
9169	SCAF11	2.16	**488**	ATP2A2	1.98
1788	DNMT3A	2.15	**23042**	PDXDC1	1.98
9088	PKMYT1	2.15	**7114**	TMSB4X	1.98
23033	DOPEY1	2.13	**9123**	SLC16A3	1.98
89882	TPD52L3	2.13	**54454**	ATAD2B	1.97
6556	SLC11A1	2.13	**23143**	LRCH1	1.97
64216	TFB2M	2.13	**4212**	MEIS2	1.97
3071	NCKAP1L	2.13	**1457**	CSNK2A1	1.97
51068	NMD3	2.13	**80012**	PHC3	1.97
509	ATP5C1	2.13	**128497**	SPATA25	1.96
953	ENTPD1	2.13	**186**	AGTR2	1.96
51105	PHF20L1	2.13	**53981**	CPSF2	1.96
5062	PAK2	2.13	**56996**	SLC12A9	1.96
9205	ZMYM5	2.12	**1584**	CYP11B1	1.96
55157	DARS2	2.12	**133619**	PRRC1	1.96
8520	HAT1	2.11	**4288**	MKI67	1.96
79739	TTLL7	2.11	**9014**	TAF1B	1.96
9495	AKAP5	2.10	**55858**	TMEM165	1.96
3181	HNRNPA2B1	2.10	**2212**	FCGR2A	1.96

Table S3. *Cont.*

NCBI gene ID	Gene Symbol	Fold Change	NCBI gene ID	Gene Symbol	Fold Change
389898	UBE2NL	2.10	**10075**	HUWE1	1.96
29850	TRPM5	2.10	**220988**	HNRNPA3	1.96
3070	HELLS	2.10	**80146**	UXS1	1.95
331	XIAP	2.09	**122011**	CSNK1A1L	1.95
55751	TMEM184C	2.09	**150468**	CKAP2L	1.95
2146	EZH2	2.09	**84624**	FNDC1	1.95
26057	ANKRD17	1.95	**7332**	UBE2L3	1.92
128061	C1orf131	1.95	**3336**	HSPE1	1.92
64090	GAL3ST2	1.94	**54800**	KLHL24	1.92
130507	UBR3	1.93	**2290**	FOXG1	1.91
2298	FOXD4	1.93	**50848**	F11R	1.91
123169	LEO1	1.93	**10627**	MYL12A	1.91
57187	THOC2	1.93	**5074**	PAWR	1.91
148789	B3GALNT2	1.93	**6476**	SI	1.91
58508	MLL3	1.92	**1009**	CDH11	1.90
5701	PSMC2	1.92	**29066**	ZC3H7A	1.90
148066	ZNRF4	1.92	**51319**	RSRC1	1.90
6670	SP3	1.92			

Table S4. Result of ROC curve analysis on our miRNA array data. ROC analysis was done to validate the diagnostic value of the miRNA in the miRNA-regulated PINs. AUC: area under (ROC) curve; * p-value < 0.05; *** p-value < 0.001.

miRBase Accession	miRNA name	AUC	p-value
MIMAT0002856	hsa-miR-520d-3p	0.49	0.549112
MIMAT0000265	hsa-miR-204-5p	0.98	6.47×10^{-10} ***
MIMAT0000272	hsa-miR-215	0.21	0.999782
MIMAT0000271	hsa-miR-214-3p	0.68	0.010387 *
MIMAT0002820	hsa-miR-497-5p	0.99	2.75×10^{-10} ***
MIMAT0000076	hsa-miR-21-5p	0.78	0.000184 ***
MIMAT0000738	hsa-miR-383	0.60	0.106284
MIMAT0000423	hsa-miR-125b-5p	0.99	2.48×10^{-10} ***
MIMAT0000064	hsa-let-7c	0.93	3.79×10^{-8} ***
MIMAT0000089	hsa-miR-31-5p	0.80	8.63×10^{-5} ***
MIMAT0000077	hsa-miR-22-3p	0.27	0.99749
MIMAT0000098	hsa-miR-100-5p	0.98	5.55×10^{-10} ***
MIMAT0000097	hsa-miR-99a-5p	0.99	2.55×10^{-10} ***
MIMAT0000443	hsa-miR-125a-5p	0.31	0.990694
MIMAT0002819	hsa-miR-193b-3p	0.41	0.86128
MIMAT0000250	hsa-miR-139-5p	0.99	2.42×10^{-10} ***
MIMAT0000437	hsa-miR-145-5p	0.96	3.14×10^{-9} ***
MIMAT0000421	hsa-miR-122-5p	0.48	0.597483

Table S5. Summary of constructed miRNA-regulated networks. L0 gene: genes connected directly to the miRNA (*i.e.*, direct target of miRNA); L1 gene: genes not connected directly to the miRNA.

miRBase Accession	miR name	Total gene count	L0 count	L1 count
MIMAT0002819	hsa-miR-193b-3p	16	1	15
MIMAT0000250	hsa-miR-139-5p	28	10	18
MIMAT0000437	hsa-miR-145-5p	86	22	64
MIMAT0000423	hsa-miR-125b-5p	211	16	195
MIMAT0000443	hsa-miR-125a-5p	206	14	192
MIMAT0000097	hsa-miR-99a-5p	14	1	13
MIMAT0000265	hsa-miR-204-5p	64	18	46
MIMAT0000076	hsa-miR-21-5p	91	16	75
MIMAT0000064	hsa-let-7c	96	20	76
MIMAT0000421	hsa-miR-122-5p	5	3	2
MIMAT0000098	hsa-miR-100-5p	14	1	13
MIMAT0000272	hsa-miR-215	3	3	0
MIMAT0000271	hsa-miR-214-3p	14	8	6
MIMAT0000738	hsa-miR-383	34	3	31
MIMAT0002856	hsa-miR-520d-3p	146	23	123
MIMAT0000077	hsa-miR-22-3p	46	11	35
MIMAT0002820	hsa-miR-497-5p	267	32	235
MIMAT0000089	hsa-miR-31-5p	34	3	31

Table S6. Specific enriched GO terms of each miRNA-regulated PINs. Genes annotated with the specific GO term in the PIN were also listed in this table. Adj. *p*-value: multiple-test adjusted p-value calculated by the method described in the work of Benjamini and Yekutieli [48].

MIMAT0002856(hsa-miR-520d-3p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0007169, Transmembrane receptor protein tyrosine kinase signaling pathway	SH3KBP1, HDAC2, RET, ABI1, LYN, GRB2, SORBS1, CLTC, CLTA, CDC42, CASP9, RAF1, SRC, AP2A1, AP2B1, MAPK3, ARHGEF7, PRKCA, RPS6, PRKAR2B, MAPK1, ARHGEF6, CDK1, SH3GL2, EIF4G1, HDAC1, ECT2, MKNK1, CASP3, PRKACA, ADRB2, PRKAR2A, EIF4B, SHC1, RAC1	2.77×10^{-29}
GO:0048011, Nerve growth factor receptor signaling pathway	HDAC2, GRB2, CLTC, CLTA, CASP9, RAF1, SRC, AP2A1, AP2B1, MAPK3, ARHGEF7, PRKCA, PRKAR2B, MAPK1, ARHGEF6, CDK1, SH3GL2, HDAC1, ECT2, CASP3, PRKACA, PRKAR2A, SHC1, RAC1	5.09×10^{-24}
GO:0007173, Epidermal growth factor receptor signaling pathway	SH3KBP1, GRB2, CLTC, CLTA, CDC42, CASP9, RAF1, SRC, AP2A1, AP2B1, MAPK3, ARHGEF7, PRKCA, PRKAR2B, MAPK1, CDK1, SH3GL2, PRKACA, PRKAR2A, SHC1	2.96×10^{-22}
GO:0043067, Regulation of programmed cell death	HDAC2, STK17B, ESR1, ABL1, LYN, TP53, GABRB3, PAK2, LCK, CASP9, RAF1, PLK1, ARHGEF7, PRKCA, RPS6, SH3RF1, MAPK1, IFT57, ARHGAP10, ARHGEF6, CDK1, APAF1, HDAC1, ECT2, CASP3, SOX10, EP300, ARAF, TFAP2A, ADRB2, HCK, KLHL20, CASP8, HIP1, RAC1	4.76×10^{-19}
GO:0042058, Regulation of epidermal growth factor receptor signaling pathway	SH3KBP1, ESR1, GRB2, CLTC, CLTA, CDC42, AP2A1, AP2B1, ARHGEF7, SH3GL2, SHC1	2.36×10^{-12}
GO:0008543, Fibroblast growth factor receptor signaling pathway	GRB2, CASP9, RAF1, SRC, MAPK3, PRKCA, PRKAR2B, MAPK1, CDK1, MKNK1, PRKACA, PRKAR2A, SHC1	2.39×10^{-12}
GO:0043068, Positive regulation of programmed cell death	STK17B, ABL1, LYN, TP53, LCK, CASP9, ARHGEF7, PRKCA, RPS6, SH3RF1, MAPK1, ARHGEF6, APAF1, ECT2, CASP3, EP300, TFAP2A, ADRB2, CASP8, HIP1, RAC1	3.09×10^{-12}
GO:0010942, Positive regulation of cell death	STK17B, ABL1, LYN, TP53, LCK, CASP9, ARHGEF7, PRKCA, RPS6, SH3RF1, MAPK1, ARHGEF6, APAF1, ECT2, CASP3, EP300, TFAP2A, ADRB2, CASP8, HIP1, RAC1	4.49×10^{-12}
GO:0006917, Induction of apoptosis	STK17B, ABL1, TP53, LCK, CASP9, ARHGEF7, PRKCA, SH3RF1, MAPK1, ARHGEF6, APAF1, ECT2, CASP3, EP300, CASP8, HIP1, RAC1	6.91×10^{-11}
GO:0012502, Induction of programmed cell death	STK17B, ABL1, TP53, LCK, CASP9, ARHGEF7, PRKCA, SH3RF1, MAPK1, ARHGEF6, APAF1, ECT2, CASP3, EP300, CASP8, HIP1, RAC1	7.42×10^{-11}

Table S6. *Cont.*

MIMAT0002856(hsa-miR-520d-3p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0042059, Negative regulation of epidermal growth factor receptor signaling pathway	SH3KBP1, GRB2, CLTC, CLTA, CDC42, AP2A1, AP2B1, ARHGEF7, SH3GL2	8.63×10^{-11}
GO:0015630, Microtubule cytoskeleton	STMN1, SORBS1, SMAD4, CLTC, CDC42, LCK, RACGAP1, PLK1, PRKAR2B, YES1, MAPK1, IFT57, CDK1, ECT2, PRKACA, RB1, EP300, CCNB1, CHAF1B, TFAP2A, CASP8, PRKAR2A	4.75×10^{-10}
GO:0060548, Negative regulation of cell death	HDAC2, ESR1, TP53, SMAD4, RAF1, PLK1, PRKCA, RPS6, SH3RF1, CDK1, HDAC1, CASP3, SOX10, ARAF, TFAP2A, HCK, KLHL20	6.27×10^{-8}
GO:0008286, Insulin receptor signaling pathway	GRB2, SORBS1, RAF1, MAPK3, RPS6, MAPK1, CDK1, EIF4G1, EIF4B, SHC1	2.15×10^{-7}
GO:0043069, Negative regulation of programmed cell death	HDAC2, ESR1, TP53, RAF1, PLK1, PRKCA, RPS6, SH3RF1, CDK1, HDAC1, CASP3, SOX10, ARAF, TFAP2A, HCK, KLHL20	3.13×10^{-7}
GO:0008284, Positive regulation of cell proliferation	HDAC2, ESR1, LYN, CDC42, E2F1, PRKCA, MAPK1, CDK1, RHOG, HDAC1, NCK1, SOX10, CCNB1, ADRB2, HCK, SHC1	8.34×10^{-7}
GO:0051988, Regulation of attachment of spindle microtubules to kinetochore	CDC42, RACGAP1, ECT2, CCNB1	3.27×10^{-5}
GO:0008629, Induction of apoptosis by intracellular signals	ABL1, TP53, CASP9, APAF1, CASP3, EP300, CASP8	5.83×10^{-5}
MIMAT0002820(hsa-miR-497-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0043067, Regulation of programmed cell death	ESR1, MEN1, ABL1, HIPK3, PPARGC1A, SIAH1, SH3RF1, PAK2, LCK, MED1, PPARG, CBX4, ARHGEF7, YWHAB, RXRA, ACVR1, MAPK1, CASP3, CASP6, AR, PTPRF, MDM2, BRCA1, MLH1, RAB27A, PIAS4, FAF1, RAC1, VHL, SKI, NR4A1, LYN, TP53, PSMC2, GATA1, GATA6, GATA3, RAF1, CDKN1B, PLK1, PSMD11, HOXA13, RPS6, ESR2, ARHGAP10, ARHGEF6, SMAD3, SKIL, RYR2, PSEN1, HCK, TRAF2	2.67×10^{-25}
GO:0043068, Positive regulation of programmed cell death	MEN1, ABL1, SIAH1, SH3RF1, LCK, PPARG, ARHGEF7, YWHAB, RXRA, MAPK1, CASP3, CASP6, PTPRF, BRCA1, MLH1, RAB27A, PIAS4, FAF1, RAC1, NR4A1, LYN, TP53, GATA6, CDKN1B, HOXA13, RPS6, ESR2, ARHGEF6, SMAD3, RYR2, PSEN1, TRAF2	1.74×10^{-17}
GO:0010942, Positive regulation of cell death	MEN1, ABL1, SIAH1, SH3RF1, LCK, PPARG, ARHGEF7, YWHAB, RXRA, MAPK1, CASP3, CASP6, PTPRF, BRCA1, MLH1, RAB27A, PIAS4, FAF1, RAC1, NR4A1, LYN, TP53, GATA6, CDKN1B, HOXA13, RPS6, ESR2, ARHGEF6, SMAD3, RYR2, PSEN1, TRAF2	3.08×10^{-17}

Table S6. *Cont.*

MIMAT0002820(hsa-miR-497-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0008285, Negative regulation of cell proliferation	MEN1, MED1, PPARG, RXRA, CASP3, AR, PTPRF, VDR, VHL, SKI, LYN, TP53, TOB1, GATA1, GATA3, RAF1, HNF4A, CDKN1B, BRD7, MED25, ESR2, ABI1, SMAD1, SMAD2, SMAD3, SMAD4, SOX7	3.85×10^{-14}
GO:0015629, Actin cytoskeleton	ABL1, SORBS1, FLNA, SEPT7, ANLN, MACF1, HAP1, SH3PXD2A, IQGAP2, BRCA1, ACTC1, ACTA1, MYL2, MYLK, SORBS2, ARPC4, ARPC5, ACTR2, ACTR3, ARPC1B, WASF1, WASF2, HCK	2.52×10^{-13}
GO:0006917, Induction of apoptosis	ABL1, SH3RF1, LCK, PPARG, ARHGEF7, YWHAB, MAPK1, CASP3, CASP6, BRCA1, MLH1, RAB27A, RAC1, NR4A1, TP53, CDKN1B, ARHGEF6, SMAD3, RYR2, PSEN1, TRAF2	9.69×10^{-11}
GO:0012502, Induction of programmed cell death	ABL1, SH3RF1, LCK, PPARG, ARHGEF7, YWHAB, MAPK1, CASP3, CASP6, BRCA1, MLH1, RAB27A, RAC1, NR4A1, TP53, CDKN1B, ARHGEF6, SMAD3, RYR2, PSEN1, TRAF2	1.06×10^{-10}
GO:0007178, Transmembrane receptor protein serine/threonine kinase signaling pathway	ACVR1, SMURF2, SKI, GDF6, BMP6, ZNF8, GATA4, HNF4A, SMAD1, SMAD2, SMAD3, SMAD4, SMAD5, RYR2	1.22×10^{-10}
GO:0007169, Transmembrane receptor protein tyrosine kinase signaling pathway	SORBS1, CDC42, SRC, MAPK3, ARHGEF7, YWHAB, MAPK1, SH3GL2, CASP3, MDM2, EIF4G1, RAC1, SH3KBP1, NR4A1, LYN, GRB2, RAF1, CDKN1B, RPS6, ABI1, ARHGEF6, MKNK1, PSEN1, EIF4B	1.67×10^{-10}
GO:0090092, Regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	MEN1, ACVR1, SMURF2, SKI, GDF6, TP53, BMP6, GATA4, GATA6, HOXA13, SMAD2, SMAD3, SMAD4, SKIL	7.51×10^{-10}
GO:0030509, BMP signaling pathway	ACVR1, SMURF2, SKI, GDF6, BMP6, ZNF8, SMAD1, SMAD4, SMAD5, RYR2	3.25×10^{-9}
GO:0060548, Negative regulation of cell death	ESR1, HIPK3, PPARGC1A, SH3RF1, MED1, CBX4, ACVR1, CASP3, AR, MDM2, VHL, TP53, GATA1, GATA6, GATA3, RAF1, CDKN1B, PLK1, RPS6, SMAD3, SMAD4, PSEN1, HCK	7.88×10^{-9}
GO:0007173, Epidermal growth factor receptor signaling pathway	CDC42, SRC, MAPK3, ARHGEF7, YWHAB, MAPK1, SH3GL2, MDM2, SH3KBP1, NR4A1, GRB2, RAF1, CDKN1B	9.22×10^{-9}
GO:0030521, Androgen receptor signaling pathway	PPARGC1A, MED14, MED1, AR, BRCA1, MED12, PIAS1, RAN, NR1I3	1.42×10^{-8}
GO:0043069, Negative regulation of programmed cell death	ESR1, HIPK3, PPARGC1A, SH3RF1, MED1, CBX4, ACVR1, CASP3, AR, MDM2, VHL, TP53, GATA1, GATA6, GATA3, RAF1, CDKN1B, PLK1, RPS6, SMAD3, PSEN1, HCK	2.44×10^{-8}
GO:0048011, Nerve growth factor receptor signaling pathway	SRC, MAPK3, ARHGEF7, YWHAB, MAPK1, SH3GL2, CASP3, MDM2, RAC1, NR4A1, GRB2, RAF1, CDKN1B, ARHGEF6, PSEN1	2.64×10^{-8}
GO:0032956, Regulation of actin cytoskeleton organization	ABL1, LRP1, ARPC4, ARPC5, ACTR3, ARPC1B, SMAD3, NCK1, SORBS3, HCK, LIMK1	6.42×10^{-6}

Table S6. *Cont.*

MIMAT0002820(hsa-miR-497-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0008543, Fibroblast growth factor receptor signaling pathway	SRC, MAPK3, YWHAB, MAPK1, MDM2, NR4A1, GRB2, RAF1, CDKN1B, MKNK1	9.06×10^{-6}
GO:0042059, Negative regulation of epidermal growth factor receptor signaling pathway	CDC42, ARHGEF7, SH3GL2, PTPRF, SH3KBP1, GRB2, PSEN1	1.11×10^{-5}
GO:0042058, Regulation of epidermal growth factor receptor signaling pathway	ESR1, CDC42, ARHGEF7, SH3GL2, PTPRF, SH3KBP1, GRB2, PSEN1	1.17×10^{-5}
GO:0007179, Transforming growth factor beta receptor signaling pathway	ACVR1, SMURF2, SKI, SMAD1, SMAD2, SMAD3, SMAD4, SMAD5	1.67×10^{-5}
GO:0015630, Microtubule cytoskeleton	STMN1, RIF1, SORBS1, CDC42, LCK, RACGAP1, YES1, YWHAB, MAPK1, SEPT7, KIF23, CDC16, MACF1, BRCA1, FEZ1, NCOR1, PLK1, CHD3, SMAD4, CEP350, CDC27, PSEN1	2.14×10^{-5}
GO:0017015, Regulation of transforming growth factor beta receptor signaling pathway	MEN1, SMURF2, SKI, TP53, SMAD2, SMAD3, SMAD4, SKIL	2.30×10^{-5}
GO:0070302, Regulation of stress-activated protein kinase signaling cascade	MEN1, ZEB2, HIPK3, SH3RF1, CDC42, MAPK3, MAPK1, LYN, NCOR1, TRAF2	2.74×10^{-5}
GO:0001959, Regulation of cytokine-mediated signaling pathway	HSP90AB1, MED1, PPARG, PTPRF, NR1H2, PIAS1, IL36RN, HIPK1	6.35×10^{-5}
GO:0008284, Positive regulation of cell proliferation	ESR1, CDC42, MED1, RARA, MAPK1, AR, MDM2, NR4A1, LYN, FZR1, BMP6, GATA1, GATA4, GATA6, CDKN1B, NCK1, HCLS1, HCK	7.29×10^{-5}

Table S6. *Cont.*

MIMAT0000423(hsa-miR-125b-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0043067, Regulation of programmed cell death	HMGA2, PML, PRNP, FGF2, XRCC4, BRCA1, IGFBP3, HDAC3, CTNNB1, CD5, CDK1, NKX2-5, MEF2C, PRKCI, CASP2, PSMA4, PSMA3, CFDP1, CAV1, FAF1, YWHAB, HIF1A, RELA, TCF7L2, TNFSF12, PSEN2, TP53, TOP2A, TNFRSF4, BID, MYC, JUN, OGT, CDKN1A, HOXA13, RNF7, PPP2R4, HDAC2, HDAC1, SNCA, PTEN, NFKBIA, IFI16, NOL3, TRAF2, HSP90B1	4.56×10^{-24}
GO:0060548, Negative regulation of cell death	HMGA2, PRNP, FGF2, XRCC4, HDAC3, CTNNB1, CDK1, NKX2-5, MEF2C, PRKCI, CFDP1, HIF1A, RELA, TCF7L2, PSEN2, TP53, TNFRSF4, MYC, JUN, CDKN1A, RNF7, HDAC2, HDAC1, SNCA, PTEN, MGMT, NFKBIA, NOL3, HSP90B1	8.04×10^{-17}
GO:0008284, Positive regulation of cell proliferation	HMGA2, FGF2, XRCC4, CDC25B, CTNNB1, EGR1, AGGF1, CDK1, NKX2-5, MEF2C, PRKCI, IRS1, HIF1A, RELA, HCLS1, TNFSF12, ARNT, PTPRC, TNFSF4, TNFRSF4, MYC, JUN, FGF1, CDKN1A, HDAC2, HDAC1, NOLC1, PTEN	2.20×10^{-15}
GO:0043069, Negative regulation of programmed cell death	HMGA2, PRNP, XRCC4, HDAC3, CTNNB1, CDK1, NKX2-5, MEF2C, PRKCI, CFDP1, HIF1A, RELA, TCF7L2, PSEN2, TP53, TNFRSF4, MYC, JUN, CDKN1A, RNF7, HDAC2, HDAC1, SNCA, PTEN, NFKBIA, NOL3, HSP90B1	3.62×10^{-15}
GO:0043068, Positive regulation of programmed cell death	HMGA2, PML, BRCA1, IGFBP3, CTNNB1, CD5, MEF2C, PRKCI, CASP2, CAV1, FAF1, YWHAB, TNFSF12, PSEN2, TP53, TOP2A, BID, JUN, OGT, CDKN1A, HOXA13, RNF7, PPP2R4, PTEN, IFI16, TRAF2	4.51×10^{-14}
GO:0010942, Positive regulation of cell death	HMGA2, PML, BRCA1, IGFBP3, CTNNB1, CD5, MEF2C, PRKCI, CASP2, CAV1, FAF1, YWHAB, TNFSF12, PSEN2, TP53, TOP2A, BID, JUN, OGT, CDKN1A, HOXA13, RNF7, PPP2R4, PTEN, IFI16, TRAF2	7.15×10^{-14}
GO:0006916, Anti-apoptosis	PRNP, HDAC3, CDK1, NKX2-5, MEF2C, PRKCI, CFDP1, RELA, TCF7L2, PSEN2, RNF7, HDAC1, SNCA, NFKBIA, NOL3, HSP90B1	1.25×10^{-10}
GO:0008285, Negative regulation of cell proliferation	SERPINF1, SRF, PML, PRNP, FGF2, CSNK2B, IGFBP3, CTNNB1, CAV1, HMGA1, VDR, CDH5, HSF1, COL18A1, TP53, MYC, JUN, CDKN1A, PAK1, PTEN	2.08×10^{-9}
GO:0015630, Microtubule cytoskeleton	STMN1, KIF1C, RANGAP1, CDC25B, BRCA1, HDAC3, CTNNB1, PIN4, HSPH1, RANBP9, CDK1, SPTAN1, YWHAQ, DVL1, FKBP4, YWHAB, CCDC85B, MAPT, PSEN2, TOP2A, SPIB, MYC, OGT, APEX1, PAFAH1B1	2.24×10^{-9}
GO:0048011, Nerve growth factor receptor signaling pathway	HDAC3, CDK1, MEF2C, PRKCI, CASP2, IRS1, YWHAB, RELA, PSEN2, HDAC2, HDAC1, PTEN, ATF1, NFKBIA	1.99×10^{-8}
GO:0050678, Regulation of epithelial cell proliferation	SERPINF1, PGR, FGF2, CTNNB1, AGGF1, CAV1, HIF1A, TNFSF12, ARNT, MYC, JUN, FGF1, PTEN	3.41×10^{-8}
GO:0007169, Transmembrane receptor protein tyrosine kinase signaling pathway	FGF2, HDAC3, CDK1, MEF2C, PRKCI, CASP2, IRS1, FIBP, PTPN1, YWHAB, RELA, PSEN2, FGF1, HDAC2, HDAC1, PTEN, ATF1, NFKBIA, EIF4EBP1	6.39×10^{-8}

Table S6. *Cont.*

MIMAT0000423(hsa-miR-125b-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0006917, Induction of apoptosis	PML, BRCA1, CD5, CASP2, CAV1, YWHAB, TNFSF12, PSEN2, TP53, BID, OGT, CDKN1A, RNF7, PTEN, IFI16, TRAF2	1.68×10^{-7}
GO:0012502, Induction of programmed cell death	PML, BRCA1, CD5, CASP2, CAV1, YWHAB, TNFSF12, PSEN2, TP53, BID, OGT, CDKN1A, RNF7, PTEN, IFI16, TRAF2	1.81×10^{-7}
GO:0035666, TRIF-dependent toll-like receptor signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	2.40×10^{-6}
GO:0034138, Toll-like receptor 3 signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	2.71×10^{-6}
GO:0051693, Actin filament capping	SPTB, SPTBN1, SPTAN1, SPTA1, ADD1, EPB49	3.43×10^{-6}
GO:0002756, MyD88-independent toll-like receptor signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	3.82×10^{-6}
GO:0034134, Toll-like receptor 2 signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	5.33×10^{-6}
GO:0034130, Toll-like receptor 1 signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	5.33×10^{-6}
GO:0030835, Negative regulation of actin filament depolymerization	SPTB, SPTBN1, SPTAN1, SPTA1, ADD1, EPB49	5.58×10^{-6}
GO:0015629, Actin cytoskeleton	WAS, CDH1, BRCA1, SPTB, SPTBN1, SPTAN1, CTDP1, STX1A, SPTA1, PAK1, SNCA, ADD1, EPB41, EPB49	5.58×10^{-6}
GO:0002755, MyD88-dependent toll-like receptor signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	7.66×10^{-6}
GO:0034142, Toll-like receptor 4 signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	1.14×10^{-5}
GO:0050679, Positive regulation of epithelial cell proliferation	FGF2, CTNNB1, AGGF1, HIF1A, TNFSF12, ARNT, MYC, JUN, FGF1	1.14×10^{-5}

Table S6. *Cont.*

MIMAT0000076(hsa-miR-21-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0030834, Regulation of actin filament depolymerization	SPTB, SPTBN1, SPTAN1, SPTA1, ADD1, EPB49	1.27×10^{-5}
GO:0030837, Negative regulation of actin filament polymerization	SPTB, SPTBN1, SPTAN1, SPTA1, ADD1, EPB49	2.71×10^{-5}
GO:0002224, Toll-like receptor signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	2.96×10^{-5}
GO:0008629, Induction of apoptosis by intracellular signals	PML, BRCA1, YWHAB, TP53, BID, CDKN1A, RNF7, IFI16	3.52×10^{-5}
GO:0043067, Regulation of programmed cell death	SPRY2, TP53, ADAMTSL4, ETS1, TDGF1, RAF1, HOXA5, HOXA13, MSX1, MSX2, NKX2-5, CBL, INHBB, COL4A3, ACVR1C, TRAF2	1.92×10^{-5}
GO:0007173, Epidermal growth factor receptor signaling pathway	SPRY2, SPRY1, GRB2, PTPN11, TDGF1, RAF1, CBL	9.44×10^{-5}
MIMAT0000250(hsa-miR-139-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0035583, Negative regulation of transforming growth factor beta receptor signaling pathway by extracellular sequestering of TGFbeta	LTBP1, FBN1, FBN2	2.61×10^{-5}
MIMAT0000089(hsa-miR-31-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0007187, G-protein signaling, coupled to cyclic nucleotide second messenger	GNA12, GNA13, DRD5, MTNR1A, S1PR3, TSHR, S1PR4	8.74×10^{-7}
GO:0019935, Cyclic-nucleotide-mediated signaling	GNA12, GNA13, DRD5, MTNR1A, S1PR3, TSHR, S1PR4	1.60×10^{-6}
GO:0007188, G-protein signaling, coupled to cAMP nucleotide second messenger	GNA12, GNA13, DRD5, S1PR3, TSHR, S1PR4	6.82×10^{-6}

Table S6. *Cont.*

MIMAT0000089(hsa-miR-31-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0048011, Nerve growth factor receptor signaling pathway	ARHGEF1, PRKCD, ARHGEF12, PRKACA, PRKCE, MCF2, ARHGEF11	6.93×10^{-6}
GO:0019933, CAMP-mediated signaling	GNA12, GNA13, DRD5, S1PR3, TSHR, S1PR4	1.05×10^{-5}
GO:0043067, Regulation of programmed cell death	PTK2B, PRKCD, TGFBR1, ARHGEF12, CTNNB1, PRKCE, TIA1, MCF2, FASTK, ARHGEF11, F2R	1.25×10^{-5}
GO:0003376, Sphingosine-1-phosphate signaling pathway	S1PR3, S1PR2, S1PR4	3.22×10^{-5}
GO:0007169, Transmembrane receptor protein tyrosine kinase signaling pathway	PTK2B, ARHGEF1, PRKCD, ARHGEF12, PRKACA, PRKCE, MCF2, ARHGEF11	5.20×10^{-5}
GO:0043068, Positive regulation of programmed cell death	TGFBR1, ARHGEF12, CTNNB1, PRKCE, TIA1, MCF2, FASTK, ARHGEF11	7.67×10^{-5}
GO:0010942, Positive regulation of cell death	TGFBR1, ARHGEF12, CTNNB1, PRKCE, TIA1, MCF2, FASTK, ARHGEF11	8.78×10^{-5}
GO:0006917, Induction of apoptosis	TGFBR1, ARHGEF12, PRKCE, TIA1, MCF2, FASTK, ARHGEF11	9.32×10^{-5}
GO:0012502, Induction of programmed cell death	TGFBR1, ARHGEF12, PRKCE, TIA1, MCF2, FASTK, ARHGEF11	9.51×10^{-5}
MIMAT0000437(hsa-miR-145-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0030509, BMP signaling pathway	BMP6, ZNF8, ACVR1, SMAD1, SMAD4, RYR2, SMAD5, SMURF2, GDF6	1.37×10^{-11}
GO:0007178, Transmembrane receptor protein serine/threonine kinase signaling pathway	BMP6, ZNF8, ACVR1, SMAD1, SMAD4, RYR2, SMAD5, SMURF2, GDF6	2.56×10^{-8}
GO:0090092, Regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	MEN1, TP53, BMP6, HOXA13, ACVR1, SMAD4, SULF1, SMURF2, GDF6	8.22×10^{-8}

Table S6. *Cont.*

MIMAT0000064(hsa-let-7c)		
GO term	**Genes**	**Adj. *p*-value**
GO:0043067, Regulation of programmed cell death	RRM2B, ACVR1, TP53, RASA1, TGFBR1, PSMA3, BIRC5, ACTN2, HOXA13, IRS2, FASTK, VAV1, PSMB6, BCL2, CDK1, HDAC1, SOX10, TIA1, AKT1, AURKB	3.43×10^{-8}
GO:0043069, Negative regulation of programmed cell death	RRM2B, ACVR1, TP53, RASA1, TGFBR1, BIRC5, IRS2, BCL2, CDK1, HDAC1, SOX10, AKT1, AURKB	6.58×10^{-6}
GO:0060548, Negative regulation of cell death	RRM2B, ACVR1, TP53, RASA1, TGFBR1, BIRC5, IRS2, BCL2, CDK1, HDAC1, SOX10, AKT1, AURKB	8.92×10^{-6}
GO:0015630, Microtubule cytoskeleton	INCENP, SNTB2, SEPT1, TACC1, BIRC5, RACGAP1, PIN4, CDCA8, CDK1, PHF1, AKT1, AURKB, NINL, CCDC85B	5.15×10^{-5}
GO:0043067, Regulation of programmed cell death	HMGA2, PML, PRNP, FGF2, XRCC4, BRCA1, IGFBP3, HDAC3, CTNNB1, CD5, CDK1, NKX2-5, MEF2C, PRKCI, CASP2, PSMA4, PSMA3, CFDP1, CAV1, FAF1, YWHAB, HIF1A, RELA, TCF7L2, TNFSF12, PSEN2, TP53, TOP2A, TNFRSF4, BID, MYC, JUN, OGT, CDKN1A, RNF7, PPP2R4, HDAC2, HDAC1, SNCA, PTEN, NFKBIA, IFI16, NOL3, TRAF2, HSP90B1	1.62×10^{-23}
GO:0060548, Negative regulation of cell death	HMGA2, PRNP, FGF2, XRCC4, HDAC3, CTNNB1, CDK1, NKX2-5, MEF2C, PRKCI, CFDP1, HIF1A, RELA, TCF7L2, PSEN2, TP53, TNFRSF4, MYC, JUN, CDKN1A, RNF7, HDAC2, HDAC1, SNCA, PTEN, MGMT, NFKBIA, NOL3, HSP90B1	4.53×10^{-17}
GO:0008284, Positive regulation of cell proliferation	HMGA2, FGF2, XRCC4, CDC25B, CTNNB1, EGR1, AGGF1, CDK1, NKX2-5, MEF2C, PRKCI, IRS1, HIF1A, RELA, HCLS1, TNFSF12, ARNT, PTPRC, TNFSF4, TNFRSF4, MYC, JUN, FGF1, CDKN1A, HDAC2, HDAC1, NOLC1, PTEN	1.23×10^{-15}
MIMAT0000443(hsa-miR-125a-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0043069, Negative regulation of programmed cell death	HMGA2, PRNP, XRCC4, HDAC3, CTNNB1, CDK1, NKX2-5, MEF2C, PRKCI, CFDP1, HIF1A, RELA, TCF7L2, PSEN2, TP53, TNFRSF4, MYC, JUN, CDKN1A, RNF7, HDAC2, HDAC1, SNCA, PTEN, NFKBIA, NOL3, HSP90B1	2.02×10^{-15}
GO:0043068, Positive regulation of programmed cell death	HMGA2, PML, BRCA1, IGFBP3, CTNNB1, CD5, MEF2C, PRKCI, CASP2, CAV1, FAF1, YWHAB, TNFSF12, PSEN2, TP53, TOP2A, BID, JUN, OGT, CDKN1A, RNF7, PPP2R4, PTEN, IFI16, TRAF2	2.80×10^{-13}
GO:0010942, Positive regulation of cell death	HMGA2, PML, BRCA1, IGFBP3, CTNNB1, CD5, MEF2C, PRKCI, CASP2, CAV1, FAF1, YWHAB, TNFSF12, PSEN2, TP53, TOP2A, BID, JUN, OGT, CDKN1A, RNF7, PPP2R4, PTEN, IFI16, TRAF2	4.35×10^{-13}
GO:0006916, Anti-apoptosis	PRNP, HDAC3, CDK1, NKX2-5, MEF2C, PRKCI, CFDP1, RELA, TCF7L2, PSEN2, RNF7, HDAC1, SNCA, NFKBIA, NOL3, HSP90B1	9.04×10^{-11}
GO:0008285, Negative regulation of cell proliferation	SERPINF1, SRF, PML, PRNP, FGF2, CSNK2B, IGFBP3, CTNNB1, CAV1, HMGA1, VDR, CDH5, HSF1, COL18A1, TP53, MYC, JUN, CDKN1A, PAK1, PTEN	1.32×10^{-9}

Table S6. *Cont.*

MIMAT0000443(hsa-miR-125a-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0015630, Microtubule cytoskeleton	STMN1, KIF1C, RANGAP1, CDC25B, BRCA1, HDAC3, CTNNB1, PIN4, HSPH1, RANBP9, CDK1, SPTAN1, YWHAQ, DVL1, FKBP4, YWHAB, CCDC85B, MAPT, PSEN2, TOP2A, SPIB, MYC, OGT, APEX1, PAFAH1B1	1.32×10^{-9}
GO:0048011, Nerve growth factor receptor signaling pathway	HDAC3, CDK1, MEF2C, PRKCI, CASP2, IRS1, YWHAB, RELA, PSEN2, HDAC2, HDAC1, PTEN, ATF1, NFKBIA	1.50×10^{-8}
GO:0050678, Regulation of epithelial cell proliferation	SERPINF1, PGR, FGF2, CTNNB1, AGGF1, CAV1, HIF1A, TNFSF12, ARNT, MYC, JUN, FGF1, PTEN	2.66×10^{-8}
GO:0007169, Transmembrane receptor protein tyrosine kinase signaling pathway	FGF2, HDAC3, CDK1, MEF2C, PRKCI, CASP2, IRS1, FIBP, PTPN1, YWHAB, RELA, PSEN2, FGF1, HDAC2, HDAC1, PTEN, ATF1, NFKBIA, EIF4EBP1	4.37×10^{-8}
GO:0006917, Induction of apoptosis	PML, BRCA1, CD5, CASP2, CAV1, YWHAB, TNFSF12, PSEN2, TP53, BID, OGT, CDKN1A, RNF7, PTEN, IFI16, TRAF2	1.22×10^{-7}
GO:0012502, Induction of programmed cell death	PML, BRCA1, CD5, CASP2, CAV1, YWHAB, TNFSF12, PSEN2, TP53, BID, OGT, CDKN1A, RNF7, PTEN, IFI16, TRAF2	1.31×10^{-7}
GO:0035666, TRIF-dependent toll-like receptor signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	2.01×10^{-6}
GO:0034138, Toll-like receptor 3 signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	2.27×10^{-6}
GO:0051693, Actin filament capping	SPTB, SPTBN1, SPTAN1, SPTA1, ADD1, EPB49	2.97×10^{-6}
GO:0002756, MyD88-independent toll-like receptor signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	3.20×10^{-6}
GO:0015629, Actin cytoskeleton	WAS, CDH1, BRCA1, SPTB, SPTBN1, SPTAN1, CTDP1, STX1A, SPTA1, PAK1, SNCA, ADD1, EPB41, EPB49	4.25×10^{-6}
GO:0034134, Toll-like receptor 2 signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	4.43×10^{-6}
GO:0034130, Toll-like receptor 1 signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	4.43×10^{-6}
GO:0030835, Negative regulation of actin filament depolymerization	SPTB, SPTBN1, SPTAN1, SPTA1, ADD1, EPB49	4.72×10^{-6}

Table S6. *Cont.*

MIMAT0000443(hsa-miR-125a-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0002755, MyD88-dependent toll-like receptor signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	6.41×10^{-6}
GO:0050679, Positive regulation of epithelial cell proliferation	FGF2, CTNNB1, AGGF1, HIF1A, TNFSF12, ARNT, MYC, JUN, FGF1	9.31×10^{-6}
GO:0034142, Toll-like receptor 4 signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	9.31×10^{-6}
GO:0030834, Regulation of actin filament depolymerization	SPTB, SPTBN1, SPTAN1, SPTA1, ADD1, EPB49	1.09×10^{-5}
GO:0030837, Negative regulation of actin filament polymerization	SPTB, SPTBN1, SPTAN1, SPTA1, ADD1, EPB49	2.37×10^{-5}
GO:0002224, Toll-like receptor signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	2.48×10^{-5}
GO:0008629, Induction of apoptosis by intracellular signals	PML, BRCA1, YWHAB, TP53, BID, CDKN1A, RNF7, IFI16	2.92×10^{-5}
GO:0050851, Antigen receptor-mediated signaling pathway	WAS, MEF2C, RELA, PSEN2, PTPRC, PAK1, PTEN, NFKBIA	8.55×10^{-5}
GO:0002221, Pattern recognition receptor signaling pathway	ATF2, CDK1, MEF2C, FOS, RELA, JUN, ATF1, NFKBIA	9.15×10^{-5}
GO:0001936, Regulation of endothelial cell proliferation	FGF2, AGGF1, CAV1, HIF1A, TNFSF12, ARNT, JUN	9.47×10^{-5}
MIMAT0000077(hsa-miR-22-3p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0035583, Negative regulation of transforming growth factor beta receptor signaling pathway by extracellular sequestering of TGFbeta	FBN2, FBN1, LTBP1	1.45×10^{-5}
MIMAT0000265(hsa-miR-204-5p)		
GO term	**Genes**	**Adj. *p*-value**
GO:0035583, Negative regulation of transforming growth factor beta receptor signaling pathway by extracellular sequestering of TGFbeta	FBN1, FBN2, LTBP1	6.31×10^{-5}

Table S7. GOBO survival analysis results. Genes which was annotated with the specified GO term of proteins in the PIN would be used as input gene set for GOBO analysis. * $p < 0.05$, ** $p < 0.01$; *** $p < 0.001$.

miRNA	GO term	*p*-value
MIMAT0000064(hsa-let-7c)	**GO:0015630, Microtubule cytoskeleton**	**9.97×10^{-6} *** **
	GO:0043067, Regulation of programmed cell death	0.268067
	GO:0043069, Negative regulation of programmed cell death	**0.0390439 ***
	GO:0060548, Negative regulation of cell death	**0.0390439 ***
MIMAT0000076(hsa-miR-21-5p)	GO:0007173, Epidermal growth factor receptor signaling pathway	0.721139
	GO:0043067, Regulation of programmed cell death	0.266312
MIMAT0000077(hsa-miR-22-3p)	GO:0035583, Negative regulation of transforming growth factor beta receptor signaling pathway by extracellular sequestering of TGFbeta	0.940727
MIMAT0000089(hsa-miR-31-5p)	GO:0003376, Sphingosine-1-phosphate signaling pathway	0.062202
	GO:0006917, Induction of apoptosis	**0.048584 ***
	GO:0007169, Transmembrane receptor protein tyrosine kinase signaling pathway	0.050408
	GO:0007187, G-protein signaling, coupled to cyclic nucleotide second messenger	0.289466
	GO:0007188, G-protein signaling, coupled to cAMP nucleotide second messenger	0.687572
	GO:0010942, Positive regulation of cell death	0.356228
	GO:0012502, Induction of programmed cell death	**0.048584 ***
	GO:0019933, CAMP-mediated signaling	0.687572
	GO:0019935, Cyclic-nucleotide-mediated signaling	0.289466
	GO:0043067, Regulation of programmed cell death	0.694486
	GO:0043068, Positive regulation of programmed cell death	0.356228
	GO:0048011, Nerve growth factor receptor signaling pathway	0.154543
MIMAT0000250(hsa-miR-139-5p)	GO:0035583, Negative regulation of transforming growth factor beta receptor signaling pathway by extracellular sequestering of TGFbeta	0.940727
MIMAT0000265(hsa-miR-204-5p)	GO:0035583, Negative regulation of transforming growth factor beta receptor signaling pathway by extracellular sequestering of TGFbeta	0.940727
MIMAT0000423(hsa-miR-125b-5p)	GO:0002224, Toll-like receptor signaling pathway	0.380928
	GO:0002755, MyD88-dependent toll-like receptor signaling pathway	0.380928
	GO:0002756, MyD88-independent toll-like receptor signaling pathway	0.380928
	GO:0006916, Anti-apoptosis	0.0593
	GO:0006917, Induction of apoptosis	0.618064
	GO:0007169, Transmembrane receptor protein tyrosine kinase signaling pathway	0.776269
	GO:0008284, Positive regulation of cell proliferation	0.882324
	GO:0008285, Negative regulation of cell proliferation	0.883393
	GO:0008629, Induction of apoptosis by intracellular signals	0.073118
	GO:0010942, Positive regulation of cell death	0.972892
	GO:0012502, Induction of programmed cell death	0.618064
	GO:0015629, Actin cytoskeleton	0.596528

Table S7. *Cont.*

miRNA	GO term	*p*-value
MIMAT0000423(hsa-miR-125b-5p)	**GO:0015630, Microtubule cytoskeleton**	**0.028245 ***
	GO:0030834, Regulation of actin filament depolymerization	0.654383
	GO:0030835, Negative regulation of actin filament depolymerization	0.654383
	GO:0030837, Negative regulation of actin filament polymerization	0.654383
	GO:0034130, Toll-like receptor 1 signaling pathway	0.380928
	GO:0034134, Toll-like receptor 2 signaling pathway	0.380928
	GO:0034138, Toll-like receptor 3 signaling pathway	0.380928
	GO:0034142, Toll-like receptor 4 signaling pathway	0.380928
	GO:0035666, TRIF-dependent toll-like receptor signaling pathway	0.380928
	GO:0043067, Regulation of programmed cell death	0.643418
	GO:0043068, Positive regulation of programmed cell death	0.972892
	GO:0043069, Negative regulation of programmed cell death	0.492576
	GO:0048011, Nerve growth factor receptor signaling pathway	0.171634
	GO:0050678, Regulation of epithelial cell proliferation	**0.002205 ****
	GO:0050679, Positive regulation of epithelial cell proliferation	0.205483
	GO:0051693, Actin filament capping	0.654383
	GO:0060548, Negative regulation of cell death	0.413665
MIMAT0000437(hsa-miR-145-5p)	GO:0007178, Transmembrane receptor protein serine/threonine kinase signaling pathway	0.196953
	GO:0030509, BMP signaling pathway	0.196953
	GO:0090092, Regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	0.843529
MIMAT0000443(hsa-miR-125a-5p)	GO:0001936, Regulation of endothelial cell proliferation	0.115146
	GO:0002221, Pattern recognition receptor signaling pathway	0.380928
	GO:0002224, Toll-like receptor signaling pathway	0.380928
	GO:0002755, MyD88-dependent toll-like receptor signaling pathway	0.380928
	GO:0002756, MyD88-independent toll-like receptor signaling pathway	0.380928
	GO:0006916, Anti-apoptosis	0.0593
	GO:0006917, Induction of apoptosis	0.618064
	GO:0007169, Transmembrane receptor protein tyrosine kinase signaling pathway	0.776269
	GO:0008284, Positive regulation of cell proliferation	0.882324
	GO:0008285, Negative regulation of cell proliferation	0.883393
	GO:0008629, Induction of apoptosis by intracellular signals	0.073118
	GO:0010942, Positive regulation of cell death	0.972892
	GO:0012502, Induction of programmed cell death	0.618064

Table S7. *Cont.*

miRNA	GO term	*p*-value
MIMAT0000443(hsa-miR-125a-5p)	GO:0015629, Actin cytoskeleton	0.596528
	GO:0015630, Microtubule cytoskeleton	**0.028245 ***
	GO:0030834, Regulation of actin filament depolymerization	0.654383
	GO:0030835, Negative regulation of actin filament depolymerization	0.654383
	GO:0030837, Negative regulation of actin filament polymerization	0.654383
	GO:0034130, Toll-like receptor 1 signaling pathway	0.380928
	GO:0034134, Toll-like receptor 2 signaling pathway	0.380928
	GO:0034138, Toll-like receptor 3 signaling pathway	0.380928
	GO:0034142, Toll-like receptor 4 signaling pathway	0.380928
	GO:0035666, TRIF-dependent toll-like receptor signaling pathway	0.380928
	GO:0043067, Regulation of programmed cell death	0.643418
	GO:0043068, Positive regulation of programmed cell death	0.972892
	GO:0043069, Negative regulation of programmed cell death	0.492576
	GO:0048011, Nerve growth factor receptor signaling pathway	0.171634
	GO:0050678, Regulation of epithelial cell proliferation	**0.002205 ****
	GO:0050679, Positive regulation of epithelial cell proliferation	0.205483
	GO:0050851, Antigen receptor-mediated signaling pathway	0.103325
	GO:0051693, Actin filament capping	0.654383
	GO:0060548, Negative regulation of cell death	0.413665
MIMAT0002820(hsa-miR-497-5p)	GO:0001959, Regulation of cytokine-mediated signaling pathway	0.06699
	GO:0006917, Induction of apoptosis	0.142401
	GO:0007169, Transmembrane receptor protein tyrosine kinase signaling pathway	0.837635
	GO:0007173, Epidermal growth factor receptor signaling pathway	0.387447
	GO:0007178, Transmembrane receptor protein serine/threonine kinase signaling pathway	0.240167
	GO:0007179, Transforming growth factor beta receptor signaling pathway	**0.017876 ***
	GO:0008284, Positive regulation of cell proliferation	0.430255
	GO:0008285, Negative regulation of cell proliferation	0.149994
	GO:0008543, Fibroblast growth factor receptor signaling pathway	0.521978
	GO:0010942, Positive regulation of cell death	0.237692
	GO:0012502, Induction of programmed cell death	0.142401
	GO:0015629, Actin cytoskeleton	0.228804
	GO:0015630, Microtubule cytoskeleton	0.18331
	GO:0017015, Regulation of transforming growth factor beta receptor signaling pathway	0.128773

Table S7. *Cont.*

miRNA	GO term	*p*-value
MIMAT0002820(hsa-miR-497-5p)	GO:0030509, BMP signaling pathway	0.39837
	GO:0030521, Androgen receptor signaling pathway	0.383811
	GO:0032956, Regulation of actin cytoskeleton organization	0.91762
	GO:0042058, Regulation of epidermal growth factor receptor signaling pathway	0.934045
	GO:0042059, Negative regulation of epidermal growth factor receptor signaling pathway	0.789492
	GO:0043067, Regulation of programmed cell death	0.111544
	GO:0043068, Positive regulation of programmed cell death	0.237692
	GO:0043069, Negative regulation of programmed cell death	0.856892
	GO:0048011, Nerve growth factor receptor signaling pathway	0.471986
	GO:0060548, Negative regulation of cell death	0.667437
	GO:0070302, Regulation of stress-activated protein kinase signaling cascade	0.561032
	GO:0090092, Regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	0.182314
MIMAT0002856(hsa-miR-520d-3p)	GO:0006917, Induction of apoptosis	0.489781
	GO:0007169, Transmembrane receptor protein tyrosine kinase signaling pathway	0.171689
	GO:0007173, Epidermal growth factor receptor signaling pathway	0.449696
	GO:0008284, Positive regulation of cell proliferation	0.05916
	GO:0008286, Insulin receptor signaling pathway	0.237933
	GO:0008543, Fibroblast growth factor receptor signaling pathway	0.159318
	GO:0008629, Induction of apoptosis by intracellular signals	0.502822
	GO:0010942, Positive regulation of cell death	0.076906
	GO:0012502, Induction of programmed cell death	0.489781
	GO:0015630, Microtubule cytoskeleton	**0.000292 *****
	GO:0042058, Regulation of epidermal growth factor receptor signaling pathway	0.762633
	GO:0042059, Negative regulation of epidermal growth factor receptor signaling pathway	0.826854
	GO:0043067, Regulation of programmed cell death	0.740092
	GO:0043068, Positive regulation of programmed cell death	0.076906
	GO:0043069, Negative regulation of programmed cell death	0.067499
	GO:0048011, Nerve growth factor receptor signaling pathway	**0.014926 ***
	GO:0051988, Regulation of attachment of spindle microtubules to kinetochore	**7.48×10^{-6} *****
	GO:0060548, Negative regulation of cell death	0.213035

Table S8. Pathophysiogical characteristics of miRNA array data used in ROC curve analysis.

Sample Name	ER	PR	HER	TNM	Stage	Grade
S621T1	0	0	0	pT1N0M0	I	2
S434T1	1	0	1	T2N1M0	IIB	3
S403T1	0	1	0	T2N1M0	IIB	2
S459T1	1	0	0	T4N0M1	IV	1
S455N1	1	0	0	T3N3M1	IV	3
S545T1	0	0	0	pT2N0M0	IIA	3
S173N1	0	0	1	T2N3M0	IIIC	3
S363T1	1	1	0	T2N1M0	IIB	1
S909T1	1	1	1	pT3N3aM0	IIIC	(Unknown)
S645T1	0	1	1	pT1bN0M0	I	3
S898T1	1	0	0	pT2N0(i-)M0	IIA	(Unknown)
S201T1	1	0	0	T1N1M0	IIA	2
S631T1	1	1	1	T2N3aM1	IV	2
S303T1	0	0	1	T2N0M0	IIA	2
S502T1	0	0	0	pT3N0M0	IIB	3
S498N1	1	1	1	pT1cN1aM0	IIA	2
S536T1	1	0	1	T1cN1miM0	IIA	2
S660T1	0	0	1	T1N0M0	I	3
S358N1	0	0	1	T2pN2M0	IIIA	2
S665T1	0	0	0	T2N3M0	IIIC	3
S475T1	0	0	1	pT1cN0M0	I	3
S423T1	1	1	0	T2N3M0	IIIC	1
S507T1	0	0	0	pT1cN1aM0	IIA	3
S891T1	1	0	0	pT2NxM0	IIA	2
S422T1	1	1	0	T2N0M0	IIA	1
S961T1	0	0	1	T2N2aM0	IIIA	2
S622T1	0	0	0	pT2N0M0	IIA	2
S454T1	0	0	0	T1cN0M0	I	2
S433T1	1	0	0	T2N0M0	IIA	1
S673T1	0	0	0	pT2N0M0	IIA	2
S450T1	0	0	1	T2N1M0	IIB	3
S430T1	1	1	0	T2N3M0	IIIC	2
S437T1	1	0	0	T3N1M0	IIIA	3
S574T1	0	0	0	T2N0M0	IIA	2
S401T1	1	1	0	T1N0M0	I	1
S427N1	0	0	0	T3N1M0	IIIA	2
S894T1	0	0	0	pT1cN0M0	I	2
S929T1	1	0	0	pT2N0M0	IIA	2
S173T1	0	0	1	T2N3M0	IIIC	3
S622N1	0	0	0	pT2N0M0	IIA	2
S562T1	1	1	1	pT1aN0M0	I	3
S602T1	0	0	0	pT1cN0M0	I	3
S490T1	1	1	1	pT2N0M0	IIA	2
S677T1	0	0	0	pT2N1M0	IIB	3
S881T1	0	0	0	pT2N1aM0	IIB	3
S619T1	0	0	0	pT2N0M0	IIA	3
S446T1	0	0	0	T2N2M0	IIIA	3
S446T2	0	0	0	T2N2M0	IIIA	3
S453T1	1	1	1	T2N1M0	IIB	2
S562N1	1	1	1	pT1aN0M0	I	3
S557T1	0	0	0	pT2N0M0	IIA	3

Table S8. *Cont.*

Sample Name	ER	PR	HER	TNM	Stage	Grade
S594T1	1	1	1	pT2N3aM0	IIIC	3
S582T1	0	0	0	pT2N0M0	IIA	3
S358T1	0	0	1	T2pN2M0	IIIA	2
S368T1	0	0	1	T3N3M0	IIIC	2
S175T1	1	1	0	T3N3Mx	IIIC	3
S357T1	0	0	0	T2pN1M0	IIIC	3
S653T1	0	0	0	pT3N3aM0	IIIC	2
S722T1	1	1	1	pT1N0M0	I	3
S593T1	0	0	0	pT1N0M0	I	3
S543T1	1	1	1	pT2N1M0	IIB	2
S498T1	1	1	1	pT1cN1aM0	IIA	2
S389T1	1	0	0	T2N1M0	IIB	3
S614T1	0	1	1	TxN0M1	IIIA	(Unknown)
S536N1	1	0	1	T1cN1miM0	IIA	2
S462T1	1	1	0	T3N3M0	IIIC	2
S477T1	0	0	0	pT1cN0M0	I	3
S917T1	0	0	0	T2N0M0	IIA	(Unknown)
S213T1	0	0	1	T1cN1M0	IIA	3
S628T1	0	0	0	T1N0M0	I	2
S291T1	0	0	0	T3pN2M0	IIIA	3
S593N1	0	0	0	pT1N0M0	I	3
S418T1	1	1	0	T2N3M0	IIIC	2
S420T1	1	1	0	T1N0M0	I	2
S363N1	1	1	0	T2N1M0	IIB	1
S629T1	1	0	1	pT1N0M0	I	2
S586T1	1	1	1	pT2N3aM0	IIIC	2
S415T1	0	1	0	T2N2M0	IIIA	2
S439T1	1	1	0	T2N2M0	IIIA	2
S941N1	0	0	0	pT1cN1aM0	IIA	3
S893T1	0	0	0	pT2N1micM0	IIB	3
S380T1	1	0	0	T2N1M0	IIB	2
S906N1	1	1	1	pT3N3aM0	IIIC	2
S918T1	0	0	0	pT2N0M0	IIA	3
S400T1	0	0	1	T2N0M0	IIA	2
S328T1	0	0	0	T1cpN0M0	I	3
S367T1	1	0	0	T1N1M0	IIIA	1
S420N1	1	1	0	T1N0M0	I	2
S922T1	0	0	0	pT1cN0(i-)M0	I	3
S896T1	0	0	1	pT2N1aM0	IIB	2
S410T1	0	1	1	T1N0M0	I	3
S572T1	0	0	1	T4N1aM0	IIIC	3
S448T1	0	0	0	T1N0M0	I	3
S207T1	0	0	1	T2N0M0	IIA	3
S604T1	0	0	1	pT2N1M0	IIB	3
S379T1	0	0	1	T2N2M0	IIIA	2
S906T1	1	1	1	pT3N3aM0	IIIC	2
S941T2	0	0	0	pT1cN1aM0	IIA	3
S941T1	0	0	0	pT1cN1aM0	IIA	3
S375T1	1	1	0	T1N0M0	I	2
S427T1	0	0	0	T3N1M0	IIIA	2
S417T1	0	0	0	pT2N0M0	IIA	3

Table S8. *Cont.*

Sample Name	ER	PR	HER	TNM	Stage
S180T1	1	0	0	T4NxM0	IIIC
S455T1	1	0	0	T3N3M1	IV
S887T1	0	0	0	pT2N0M0	IIA
S445T1	1	0	1	T2N1M0	IIB
S469T1	1	1	0	T2N0M0	IIA
S469T2	1	1	0	T2N0M0	IIA
S483T1	1	1	0	T2N3M0	IIIC
S444T1	1	1	0	T1NxM0	(Unknown)
S909N1	1	1	1	pT3N3aM0	IIIC
S464T1	1	1	0	T1N0M0	I
S894N1	0	0	0	pT1cN0M0	I
S698T1	1	1	1	pT1cN0M0	I
S452T1	0	0	0	T1N0M0	I
S474T1	(Unknown)	(Unknown)	(Unknown)	(Unknown)	(Unknown)

References

1. Jemal, A.; Center, M.M.; DeSantis, C.; Ward, E.M. Global patterns of cancer incidence and mortality rates and trends. *Cancer Epidemiol. Biomark. Prev.* **2010**, *19*, 1893–1907.
2. Veronesi, U.; Boyle, P.; Goldhirsch, A.; Orecchia, R.; Viale, G. Breast cancer. *Lancet* **2005**, *365*, 1727–1741.
3. Cho, E.; Spiegelman, D.; Hunter, D.J.; Chen, W.Y.; Stampfer, M.J.; Colditz, G.A.; Willett, W.C. Premenopausal fat intake and risk of breast cancer. *J. Natl. Cancer Inst.* **2003**, *95*, 1079–1085.
4. Huttenhofer, A.; Schattner, P.; Polacek, N. Non-coding RNAs: Hope or hype? *Trends Genet.* **2005**, *21*, 289–297.
5. Kim, V.N.; Han, J.; Siomi, M.C. Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 126–139.
6. Kwak, P.B.; Iwasaki, S.; Tomari, Y. The microRNA pathway and cancer. *Cancer Sci.* **2010**, *101*, 2309–2315.
7. Hutvagner, G.; Zamore, P.D. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **2002**, *297*, 2056–2060.
8. Mourelatos, Z.; Dostie, J.; Paushkin, S.; Sharma, A.; Charroux, B.; Abel, L.; Rappsilber, J.; Mann, M.; Dreyfuss, G. miRNPs: A novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev.* **2002**, *16*, 720–728.
9. Nelson, P.T.; Hatzigeorgiou, A.G.; Mourelatos, Z. miRNP:mRNA association in polyribosomes in a human neuronal cell line. *RNA* **2004**, *10*, 387–394.
10. Meister, G.; Landthaler, M.; Patkaniowska, A.; Dorsett, Y.; Teng, G.; Tuschl, T. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* **2004**, *15*, 185–197.

11. Calin, G.A.; Dumitru, C.D.; Shimizu, M.; Bichi, R.; Zupo, S.; Noch, E.; Aldler, H.; Rattan, S.; Keating, M.; Rai, K.; *et al*. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15524–15529.
12. Zhang, B.; Pan, X.; Cobb, G.P.; Anderson, T.A. microRNAs as oncogenes and tumor suppressors. *Dev. Biol.* **2007**, *302*, 1–12.
13. Meng, F.; Henson, R.; Wehbe-Janek, H.; Ghoshal, K.; Jacob, S.T.; Patel, T. MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer. *Gastroenterology* **2007**, *133*, 647–658.
14. Rask, L.; Balslev, E.; Jorgensen, S.; Eriksen, J.; Flyger, H.; Moller, S.; Hogdall, E.; Litman, T.; Nielsen, B.S. High expression of miR-21 in tumor stroma correlates with increased cancer cell proliferation in human breast cancer. *APMIS* **2011**, *119*, 663–673.
15. Zaman, M.S.; Shahryari, V.; Deng, G.; Thamminana, S.; Saini, S.; Majid, S.; Chang, I.; Hirata, H.; Ueno, K.; Yamamura, S.; *et al*. Up-regulation of microRNA-21 correlates with lower kidney cancer survival. *PLoS One* **2012,** *7*, e31060.
16. Nadiminty, N.; Tummala, R.; Lou, W.; Zhu, Y.; Shi, X.B.; Zou, J.X.; Chen, H.; Zhang, J.; Chen, X.; Luo, J.; *et al*. MicroRNA let-7c is downregulated in prostate cancer and suppresses prostate cancer growth. *PLoS One* **2012,** *7*, e32832.
17. Wang, X.F.; Shi, Z.M.; Wang, X.R.; Cao, L.; Wang, Y.Y.; Zhang, J.X.; Yin, Y.; Luo, H.; Kang, C.S.; Liu, N.; *et al*. MiR-181d acts as a tumor suppressor in glioma by targeting K-ras and Bcl-2. *J. Cancer Res. Clin. Oncol.* **2012**, *138*, 573–584.
18. Zhang, Y.; Yan, L.X.; Wu, Q.N.; Du, Z.M.; Chen, J.; Liao, D.Z.; Huang, M.Y.; Hou, J.H.; Wu, Q.L.; Zeng, M.S.; *et al.* miR-125b is methylated and functions as a tumor suppressor by regulating the ETS1 proto-oncogene in human invasive breast cancer. *Cancer Res.* **2011**, *71*, 3552–3562.
19. Sachdeva, M.; Mo, Y.Y. MicroRNA-145 suppresses cell invasion and metastasis by directly targeting mucin 1. *Cancer Res.* **2010**, *70*, 378–387.
20. Guo, X.; Wu, Y.; Hartley, R.S. MicroRNA-125a represses cell growth by targeting HuR in breast cancer. *RNA Biol.* **2009**, *6*, 575–583.
21. Jin, H.; Tuo, W.; Lian, H.; Liu, Q.; Zhu, X.Q.; Gao, H. Strategies to identify microRNA targets: New advances. *N. Biotechnol.* **2010**, *27*, 734–738.
22. Friedman, R.C.; Farh, K.K.; Burge, C.B.; Bartel, D.P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **2009**, *19*, 92–105.
23. Huang, J.C.; Babak, T.; Corson, T.W.; Chua, G.; Khan, S.; Gallie, B.L.; Hughes, T.R.; Blencowe, B.J.; Frey, B.J.; Morris, Q.D. Using expression profiling data to identify human microRNA targets. *Nat. Methods* **2007**, *4*, 1045–1049.
24. Liu, B.; Li, J.; Tsykin, A.; Liu, L.; Gaur, A.B.; Goodall, G.J. Exploring complex miRNA-mRNA interactions with Bayesian networks by splitting-averaging strategy. *BMC Bioinforma.* **2009**, *10*, 408.

25. Tseng, C.W.; Lin, C.C.; Chen, C.N.; Huang, H.C.; Juan, H.F. Integrative network analysis reveals active microRNAs and their functions in gastric cancer. *BMC Syst. Biol.* **2011**, *5*, 99.
26. Farazi, T.A.; Horlings, H.M.; Ten Hoeve, J.J.; Mihailovic, A.; Halfwerk, H.; Morozov, P.; Brown, M.; Hafner, M.; Reyal, F.; van Kouwenhove, M.; *et al*. MicroRNA sequence and expression analysis in breast tumors by deep sequencing. *Cancer Res.* **2011**, *71*, 4443–4453.
27. Derfoul, A.; Juan, A.H.; Difilippantonio, M.J.; Palanisamy, N.; Ried, T.; Sartorelli, V. Decreased microRNA-214 levels in breast cancer cells coincides with increased cell proliferation, invasion and accumulation of the Polycomb Ezh2 methyltransferase. *Carcinogenesis* **2011**, *32*, 1607–1614.
28. Heyn, H.; Engelmann, M.; Schreek, S.; Ahrens, P.; Lehmann, U.; Kreipe, H.; Schlegelberger, B.; Beger, C. MicroRNA miR-335 is crucial for the BRCA1 regulatory cascade in breast cancer development. *Int. J. Cancer* **2011**, *129*, 2797–2806.
29. Sempere, L.F.; Christensen, M.; Silahtaroglu, A.; Bak, M.; Heath, C.V.; Schwartz, G.; Wells, W.; Kauppinen, S.; Cole, C.N. Altered MicroRNA expression confined to specific epithelial cell subpopulations in breast cancer. *Cancer Res.* **2007**, *67*, 11612–11620.
30. Zhu, X.M.; Wu, L.J.; Xu, J.; Yang, R.; Wu, F.S. Let-7c microRNA expression and clinical significance in hepatocellular carcinoma. *J. Int. Med. Res.* **2011**, *39*, 2323–2329.
31. Iyevleva, A.G.; Kuligina, E.; Mitiushkina, N.V.; Togo, A.V.; Miki, Y.; Imyanitov, E.N. High level of miR-21, miR-10b, and miR-31 expression in bilateral *vs.* unilateral breast carcinomas. *Breast Cancer Res. Treat.* **2012**, *131*, 1049–1059.
32. Rutnam, Z.J.; Yang, B.B. The involvement of microRNAs in malignant transformation. *Histol. Histopathol.* **2012**, *27*, 1263–1270.
33. Liu, H. MicroRNAs in breast cancer initiation and progression. *Cell Mol. Life Sci.* **2012**, *69*, 3587–3599.
34. Zhang, Z.J.; Ma, S.L. miRNAs in breast cancer tumorigenesis (Review). *Oncol. Rep.* **2012**, *27*, 903–910.
35. Mathivanan, S.; Periaswamy, B.; Gandhi, T.K.; Kandasamy, K.; Suresh, S.; Mohmood, R.; Ramachandra, Y.L.; Pandey, A. An evaluation of human protein-protein interaction data in the public domain. *BMC Bioinforma.* **2006**, *7*, S19.
36. Halbert, C.H.; Kumanyika, S.; Bowman, M.; Bellamy, S.L.; Briggs, V.; Brown, S.; Bryant, B.; Delmoor, E.; Johnson, J.C.; Purnell, J.; *et al*. Participation rates and representativeness of African Americans recruited to a health promotion program. *Health Educ. Res.* **2010**, *25*, 6–13.
37. Nicolas Diaz-Chico, B.; German Rodriguez, F.; Gonzalez, A.; Ramirez, R.; Bilbao, C.; Cabrera de Leon, A.; Aguirre Jaime, A.; Chirino, R.; Navarro, D.; Diaz-Chico, J.C. Androgens and androgen receptors in breast cancer. *J. Steroid Biochem. Mol. Biol.* **2007**, *105*, 1–15.
38. Sethi, N.; Kang, Y. Dysregulation of developmental pathways in bone metastasis. *Bone* **2011**, *48*, 16–22.

39. Liao, S.J.; Zhou, Y.H.; Yuan, Y.; Li, D.; Wu, F.H.; Wang, Q.; Zhu, J.H.; Yan, B.; Wei, J.J.; Zhang, G.M.; *et al*. Triggering of Toll-like receptor 4 on metastatic breast cancer cells promotes alphavbeta3-mediated adhesion and invasive migration. *Breast Cancer Res. Treat.* **2011**, *133*, 853–863.
40. Jiang, P.; Enomoto, A.; Takahashi, M. Cell biology of the movement of breast cancer cells: Intracellular signalling and the actin cytoskeleton. *Cancer Lett.* **2009**, *284*, 122–130.
41. Garcia, D.M.; Baek, D.; Shin, C.; Bell, G.W.; Grimson, A.; Bartel, D.P. Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nat. Struct. Mol. Biol.* **2011**, *18*, 1139–1146.
42. Lewis, B.P.; Burge, C.B.; Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **2005**, *120*, 15–20.
43. Lall, S.; Grun, D.; Krek, A.; Chen, K.; Wang, Y.L.; Dewey, C.N.; Sood, P.; Colombo, T.; Bray, N.; Macmenamin, P.; *et al*. A genome-wide map of conserved microRNA targets in *C. elegans*. *Curr. Biol.* **2006**, *16*, 460–471.
44. Krek, A.; Grun, D.; Poy, M.N.; Wolf, R.; Rosenberg, L.; Epstein, E.J.; MacMenamin, P.; da Piedade, I.; Gunsalus, K.C.; Stoffel, M.; *et al*. Combinatorial microRNA target predictions. *Nat. Genet.* **2005**, *37*, 495–500.
45. Betel, D.; Wilson, M.; Gabow, A.; Marks, D.S.; Sander, C. The microRNA.org resource: Targets and expression. *Nucleic Acids Res.* **2008**, *36*, D149–D153.
46. Tusher, V.G.; Tibshirani, R.; Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 5116–5121.
47. Keshava Prasad, T.S.; Goel, R.; Kandasamy, K.; Keerthikumar, S.; Kumar, S.; Mathivanan, S.; Telikicherla, D.; Raju, R.; Shafreen, B.; Venugopal, A.; *et al*. Human Protein Reference Database—2009 update. *Nucleic Acids Res.* **2009**, *37*, D767–D772.
48. Benjamini, Y.; Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **2001**, *29*, 1165–1188.
49. Ringner, M.; Fredlund, E.; Hakkinen, J.; Borg, A.; Staaf, J. GOBO: Gene expression-based outcome for breast cancer online. *PLoS One* **2011**, *6*, e17911.
50. Sing, T.; Sander, O.; Beerenwinkel, N.; Lengauer, T. ROCR: Visualizing classifier performance in R. *Bioinformatics* **2005**, *21*, 3940–3941.
51. Hanley, J.A.; McNeil, B.J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **1982**, *143*, 29–36.

Reprinted from *IJMS*. Cite as: Shuang, T.; Shi, C.; Chang, S.; Wang, M.; Bai, C.H. Downregulation of miR-17~92 Expression Increase Paclitaxel Sensitivity in Human Ovarian Carcinoma SKOV3-TR30 Cells via BIM Instead of PTEN. *Int. J. Mol. Sci.* **2013**, *14*, 3802-3816.

Article

Downregulation of miR-17~92 Expression Increase Paclitaxel Sensitivity in Human Ovarian Carcinoma SKOV3-TR30 Cells via BIM Instead of PTEN

Ting Shuang [1], Chunxue Shi [1,2,†], Shuang Chang [1,†], Min Wang [1,*] and Cui Hong Bai [1]

[1] Department of Obstetrics and Gynecology, Shengjing Hospital of China Medical University, Shenyang 110004, China; E-Mails: shuangting87@163.com (T.S.); c.shuang-0519@163.com (S.C.); myearth126@126.com (C.H.B.)

[2] Shenyang Maternity and infant Hospital, 90 Qingnian Road, Shenhe District, Shenyang 110014, China; E-Mail: shichunxue_502@163.com

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: wm21st@hotmail.com; Tel.: +86-18940251222; Fax: +86-024-8395-5092.

Received: 24 December 2012; in revised form: 23 January 2013 / Accepted: 4 February 2013 / Published: 8 February 2013

Abstract: To better understand the molecular mechanisms of paclitaxel resistance in ovarian carcinoma, we evaluated the expression of miRNAs using miRNA microarray between human ovarian carcinoma SKOV3 cells and paclitaxel resistant SKOV3-TR30 cells. Results showed that 69 miRNAs were upregulated while 102 miRNAs were downregulated in SKOV3-TR30 cells. Using real-time PCR, we further clarified that miR-17~92 was overexpressed in SKOV3-TR30 cells compared with that in SKOV3 cells. We then established stable virally transduced SKOV3-TR30-m-PTIP-Sponge all SKOV3-TR30 cells and its vector-only control SKOV3-TR30-m-PTIP-GFP cells. Real time-PCR revealed that SKOV3-TR30-m-PTIP-Sponge all cells expressed approximately 6.18-fold lower levels of miR-17~92 compared with the control group. Decreased expression of miR-17~92 resulted in cell cycle arrest in the G2/M phase and growth inhibition. After the transduction, the BIM protein level was increased in SKOV3-TR30

cells and luciferase reporter assays revealed that miR-17~92 binds directly to the 3'-UTR of BIM. Results of luciferase reporter assays accompanied with Western Blot showed that although miR-17~92 binds directly to the 3'-UTR of PTEN, the PTEN protein expression level was upregulated slightly while the result is of no statistical significance. Our results showed that miR-17~92 could be a causal factor of the downregulation of BIM in SKOV3-TR30 cells and thus induce the paclitaxel resistance in SKOV3-TR30 cells.

Keywords: miR-17~92 cluster; transduction; ovarian carcinoma; paclitaxel resistance; PTEN; BIM

1. Introduction

Ovarian carcinoma is one of the most common gynecological malignant tumors. Its incidence is ranked second in malignant tumors of the female reproductive system and is gradually increasing in recent years. The most recommended treatment worldwide is surgery, along with paclitaxel and platinum–based (first-line chemotherapy) adjuvant chemotherapy; however, its mortality rate is still about 70%, which is the highest in gynecological malignancies [1]. One of the most significant reasons for such high morality rate is that about 30%~40% of patients with ovarian carcinoma are resistant to chemotherapy and, moreover, 60% of first-line chemotherapy-sensitive patients are resistant to chemotherapy after six months. Therefore, clarification of the exact mechanism of resistance and resistance reversal of ovarian carcinoma has become extremely urgent and important research topic.

MicroRNAs (miRNAs) contain 18–24 nucleotides and are small, non-coding RNAs that post transcriptionally regulate gene expression through translational repression and mRNA degradation [2]. Recent research has shown the tumor suppressant and oncogenic potential of a number of miRNAs, underscoring their importance in human cancer therapeutic and diagnostic applications [3–7].

BIM is a member of the BH3-only family of pro-apoptotic proteins and is expressed in a wide variety of tissues. BIM can initiate apoptosis by directly activating Bax through interaction with the Bcl-2/Bax herterodimer complex, which can further induces mitochondrial cell death [8,9]. It plays a critical role in tumor cell biology, including the regulation of tumorigenesis through activities as a tumor suppressor, tumor metastasis, and tumor cell survival. Therefore, it has gradually become an interesting target for cancer chemotherapy. PTEN plays a well-established role in the negative regulation of the PI3K pathway, which is frequently activated in several cancer types, including ovarian cancer. PTEN loss of function occurs in a wide spectrum of human cancers through mutations, deletions and transcriptional silencing. The expression of BIM and PTEN is highly regulated by its transcriptional and post-translational levels.

Overexpression of miR-17~92 has been observed in lymphomas and solid tumors [10] and is related to cell proliferation. The gene cluster of miR-17~92 resides with intron 3 of cl3orf25

non-protein-coding gene at 13q31.3 gene [8,11]. Studies by Lewis, BP, and others [12] have shown that miR-17~92 may play a role in PTEN and BIM. In recent years, more and more studies have reported that abnormal expression of PTEN and BIM has participated in the formation of tumor drug-resistance [13–15].

Yet it is not clear if miR-17~92 gene cluster has an impact on the paclitaxel resistance of ovarian carcinoma through affecting the expression of the BIM or PTEN protein. Furthermore, the effect of both BIM and PTEN on paclitaxel resistant in ovarian cancer cells has not been thoroughly researched; particularly, the mechamism involved in their expression regulation has barely been studied in ovarian carcinoma chemoresistance.

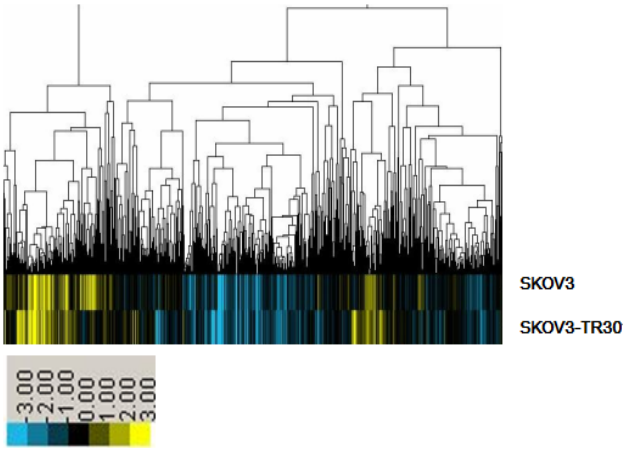
In this study, we investigated whether BIM or PTEN gene was post-transcriptionally regulated by miR-17~92 and the contribution of miR-17~92 to BIM or PTEN protein levels in SKOV3-TR30 cells. In addition, we also investigated the impact of miR-17~92 on SKOV3-TR30 cell proliferation and cell cycle. We aim to get a better understanding of the molecular mechanisms of paclitaxel resistance in SKOV3-TR30 cells, to provide a clue for further investigation of the function of miR-17~92 and its target genes, and their correlation in ovarian carcinoma paclitaxel resistance.

2. Results

2.1. MiRNA Expression Profiling

A microarray platform was used to analyze and compare the pattern of miRNA expression between the parental SKOV3 cell line and its counterpart paclitaxel resistant SKOV3-TR30 cell line. The expression profiles of 171 miRNAs changed significantly: 69 upregulated miRNAs (miR-17, miR-19b, miR-92-1) and 102 downregulated miRNAs (miR-134, miR-34, miR-196b) in SKOV3-TR30 cells as compared with SKOV3 cells (Figure 1).

Figure 1. Analysis of miRNAs Expression Profile of Sensitive and Drug-resistant Human Ovarian carcinoma Cells. Data were analyzed by the CLUSTER analysis software CLUSTER 3.0. Yellow means high expression, blue represents the low expression.



2.2. miR-17~92 Was Overexpressed in SKOV3-TR30 Cells Compared with SKOV3 Cells

We next studied the expression level of miR-17~92 in SKOV3 and SKOV3-TR30 cells using real-time PCR. Amplification Curve and Melting Curve of miR-17~92 and internal control β-actin see Figure 2. The expression level of miR-17~92 in the SKOV3-TR30 cells is indicated as "1" as control group. Compared to SKOV3-TR30 cells, the expression level of miRNA-17~92 in SKOV3 cells is 0.557, and the difference is of statistically significant ($t = 9.193$, $p < 0.05$) (Table 1).

The result revealed that the expression level of miR-17~92 was markedly increased in paclitaxel resistant SKOV3-TR30 cells compared with paclitaxel sensitive SKOV3 cells.

Figure 2. Amplification Curve and Melting Curve of miR-17~92 Expression in SKOV3, SKOV3-TR30, SKOV3-TR30-m-PTIP-Sponge all cells and the SKOV3-TR30-m-PTIP-GFP cells Detected by real-time PCR. (**a**) Amplification Curve of Expression of miR-17~92; (**b**) Melting Curve of Expression of miR-17~92.

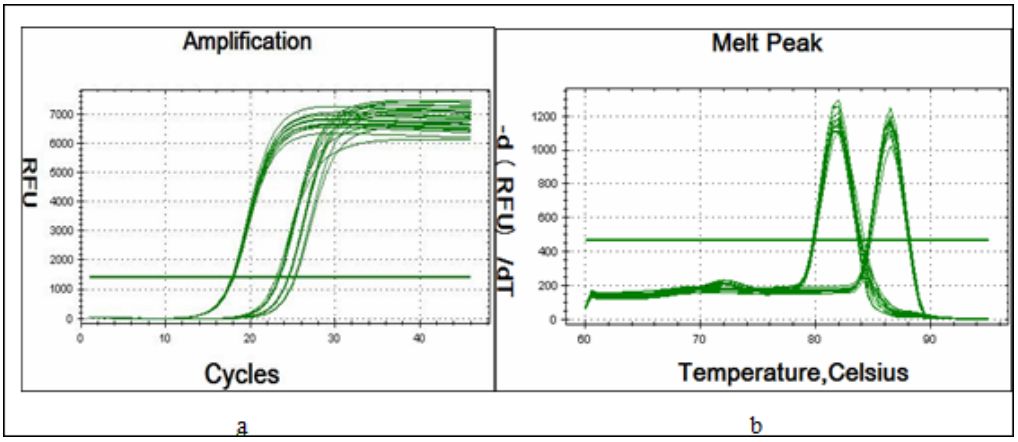


Table 1. Expression level of miR-17~92 in Different Groups.

Cells line	Δct	ΔΔct	2-ΔΔct
SKOV3	9.85 ± 0.198	0.845 ± 0.198	0.55 (0.483–0.642) *
SKOV3-TR30	10.695 ± 0.488	0.0 ± 0.488	1.0 (0.711–1.412)
SKOV3-TR30-m-PTIP-GFP	9.715 ± 0.007	0.98 ± 0.007	0.507 (0.642–0.801) *
SKOV3-TR30-m-PTIP-Sponge all	7.085 ± 0.304	3.61 ± 0.304	0.082 (0.075–0.09) *

* represents compared with the expression level of miRNA-17~92 in SKOV3, the different expression level of miRNA-17~92 in SKOV3-TR30, SKOV3-TR30-m-PTIP-GFP and SKOV3-TR30-m-PTIP-Sponge all is of great significance, $p < 0.05$.

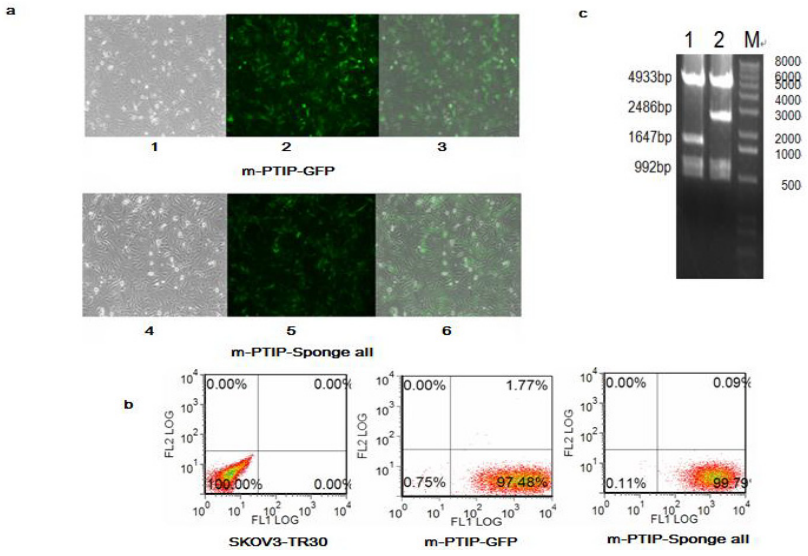
2.3. Comparison of miR-17~92 Expression Levels in SKOV3-TR30-m-PTIP-Sponge all and the SKOV3-TR30-m-PTIP-GFP

In order to determine whether the resistance to paclitaxel could be caused by miR-17~92 in SKOV3-TR30, we established stable virally transduced SKOV3-TR30 cell line

SKOV3-TR30-m-PTIP-Sponge all cells and its vector-only control SKOV3-TR30-m-PTIP-GFP cells. The m-PTIP-Sponge all inhibit plasmids were identified using restriction enzyme digestion shown in (Figure 3c). Using Flow Cytometry, the purity of stable transfectants identified by the expression of plasmid green fluorescent protein was greater than 95% (Figure 3a,b). We next studied the expression level of miR-17~92 in SKOV3-TR30 -m-PTIP-GFP cells and SKOV3-TR30-m-PTIP-Sponge all cells using real-time PCR. Real-time PCR results revealed that SKOV3-TR30-m-PTIP-Sponge all expressed approximately 6.18-fold lower levels of miR-17~92 (Table 1) compared with which in SKOV3-TR30-m-PTIP-GFP cells.

Amplification Curve and Melting Curve of miR-17~92 and internal control β-actin see Figure 2. The expression level of miR-17~92 in the SKOV3-TR30 cells is indicated as "1" as control. Compared with SKOV3-TR30 cells, the expression level of miRNA-17~92 in SKOV3-TR30-m-PTIP-GFP cells and SKOV3-TR30-m-PTIP-Sponge all cells is 0.507 (t = 9.417, p < 0.05) and 0.082 (t = 23.659, p < 0.05), respectively. Contrast the expression level of miR-17~92 in SKOV3-TR30-m-PTIP-GFP cells, which is 0.507 and in SKOV3-TR30-m-PTIP-Sponge all cells which 0.082, the difference is of statistical significance (t = 10.726, p < 0.05).

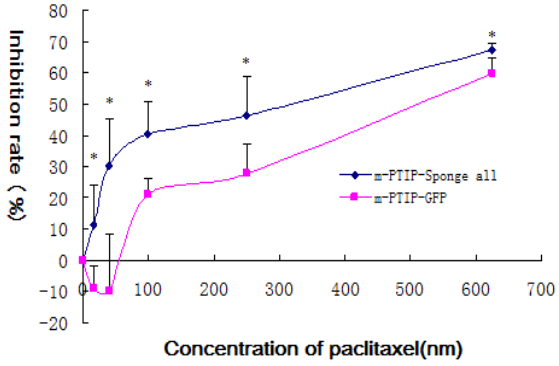
Figure 3. (**a**) Determination of stable virally transduced SKOV3-TR30-m-PTIP-Sponge all cell line and its vector-only control SKOV3-TR30-m-PTIP-GFP by Fluorescence Microscope. (100×) 1, 2, 3: transduced by m-PTIP-GFP empty plasmid 1, white field, 2, with fluorescence, 3, the integration of 1 and 2. 4,5,6: transduced by m-PTIP-Sponge all inhibit plasmid, 4, white field, 5, with fluorescence, 6, the integration of 4 and 5; (**b**) Determination of stable virally transduced SKOV3-TR30-m-PTIP-Sponge all cell line and its vector-only control SKOV3-TR30-m-PTIP-GFP using Flow Cytometry. Ratio of cells expression green fluorescent protein for SKOV3-TR30, SKOV3-TR30-m-PTIP-GFP and SKOV3-TR30-m-PTIP-Sponge all were 0.00, 97.48% and 99.79%; (**c**) Plasmid Restriction Enzyme Digestion Results 1: m-PTIP-GFP empty plasmid, 2: m-PTIP-Sponge all inhibit plasmid, M: Marker.



2.4. Inhibition of miR-17~92 could Inhibit the Proliferation of SKOV3-TR30 Cells

We further study the effects of miR-17~92 expression on the proliferation of SKOV3-TR30 using MTT assay. Paclitaxel inhibited cell proliferation in both SKOV3-TR30-m-PTIP-Sponge all and SKOV3-TR30-m-PTIP-GFP, and with increased concentration of paclitaxel the inhibitory effects was more obvious. Further more, SKOV3-TR30-m-PTIP-Sponge all cells had significantly reduced cell proliferation ($p < 0.05$) compared with SKOV3-TR30-m-PTIP-GFP (Figure 4) and when the concentration of paclitaxel was 100 nm, the inhibition rate difference was most obvious, which is 40.407% of SKOV3-TR30-m-PTIP-Sponge all cells and 20.934% of SKOV3-TR30-m-PTIP-GFP. IC50 of paclitaxel of SKOV3-TR30-m-PTIP-Sponge all cells and SKOV3-TR30-m-PTIP-GFP cells is 470.80 nM and 170.13 nM respectively.

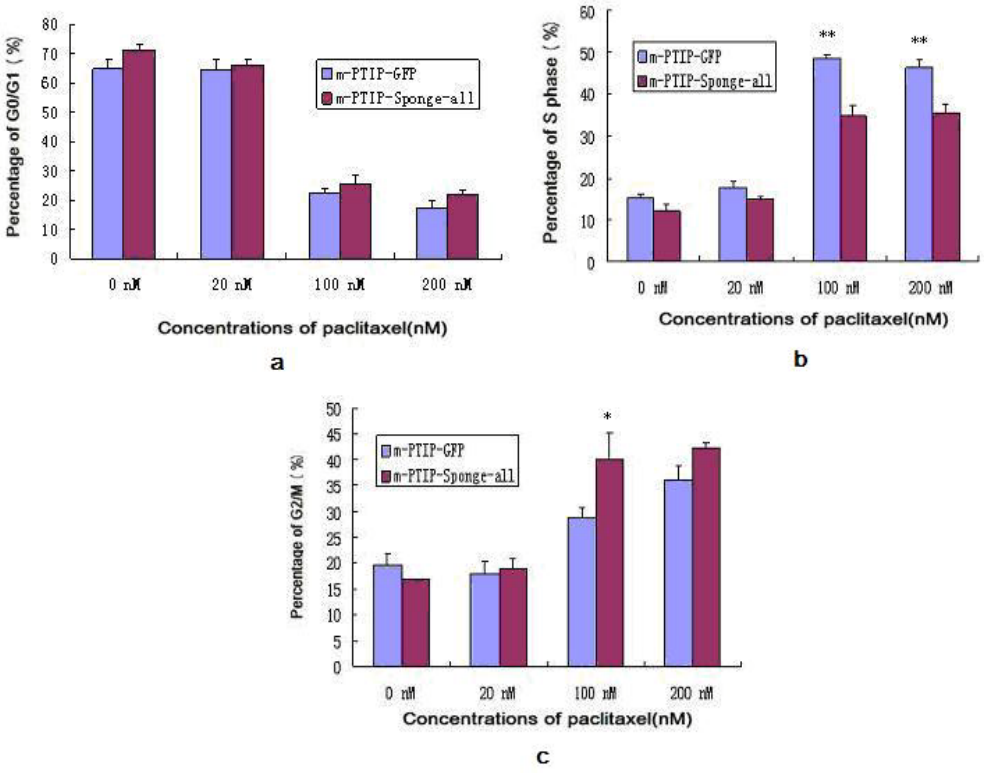
Figure 4. MTT assay was used to detect the effects of miR-17~92 on cell growth in SKOV3-TR30 cells after transduced with either m-PTIP-Sponge all inhibit plasmid or m-PTIP-GFP empty plasmid. Inhibition rate of stable virally transduced SKOV3-TR30-m-PTIP-Sponge all cells and its vector-only control SKOV3-TR30-m-PTIP-GFP cells is of great difference, which has statistical significance ($p < 0.05$). The difference is most obvious when the concentration of paclitaxel is 100 nM (* represents that under the same concentration of paclitaxel, SKOV3-TR30 cells transduced by m-PTIP-Sponge all inhibit plasmid *vs.* SKOV3-TR30 cells transduced by m-PTIP-GFP empty plasmid, $p < 0.05$).



2.5. Decreased Expression of miR-17~92 Resulted in Cell Cycle Arrest in the G2/M Phase

We accessed cell cycle distribution profiles after the transduction of SKOV3-TR30 with either miR-17~92-PTIP-Sponge all plasmids or negative control miR-17~92-PTIP-GFP (m-PTIP-GFP) plasmids. Both SKOV3-TR30-m-PTIP-Sponge all and SKOV3-TR30-m-PTIP-GFP cells were treated with paclitaxel (the concentration was 0 nM, 20 nM, 100 nM, 200 nM). Decreased expression of miR-17~92 in SKOV3-TR30-m-PTIP-Sponge all cells resulted in a significant increase in the fraction of cells arrested at G2/M as well as a concomitant decrease in the fraction of cells arrested at S phase compared with SKOV3-TR30-m-PTIP-GFP. These effects were obvious when the concentration of paclitaxel was 100 nm (Figure 5).

Figure 5. Results of Cell Cycle of SKOV3-TR30 after transduction. (**a**) The Proportion of G0/G1 Phase; (**b**) The Proportion of S Phase; (**c**) The Proportion of G2/M Phase (* respects the group transduced by m-PTIP-Sponge all inhibit plasmid *vs.* the group transduced by m-PTIP-GFP empty plasmid in the same concentration of paclitaxel, $p < 0.05$; ** respects the group transduced by m-PTIP-Sponge all inhibit plasmid *vs.* the group transduced by m-PTIP-GFP empty plasmid in the same concentration of paclitaxel, $p < 0.01$).



2.6. BIM, a BH3-only Propoapoptotic Protein, Is a Direct Target of the miR-17~92

As result shows miR-17~92 expression is significantly higher in SKOV3-TR30 than in SKOV3. To demonstrate the BIM protein level was directly mediated by miR-17~92 through biding to 3'-UTR of BIM, we co-transfected the BIM 3'-UTR (B1,B2,B3) along with TMP2-miR-17~92 into HEK293 cells and than performed luciferase reporter assays. The luciferase reporter plasmid contains a segment of the BIM 3'-UTR, among the three cloned fragments of the BIM 3'-UTR. B2 containing binding sites for miR-17-5p/-20a and miR-92, and B3 contains binding sites for miR-19 and miR-92 while there is not binding site of B1 with miR-17~92. The luciferase activity of the reporter plasmid containing the B1, B2 or B3 was decreased to 91.5%, 23.75% and 6.25% separately compared with the control construct (Figure 6b). We then used SKOV3 and SKOV3-TR30 cells to clarify if miR-17~92 has an effect on BIM protein expression which then trigger paclitaxel resistance in ovarian carcinoma cells. In contrast to SKOV3 cells, SKOV3-TR30 cells shows significantly decreased expression of BIM (Figure 7a, Table 2). We further applied SKOV3-TR30-m-PTIP-Sponge all cells and the control

SKOV3-TR30-m-PTIP-GFP cells for a better understanding about the effect of miR-17~92 on BIM protein. Both of the cell lines were treated with paclitaxel (the concentration was 0 nM, 20 nM, 100 nM, 200 nM). In contrast to SKOV3-TR30-m-PTIP-GFP cells, SKOV3-TR30-m-PTIP-Sponge all cells with a significant lower expression of miR-17~92 showed a increase of the BIM protein level both at the steady-state condition without any paclitaxel (Figure 7b, Table 3) and after the treatment with paclitaxel in different concentrations (Figure 7c, Table 3). These findings suggest that miR-17~92 plays an important role in BIM protein expression decrease through the post-transcriptional regulation.

Figure 6. (**a**) Overexpression of miR-17~92 in HEK293 cells inhibited the luciferase activity of the PTEN 3'-UTR reporter plasmid. The 3'-UTR reporter plasmids PTEN or PTEN mut were co-transfected with TMP2-miR-17~92 into HEK293 cells. The activity of firefly luciferase was normalized to that of Renilla luciferase. (**b**) Overexpression of miR-17~92 in HEK293 cells inhibited the luciferase activity of the BIM 3'-UTR reporter plasmid. The 3'-UTR reporter plasmids B1, B2 orB3 were co-transfected with TMP2-miR-17~92 into the HEK293 cells. The activity of firefly luciferase was normalized to that of Renilla luciferase. * $p < 0.05$ compared with the control.

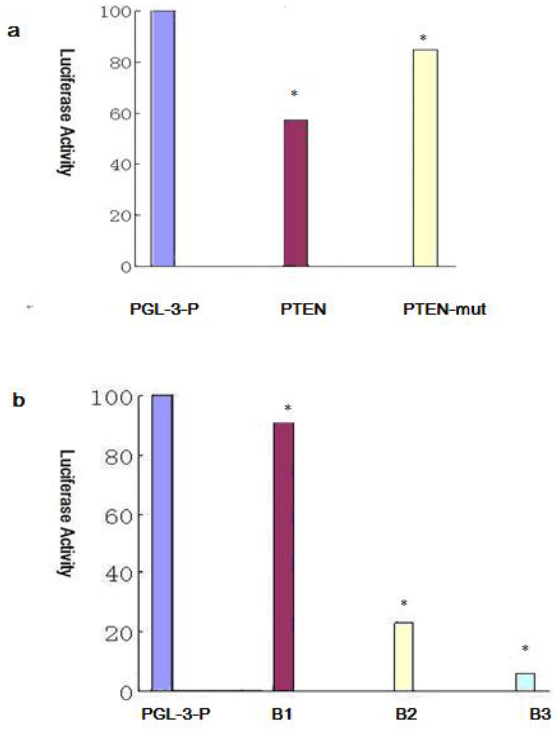


Table 2. Relative Content of BIM Protein in SKOV3-TR30 and SKOV3 cells ($\bar{x}$ ± s).

Cell group	BIM/actin
SKOV3-TR30	0.2118 ± 0.0923
SKOV3	0.2735 ± 0.1233

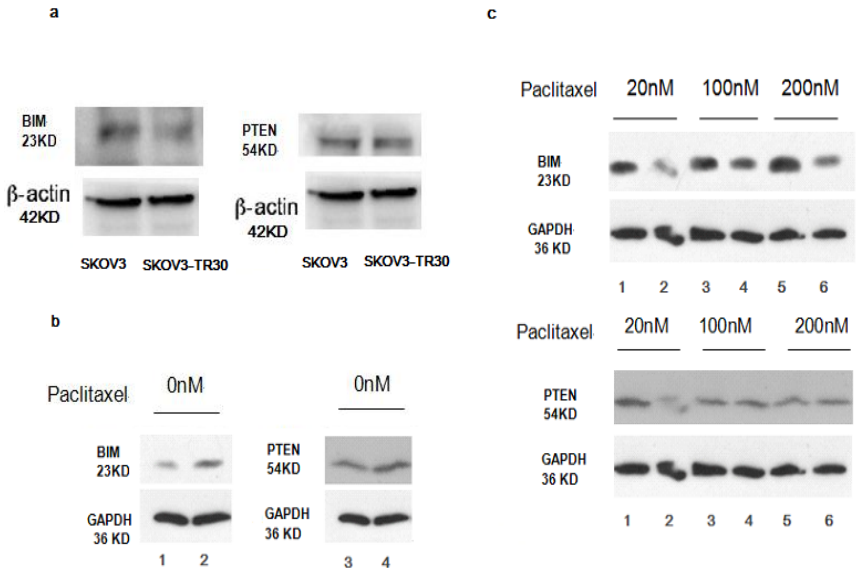
($t = 3.983$, $p = 0.028$).

Table 3. Relative Content of BIM and PTEN Protein in Each Cell Group (Relative Gray Rate, $\bar{x} \pm s$).

Concentrations of Paclitaxel (nM)	BIM/GAPDH		PTEN/GAPDH	
	m-PTIP-GFP	m-PTIP-Sponge all	m-PTIP-GFP	m-PTIP-Sponge all
0	70.59 ± 3.4215	107.70 ± 2.3265 *	122.77 ± 2.3265	137.67 ± 1.2452
20	94.13 ± 5.3265	154.08 ± 2.2641 *	98.15 ± 5.8736	121.96 ± 1.4737
100	123.72 ± 3.1749	136.83 ± 1.5821 *	122.59 ± 3.5521	130.93 ± 1.2481
200	109.54 ± 1.1399	164.46 ± 3.7326 *	116.54 ± 1.3357	125.28 ± 3.2158

* represents under the same concentration of paclitaxel, the group infected by m-PTIP-Sponge all inhibit plasmid *vs.* the group infected by m-PTIP-GFP empty plasmid, $p < 0.05$.

Figure 7. (**a**) The samples from SKOV3 cells and SKOV3-TR30 cells were subjected to Western Blot analysis with a BIM and PTEN antibody while β-actin levels were used as loading control; (**b**, **c**) The samples from SKOV3-TR30-m-PTIP-Sponge all cells and its vector-only control SKOV3-TR30-m-PTIP-GFP cells were subjected to Western Blot with BIM and PTEN antibody while GAPDH levels were used as loading control; (**b**) Decreased expression of miR-17~92 upregulates the expression of BIM instead of PTEN in SKOV3-TR30 cells without treatment of paclitaxel. 1,3 Protein samples from SKOV3-TR30 cells transduced by m-PTIP-GFP empty plasmid. 2,4 Protein samples from SKOV3-TR30 cells transduced by m-PTIP-Sponge all inhibit plasmid; (**c**) Decreased expression of miR-17~92 upregulates the expression of BIM instead of PTEN in SKOV3-TR30 cells after the treatment of paclitaxel with different concentration.1, 3, 5, Protein samples from SKOV3-TR30 cells transduced by m-PTIP-Sponge all inhibit plasmid 2,4,6: Protein samples from SKOV3-TR30 cells transduced by m-PTIP-GFP empty plasmid. The concentration of paclitaxel: 1,2: 20 nM, 3,4: 100 nM, 5,6: 200 nM.



2.7. Effect of miR-17~92 on PTEN Expression

In order to clarify whether miR-17~92 has effect on PTEN expression, we constructed a luciferase reporter plasmid containing point mutations in the predicted miRNA binding sites within the PTEN 3'-UTR (PTEN mut). We then performed luciferase reporter assays in HEK293 cells. The luciferase activity of the luciferase reporter gene containing the wild-type PTEN 3'-UTR decreased 42.6% compared with the control construct pGL-3-P, but the decrease was not seen with the luciferase reporter gene PTEN mut (Figure 6a). That means miR-17~92 binds directly to 3'-UTR of PTEN. We then used SKOV3 and SKOV3-TR30 cells to see if there is a different expression of PTEN. In contrast to SKOV3 cells, expression of PTEN protein is decreased in SKOV3-TR30 cells but the result is of no statistical significance (Figure 7a, Table 4). We then use SKOV3-TR30 to make further clarifications. Results showed PTEN protein level is slightly upregulated in SKOV3-TR30-m-PTIP-Sponge all cells compared with SKOV3-TR30-m-PTIP-GFP cells while the difference is of no statistical significance.

To sum up, expression level of PTEN was not effected; even miR-17~92 binds directly to the 3'-UTR of PTEN. In addition, although miR-17~92 expression level was much lower in SKOV3-TR30-m-PTIP-Sponge all cells, there shows no significant increased expression of PTEN protein. These observations suggested that not only miR-17~92 but also other mechanisms may be responsible for the regulation of PTEN expression, including other miRNAs involved in the post-transcription of PTEN.

Table 4. Relative Content of PTEN Protein in SKOV3-TR30 and SKOV3 cells ($\bar{x}$ ± s).

Cell group	PTEN/actin
SKOV3-TR30	0.4043 ± 0.1266
SKOV3	0.4262 ± 0.1422

($t = 2.81$, $p = 0.067$).

3. Discussion

The current work shows that miR-17~92 cluster expression correlates with paclitaxel resistance of SKOV3-TR30 cells. It's also shown that BIM instead of PTEN is suppressed by miR-17~92 cluster via direct binding to the BIM 3'-UTR.

In this study, we established a stable virally transduced SKOV3-TR30 cell line SKOV3-TR30-m-PTIP-Sponge all, which expressed approximately 6.18-fold lower levels of miR-17~92 compared with its vector-only control SKOV3-TR30-m-PTIP-GFP, in order to seek an understanding of the molecular mechanisms of paclitaxel resistance with respect to miR-17~92. At first, we approached this issue by obtaining the miRNA differential expression profile between the parental SKOV3 cells and its paclitaxel resistance SKOV3-TR30 cells by using the array-based miRNA assay. Our results showed that 69 miRNAs are upregulated while 102 miRNAs were downregulated in SKOV3-TR30 cells. Among them miR-17~92 expression was significantly upregulated in paclitaxel resistance SKOV3-TR30 cells compared with that in the parental SKOV3 cells. Using real-time PCR we further clarified that miR-17~92 was overexpressed in SKOV3-TR30 cells compared with SKOV3 cells.

A miR-17~92 cluster comprising miR-17, miR-18a, miR-20a, miR-19a, miR-19b, and miR-92-1 is overexpressed in a large fraction of lymphomas [11]. Besides, miR-17~92 was detected as overexpressed also in multiple myeloma [16], breast cancer [17] and osteosarcoma [18]. Ohyashiki [19] states that miR-92a is not only used for the diagnosis of non-Hodgkin's lymphoma, but also used as an indicator for monitoring the recurrence of non-Hodgkin's lymphoma after chemotherapy. Other studies show that the tumorigenic effect of miR-17~92 may be due to synergism between its family members [20].

Significant decreased expression of miR-17~92 in SKOV3-TR30 cells by transduction with miR-17~92 inhibitory plasmids (miR-17-92-PTIP-Sponge all) markedly inhibited cell growth and the inhibition rate is most obvious (40.4%, $p < 0.05$) when the concentration of paclitaxel was 100 nm. Decreased expression of miR-17~92 also resulted in cell cycle arrest in the G2/M phase and is most obvious when the concentration of paclitaxel was 100 nm ($p < 0.05$). At that time, the expression of BIM protein was significantly increased. Thus, it is likely that the downregulation of miR-17~92 is closely associated with the sensitivity to paclitaxel through the upregulation BIM. To further clarify whether the BIM protein level was mediated through miR-17~92, we co-transfected the BIM 3'-UTR along with TMP2-miR-17~92 plasmid into HEK293 cells, than performed luciferase reporter assays. The reporter assays revealed that the firefly luciferase activity from HEK293 cells transfected with the reporter gene containing the wild-type B1, B2 or B3 was decreased to 91.5%, 23.75% and 6.25% separately compared with the control construct (Figure 6b).These findings suggest that BIM is likely to be direct target of miR-17~92 and that BIM protein was regulated at the post-transcriptional level in SKOV-TR30 cells.

So far, the research of the influence of miRNA-17~92 on ovarian carcinoma drug-resistance is rarely reported. Using bioinformatics analysis software, we predict that PTEN and BIM are most likely potential target genes of miR17~92. The newest transgenic animal experiments in 2008 shows that miR-17~92 gene clusters down regulate the expression of PTEN and BIM, which are tumor suppressor factors [21]. In recent years, more and more studies have shown that abnormal expression of PTEN and BIM is mostly related to the formation of tumor's drug resistance [13–15]. There are also studies show down regulation of PTEN and BIM in certain ovarian carcinoma cells with feature of chemoresistance [15,22], besides studies of Lewis and BP [10] show that miR-17~92 takes effect through PTEN and BIM. So in order to further prove whether miR-17~92 cause paclitaxel resistance through PTEN in SKOV3-TR30,we tested the expression of PTEN protein in paclitaxel resistant ovarian carcinoma SKOV3-TR30 cells after transduced with miR-17~92-PTIP-Sponge all. Different from the significant upregulation of BIM protein, PTEN was upregulated slightly while the result is of no statistical significance. The luciferase activity of the reporter gene containing the wild-type PTEN 3'-UTR was decreased 42.6% compared with the control construct pGL-3-P, but the decrease was not seen with the reporter gene PTEN mut (Figure 6a). These observations altogether suggested that PTEN expression is responsible for not only miR-17~92 but also other mechanisms regulating PTEN expression including other miRNAs or other mechanism involved in post-transcription.

4. Experimental Section

4.1. Cell Culture and Plasmids

The ovarian carcinoma cell line SKOV3 was provided by Tumor Cell Bank Research Institute of the Chinese Academy of Medical Sciences. The paclitaxel-resistant ovarian carcinoma cell line, SKOV3-TR30 with the resistant ability of 27-fold greater than its parental cell line, was derived from SKOV3 cell line and provided by Zhejiang University affiliated Obstetrics and Gynecology Hospital. SKOV3 cells were maintained in McCoy's 5a medium containing fetal bovine serum with 10%, penicillin with 100 μg/mL and streptomycin with 100 μg/mL. SKOV3-TR30 cells were maintained in McCoy's 5a medium supplemented with 10% fetal bovine serum, penicillin with 100 μg/mL and streptomycin with 100 μg/mL and paclitaxel with 30 nmol/L, paclitaxel was withdrawn a week before the experiment. All cells were maintained in a humidified atmosphere containing 5% CO_2 at 37 °C. Cells in the logarithmic phase of growth were used for all studies described. HEK293 cells were maintained in a DMEM medium containing fetal bovine serum with 10% in a humidified atmosphere containing 5% CO_2 at 37 °C.

MiR-17~92 inhibitory plasmids miR-17~92-PTIP-Sponge all (m-PTIP-Sponge all), empty plasmids miR-17~92-PTIP-GFP (m-PTIP-GFP) and TMP2-miR-17~92 were provided by the Zoology Institute of Chinese Academy of Sciences.

Three plasmids named as B1, B2, B3 are cloned from the 3'-UTR of BIM. These vectors based on the pGL3-P vector were kind gifts from Professor Fu kai. B2 fragment contains miR-17-5p/-20a and miR-92 binding site and B3 fragment contains miR-19 and miR-92 binding site. There is not binding site of B1 with miR-17~92.

PTEN 3'-UTR and PTEN mut 3'-UTR were also gifts from Professor Fu kai. PTEN 3'-UTR contains putative binding sites for both miR20a/17-5p and miR-19.

4.2. MicroRNA Gene Chip

We analyze the expression profiles of miRNA on the sensitive and resistant ovarian carcinoma cell lines with using Affymetrix, the Gene Chip® analysis of miRNA chip.

4.3. Quantitative Real-Time PCR of miR-17~92 within SKOV3 and SKOV3-TR30 Cells

The expression level of mature miRNAs was determined using the TaqMan real-time quantitative PCR. Briefly, single-stranded cDNA was synthesized from 10ng of total RNA using the TaqMan MicroRNA Reverse Transcription Kit. Each cDNA generated was amplified by quantitative PCR using sequence-specific primers from the TaqMan MicroRNA Assays. PCR primers are as follows: (1) miR-17~92 gene: upstream: 5'-CAGTAAAGGTAAGGAGAGCTCAATCTG-3', downstream: 5'-CATACAACCACTAAGCTAAAGAATAATCTGA-3'; (2) internal control β-actin gene: upstream: 5'-GCAAAGACCTGTACGCCAACA-3', downstream: 5'-TGCATCCTGTCGGCAATG-3'. The

relative quantity of the target miRNAs was estimated by the $2^{-\Delta\Delta CT}$ after normalizing to the expression level of β-actin, which was detected by a TaqMan gene expression Assay.

4.4. Establishment of Stable SKOV3-TR30 Cell Lines with Induced Expression of miR-17~92 Cluster

To establish miR-17~92-PTIP-Sponge all cell line, the HEK293T cell line was co-transfected with the miR-17~92-PTIP-Sponge all plasmids and pCL packaging plasmid or empty plasmids miR-17~92-PTIP-GFP (m-PTIP-GFP) and pCL packaging plasmid by the calcium phosphate method. The virus supernatant was collected and used to infect the SKOV3-TR30 cells. The stable cell lines which have decreased expression of miR-17~92 and cell lines transduced by empty plasmids miR-17~92-PTIP-GFP (m-PTIP-GFP) were maintained in the presence of 1 μg/mL doxycycline. Stable transfectants expressing green fluorescent protein were identified by Flow Cytometry and fluorescence microscopy.

4.5. Cell Proliferation Assays

Cells were plated at a density of 5×10^3 cells/well in a 96-well plate and incubated with indicated concentrations of paclitaxel (0 nM, 20 nM, 100 nM, 200 nM, quadrupled transfected wells were analyzed for each concentration) for 48 h. The cells were then incubated with MTT reagent and the absorbance at 490 nm was determined.

4.6. Cell Cycle Analysis

Cells in the logarithmic phase of growth were selected and paclitaxel were added to the cells with concentration of 0 nM, 20 nM, 100 nM, 200 nM separately. After 48 h, cells were collected and washed in cold phosphate-buffered saline (PBS). After centrifugation at 1000*g* for 5 min, cells were fixed in cold ethanol (70%) over night, washed with cold PBS and stained with 15 μL propidium iodide (50 μg/mL) in the presence of 15 μL of RNase A (10 mg/mL) .The cells were incubated for 30 min in the dark and analyzed using Flow Cytometery.

4.7. Western Bloting

Cells were collected and washed twice in cold phosphate-buffered saline (PBS). Then cells were solubilized with lysis buffer, on ice for 30 min, and then centrifuged at 12,000 rpm for 15 min at 4 °C. The supernatants were collected, and the protein concentrations were determined using a protein assay CBB kit with a BSA standard.

Samples were loaded for gel electrophoresis at 20 μg/sample, and after electrophoresis, they were blotted onto PVDF membranes. Membranes were blocked for 2 h with blocking buffer at room temperature and then were incubated overnight at 4 °C with rabbit primary antibodies to PTEN and BIM (both from Cell Signaling 1:1000), GAPDH (from Abcam) and then for 1 h at room temperature

with HRP-conjugated anti-rabbit (1:2000) or antimouse IgG (1:2000) (both from Abcam). Membranes were then washed three times in TBS-Tween, and specific bands were visualized using the ECL system.

4.8. Luciferase Reporter Assays

HEK 293 cells were plated at 2×10^5 cells per well in a 24-well plate 24 h before transfection. The pGL-3-P promoter plasmids containing the wild type or mutated 3'-UTRs of PTEN and 3'-UTRBIM (B1,B2,B3)were co-transfected into HEK 293 cells with TMP-miR-17~92 according to the manufacturer's instructions using LIPOFECTAMINE 2000, pRL-SV40 (Promega, Madison, WI, USA) was also transfected as a normalization control. Luciferase Assays were performed 24 h after transfection using the Dual Luciferase Assay System. Firefly Luciferase activity was normalized to renilla luciferase activity for each reaction. Quintupled transfected wells were analyzed for each group.

4.9. Statistical Analysis

All values are in the expression of mean ± SD; Analysis was done using variance analysis, and comparisons among different groups were made using the Least Significant Difference (LSD) test. It is considered statistically significant when p value is less than 0.05.

5. Conclusions

In this study, we have demonstrated the significant contribution of miR-17~92/BIM to paclitaxel resistance in human ovarian carcinoma SKOV3-TR30 cells. However, we could not clearly clarify if there is a link from miR-17~92 and PTEN to paclitaxel resistance mechanisms. There are studies to show that decreased PTEN expression is partly responsible for ovarian carcinoma drug resistance [13]. In this present study, however, PTEN was not over-expressed in the resistant SKOV3-TR30 cells after transduced with miR-17~92-PTIP-Sponge all plasmid. Further studies to clarify the possible involvement of other miRNAs that might regulate the expression of PTEN in SKOV3-TR30 cells and the detailed molecular mechanism are needed in our future research.

Acknowledgments

This work is supported by the Natural Science Foundation of China (Grant Number: 30973189).

Conflict of Interest

The authors declare no conflict of interest.

References

1. Jemal, A.; Siegel, R.; Xu, J.; Ward, E. Cancer statistics, 2010. *CA Cancer J. Clin.* **2010**, *60*, 277–300.

2. Filipowicz, W.; Bhattacharyya, S.N.; Sonenberg, N. Mechanisms of posttranscriptional regulation by microRNAs: Are the answers in sight? *Nat. Rev. Genet.* **2008**, *9*, 102–114.
3. Caldas, C.; Brenton, J.D. Sizing up miRNAs as cancer genes. *Nat. Med*. **2005**, *11*, 712–714.
4. Calin, G.A.; Croce, C.M. MicroRNA signatures in human cancers. *Nat. Rev. Cancer* **2006**, *6*, 857–866.
5. Esquela-Kerscher, A.; Slack, F.J. Oncomirs-microRNAs with a role in cancer. *Nat. Rev. Cancer* **2006**, *6*, 259–269.
6. Cho, W.C. MicroRNAs: Potential biomarkers for cancer diagnosis prognosis and targets for therapy. *Int. J. Biochem. Cell Biol.* **2010**, *42*, 1273–1281.
7. William, C.S.; Cho, W.C. MicroRNAs in cancer—from research to therapy. *Biochem. Biophys. Acta* **2010**, *1805*, 209–217.
8. Kuwana, T.; Bouchier-Hayes, L.; Chipuk, J.E.; Bonzon, C.; Sullivan, B.A. Green, D.R.; Newmeyer, D.D. BH3 domains of BH3-only proteins differentially regulate Bax-mediated mitochondrial membrane permeabilization both directly and indirectly. *Mol. Cell* **2005**, *17*, 525–535.
9. Letai, A.; Bassik, M.C.; Walensky, L.D.; Sorcinelli, M.D.; Weiler, S.; Korsmeyer, S.J. Distinct BH3 domains either sensitize or activate mitochondrial apoptosis, serving as prototype cancer therapeutics. *Cancer Cell* **2002**, *2*, 183–192.
10. Ota, A.; Tagawa, H.; Karnan, S.; Tsuzuki, S.; Karpas, A.; Kira, S.; Yoshida, Y.; Seto, M. Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma. *Cancer Res.* **2004**, *64*, 3087–3095.
11. He, L.; Thomson, J.M.; Hemann, M.T.; Hernando-Monge, E.; Mu, D.; Goodson, S.; Powers, S.; Cordon-Cardo, C.; Lowe, S.W.; Hannon, G.J.; Hammond, S.M. A microRNA polycistron as a potential human oncogene. *Nature* **2005**, *435*, 828–833.
12. Lewis, B.P.; Shih, I.H.; Jones-Rhoades, M.W.; Bartel, D.P.; Burge, C.B. Prediction of mammalian microRNA targets. *Cell* **2003**, *115*, 787–798.
13. Willis, S.N.; Fletcher, J.I.; Kaufmann, T.; van Delft, M.F.; Chen, L.; Czabotar, P.E.; Ierino, H.; Lee, E.F.; Fairlie, W.D.; Bouillet, P.; *et al*. Apoptosis initiated when BH3 ligands engage multiple Bcl-2 homologs, not Bax or Bak. *Science* **2007**, *315*, 856–859.
14. Tanaka, M.; Grossman, H.B. *In vivo* gene therapy of human bladder cancer with PTEN suppresses tumor growth; down regulates phosphorylated Akt, and increases sensitivity to doxorubicin. *Gene Ther.* **2003**, *10*, 1636–1642.
15. Wu, H.; Cao, Y.; Weng, D.; Xing, H.; Song, X.; Zhou, J.; Xu, G.; Lu, Y.; Wang, S.; Ma, D. Effect of tumor suppressor gene PTEN on the resistance to cisplatin in human ovarian carcinoma cell lines and related mechanisms. *Cancer Lett.* **2008**, *271*, 260–271.
16. Gao, X.; Zhang, R.; Qu, X. MiR-15a, miR-16-1 and miR-17~92 cluster expression are linked to poor prognosis in multiple myeloma. *Leuk. Res.* **2012**, *36*, 1505–1509.
17. Ouchida, M.; Kanzaki, H. Novel direct targets of miR-19a identified in breast cancer cells by a quantitative proteomic approach. *PLoS One* **2012**, *7*, e44095.

18. Baumhoer, D.; Zillmer, S.; Unger, K.; Rosemann, M.; Atkinson, M.J.; Irmler, M.; Beckers, J.; Siggelkow, H.; von Luettichau, I.; Jundt, G.; *et al*. MicroRNA profiling with correlation to gene expression revealed the oncogenic miR-17~92 cluster to be up-regulated in osteosarcoma. *Cancer Genet.* **2012**, *205*, 212–219.
19. Ohyashiki, K.; Umezu, T.; Yoshizawa, S.; Ito, Y.; Ohyashiki, M.; Kawashima, H.; Tanaka, M.; Kuroda, M.; Ohyashiki, J.H. Clinical impact of down-regulated plasma miR-92a levels in non-hodgkin's lymphoma. *PLoS One* **2011**, *6*, e16408.
20. Grimson, A.; Farh, K.K.; Johnston, W.K.; Garrett-Engele, P.; Lim, L.P.; Bartel, D.P. MicroRNA targeting specificity in mammals: Determinants beyond seed pairing. *Mol. Cell.* **2007**, *27*, 91–105.
21. Xiao, C.; Srinivasan, L.; Calado, D.P; Patterson, H.C.; Zhang, B.; Wang, J.; Henderson, J.M.; Kutok, J.L.; Rajewsky, K. Lymphoproliferative disease and autoimmunity in mice with increased miR-17~92 expression in lymphocytes. *Nat. Immunol.* **2008**, *9*, 405–414.
22. Wang, J.; Zhou, J.Y.; Wu, G.S. Bim protein degradation contributes to cisplatin resistance. *J. Biol. Chem.* **2011**, *286*, 22384–22392.

Reprinted from *IJMS*. Cite as: Pacilli, A.; Ceccarelli, C.; Treré, D.; Montanaro, L. SnoRNA U50 Levels Are Regulated by Cell Proliferation and rRNA Transcription. *Int. J. Mol. Sci.* **2013**, *14*, 14923-14935.

Article

SnoRNA U50 Levels Are Regulated by Cell Proliferation and rRNA Transcription

Annalisa Pacilli, Claudio Ceccarelli, Davide Treré and Lorenzo Montanaro *

Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Sant'Orsola-Malpighi University Hospital, via Massarenti, 9, Bologna 40138, Italy; E-Mails: annalisa.pacilli@gmail.com (A.P.); claudio.ceccarelli@unibo.it (C.C.); davide.trere@unibo.it (D.T.)

* Author to whom correspondence should be addressed; E-Mail: lorenzo.montanaro@unibo.it; Tel.: +39-051-636-4-524; Fax: +39-051-306-861.

Received: 28 May 2013; in revised form: 1 July 2013 / Accepted: 2 July 2013 / Published: 17 July 2013

Abstract: rRNA post transcriptional modifications play a role in cancer development by affecting ribosomal function. In particular, the snoRNA U50, mediating the methylation of C2848 in 28S rRNA, has been suggested as a potential tumor suppressor-like gene playing a role in breast and prostate cancers and B-cell lymphoma. Indeed, we observed the downregulation of U50 in colon cancer cell lines as well as tumors. We then investigated the relationship between U50 and proliferation in lymphocytes stimulated by phytohemagglutinin (PHA) and observed a strong decrease in U50 levels associated with a reduced C2848 methylation. This reduction was due to an alteration of U50 stability and to an increase of its consumption. Indeed, the blockade of ribosome biogenesis induced only an early decrease in U50 followed by a stabilization of U50 levels when ribosome biogenesis was almost completely blocked. Similar results were found with other snoRNAs. Lastly, we observed that U50 modulation affects ribosome efficiency in IRES-mediated translation, demonstrating that changes in the methylation levels of a single specific site on 28S rRNA may alter ribosome function. In conclusion, our results link U50 to the cellular proliferation rate and ribosome biogenesis and these findings may explain why its levels are often greatly reduced in cancers.

Keywords: snoRNA U50; ribosome methylation; IRES-mediated translation; colon cancer

1. Introduction

Ribosome biogenesis is a highly coordinated process occurring in the nucleolus, where a polycistronic pre-ribosomal RNA (pre-rRNA) transcript is processed to generate the mature 18S, 5.8S, and 28S rRNA. During this processing, the rRNA sequences undergo extensive covalent nucleotide modification, largely directed by small nucleolar RNA (snoRNA)-protein complexes (snoRNP) [1,2]. SnoRNA may be divided into two classes: The H/ACA and the C/D box, mediating pseudouridylation and 2'*O*-methylation of specific sites, respectively [3,4]. In particular, the methylation reaction is guided by an extensive region (10–21 nt) of complementarity between the C/D box snoRNA and rRNA sequences flanking the modification site [5–8]. In mammals, snoRNAs are transcribed by the RNA Polymerase II being localized within the introns of snoRNA host genes. These host genes are also transcribed for either protein coding or noncoding mRNAs [9] which often contain a 5' terminal oligopyrimidine (5' TOP) sequence responsible for their translational upregulation in response to growth factors or other conditions requiring increased protein synthesis (reviewed by Meyuhas *et al.* [10]); however, the precise function of 5' TOP motif with respect to snoRNA synthesis is unknown [11]. Furthermore, the development of three-dimensional maps of the modified nucleotides in the ribosomes of Escherichia coli and yeast has revealed that rRNA modifications occur in conserved and functionally important regions for subunit–subunit and nascent protein interactions, for tRNA and mRNA binding, but not in those interacting with proteins (see [12,13]). This correlation indicates that modifications influence both the structure and the function of the ribosome [14]. Indeed, there is evidence that post-transcriptional rRNA modifications, including pseudouridylation and methylation, affect ribosomal function [15–17] and that alterations in this modification pattern might be involved in human diseases, such as ribosomopathies and tumorigenesis [18,19]. Recent reports have demonstrated that somatic rearrangements, mutations, or the reduction in the expression of the C/D snoRNA U50 have been found in breast carcinomas, prostate cancer, and B-cell lymphomas [20–23]. snoRNA U50 is known to mediate the methylation of ribose residues corresponding to the cytosines in positions 2848 and 2863 in 28S ribosomal RNA [5,23]. In breast cancer cell lines, the reintroduction of U50 is able to induce cell death, suggesting a tumor-suppressor-like behavior for this snoRNA [23]. The human snoRNA U50 sequence is localized in the 5th intron of the non-coding host gene named small nucleolar RNA host gene 5 (SNHG5) [20,24], which is a member of the 5' TOP gene family.

In this paper we investigated the relationship between snoU50 and cancer in colon cancer cell lines and tumors with particular regard to proliferation.

2. Results and Discussion

2.1. Evaluation of U50 Levels in Colorectal Cancer Tissues and Cell Lines

The evaluation of snoRNA U50 levels on colon cancer tissues was performed on both tumor and normal tissues in a series of 34 patients. We found that U50 was downregulated in tumor tissues if compared to the normal counterpart and this reduction was statistically significant in a subgroup of low-stage tumors ($p = 0.047$) (Figure 1A, left). The decrease in U50 levels in tumors was in line with previous reports, demonstrating the same behavior in prostate and breast cancers [21,23]. In order to find a key element linking U50 and tumorigenesis, we performed a correlation analysis with the available clinical and bio-pathological features of tumors (see Table S1) and we found a significant association with the tumor grade. Indeed, high-grade tumors displayed lower U50 levels in comparison to those observed in low-grade tumors ($p = 0.049$) (Figure 1A, right). We then evaluated U50 expression in a panel of eleven colon cancer cell lines. We found that U50 expression is highly variable between lines, but always lower than normal colon tissues (NT) (Figure 1B, left), while the overall comparison between NT and colon cancer cell lines showed a statistically significant difference for U50 expression ($p = 0.0004$) (Figure 1B, right-top). Furthermore, we found that the U50 expression in lines derived from primary tumors (HCT, SW480, RKO, HCA7, CaCo-2, La174T, HT29, SW48) was significantly different from that in those derived from metastatic tumors (Colo205, SW620, LoVo-$p = 0.0121$) (Figure 1B, right-bottom).

Figure 1. Downregulation of U50 levels in colon cancer tumors and cell lines. (**A**) Evaluation of U50 levels in both tumoral and adjacent untransformed colon tissues on all samples ($n = 34$) and in a subgroup of low-stage tumors ($n = 20$) (left); comparison of U50 levels between low- and high-grade tumors ($n = 27$ and $n = 7$, respectively); (**B**) Evaluation of U50 levels in a cohort of colon cancer cell lines (left); comparison between U50 levels in normal colon tissues and cell lines (right-top) and between cell lines derived from primary or metastatic tumors (right-bottom). The results correspond to means ± S.E.M. of three different experiments. * $p \leq 0.05$; *** $p \leq 0.001$; NT, normal tissues; TT, tumoral tissues; T, tumors.

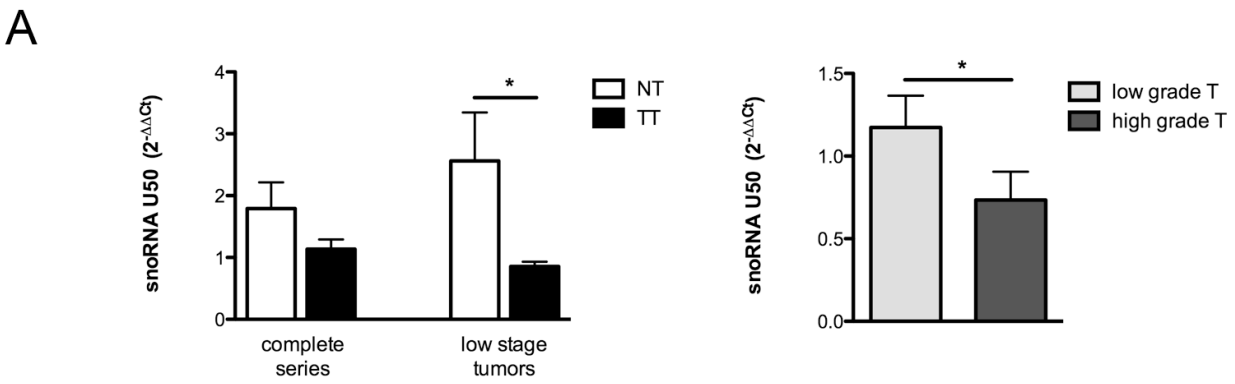
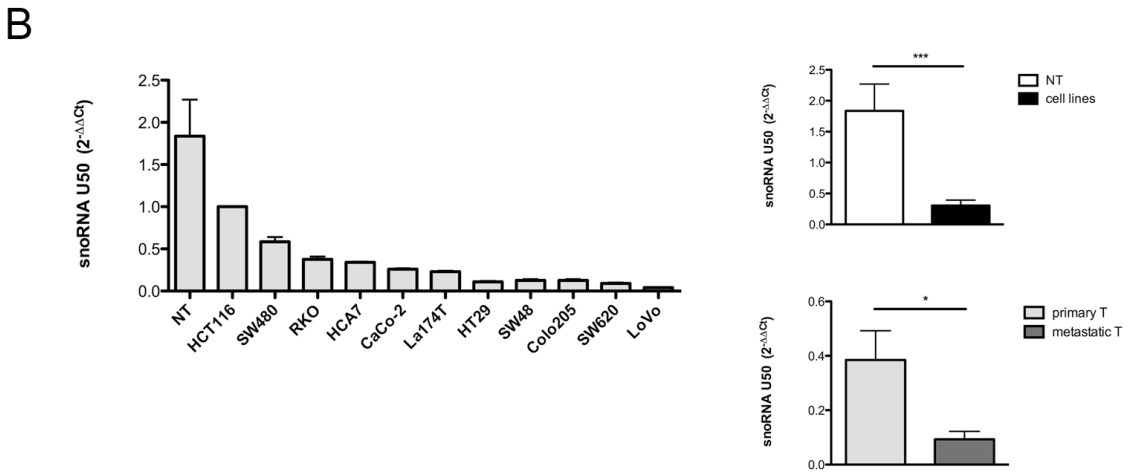


Figure 1. *Cont.*



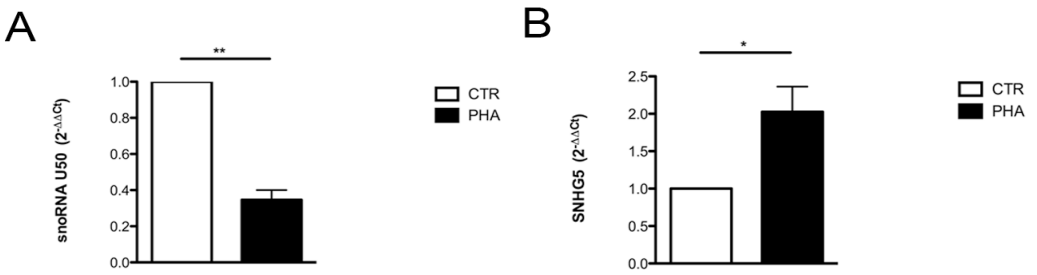
2.2. Relationship between U50 and Proliferation

Compared to untransformed tissues, U50 was downregulated in tumors. Being that one of the major distinguishing characteristics between these tissues is their proliferation activity, we investigated whether there is a relationship between U50 and proliferation.

In order to study changes in U50 synthesis depending on cellular proliferation, we used a well-established model of investigation, *i.e.*, the comparison of resting and stimulated primary human lymphocytes. Thus we isolated lymphocytes from healthy donors and performed lymphocyte proliferation assays by using phytohemoagglutinin (PHA). After PHA stimulation, we evaluated the expressions of both U50 and its host gene SNHG5 and observed a strong reduction of U50 levels in PHA stimulated cells, compared to controls (Figure 2A). Conversely, the SNHG5 host gene was upregulated (Figure 2B). Since snoU50 and SNHG5 are simultaneously transcribed, this discrepancy might be explained by a different stability of the two RNAs. The massive decrease in U50 led us to hypothesize that during cellular proliferation U50 stability might be extremely reduced. To prove this hypothesis, we performed mRNA stability assay on lymphocytes—Both stimulated and not stimulated by PHA—By treating cells with the transcriptional inhibitor Actinomycin D (Act-D) at an 80 nM concentration. At this concentration, Act-D is known to abolish the activity of RNA polymerases I and II. We observed that the Act-D treatment did not reduce U50 stability in non-proliferative control cells, but rather induced a progressive accumulation of this snoRNA depending on the length of Act-D treatment (Figure 3). On the contrary, in PHA stimulated cells, we found an early decrease in U50 levels up to 3 h of treatment, followed by a general stabilization of U50 levels (Figure 3). The fluctuation of U50 levels observed at 3 h may be the result of the combined effect of time needed to obtain the full inhibition of ribosome processing by Act-D treatment and the regulation of SNHG5 transcription as documented in Figure S1. Furthermore, PHA treatment induced a notable decrease in U50 levels but an increase in SNHG5 levels. The different levels of our transcripts in the two conditions before Act-D treatment and the behavior of mature SNHG5 mRNA

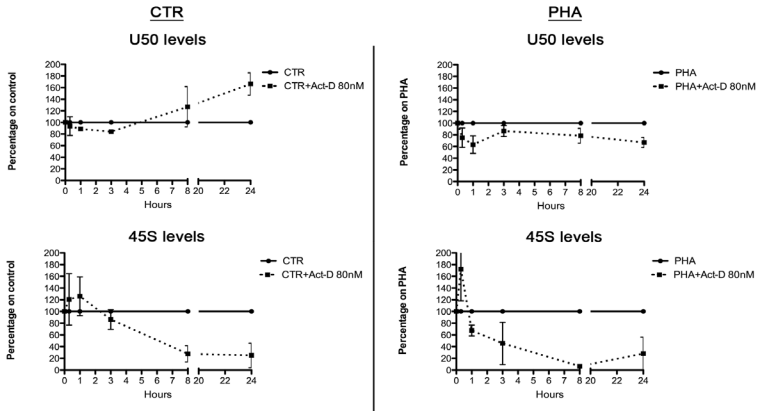
levels observed after 80 nM Act-D treatment excludes the possibility that the increase of U50 levels after treatment was due to an increase of its transcription.

Figure 2. Regulation of U50 levels in response to cellular proliferation. Evaluation of U50 (**A**); and its host gene SNHG5 (**B**) levels in human lymphocytes stimulated or not to proliferate with PHA for 72 h. The results correspond to means ± S.E.M. of four different experiments. * $p \leq 0.05$; ** $p \leq 0.01$. CTR, controls; PHA, phytohemoagglutinin.



We hypothesized that the observed stabilization of U50 was a consequence of the concomitant transcriptional inhibition of rRNA genes, which implied a reduced post-transcriptional modification of rRNA, and thus a lower U50 consumption. This would be in line with the variation of 45S rRNA levels found in both control and PHA-treated lymphocytes (Figure 3). The different regulation of 45S levels over time observed in control and PHA treated cells can be explained by the different activity of rRNA processing in the two conditions. Taken together, these results showed that U50 and 45S had opposite behaviors. This is particularly clear in control cells, since the ribosome biogenesis is not as intensive and U50 did not have to be massively consumed.

Figure 3. Relationship between U50 and rRNA transcription. Evaluation of U50 and 45S levels in human lymphocytes controls or PHA-stimulated in the presence or absence of high doses of the transcriptional inhibitor Act-D (80 nM) in order to block both polymerases I and II. The analyses were performed 20 min, 1 h, 3 h, 8 h, and 24 h after Act-D treatment. The results correspond to means ± S.E.M. of three different experiments. CTR, controls; PHA, phytohemoagglutinin; Act-D, actinomycin-D.

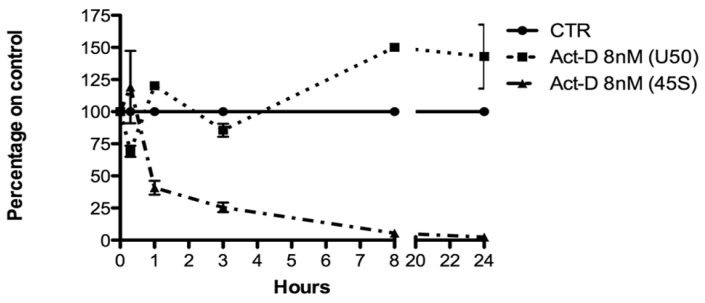


2.3. Relationship between U50 and rRNA Transcription

The accumulation of U50 induced by the block of total transcription and the inverse relationship between U50, proliferation, and 45S levels led us to suppose that U50 synthesis could be regulated by the ribosome biogenesis rate, and so we selectively reduced rRNA transcription in SW620 colon cancer cells. We chose this line since the levels of U50 are quite low in presence of a very active ribosome biogenesis activity.

A tenfold lower Act-D concentration was used to selectively inhibit polymerase I only. This selectivity was proved by measuring the transcription levels of a housekeeping mRNA (β-glucuronidase-GUS) and SNHG5, whose genes were both transcribed by polymerase II. Results demonstrated that none were affected by Act-D 8nM treatment (Figure S2). The SNHG5 transcript was also measured to exclude the possibility that U50 levels could be regulated by its transcription. After 45S synthesis inhibition, U50 levels increased (Figure 4). This result obtained after the blockade of rRNA transcription confirmed our hypothesis that U50 and cellular proliferation are inversely associated because of its massive consumption and not because of its transcription. Indeed, U50 is implicated in post-transcriptional methylation of 28S rRNA, and it is well known that during proliferation there is an increase in ribosome biogenesis requiring modification of rRNA molecules. The levels of U50 and 45S transcript displayed opposite behaviors, thus confirming that the strict dependence of U50 on the rRNA transcription rate might be due to its biological function. Taken together, these data indicated a putative role of U50 and snoRNAs in growth, tumorigenicity, and metastasis. We then investigated whether the relationship between ribosome biogenesis and U50 is shared by other snoRNAs and evaluated the cellular levels of 3 additional C/D box snoRNAs such as U33, U34, and U56 housed in genes coding for the ribosomal protein L13A (U33 and U34) and for the nucleolar protein 5A, which is a component of snoRNPs complex (Nop56). We found that these selected snoRNAs displayed the same regulation and that, in all cases, the accumulation started after 3 h and progressively increased with the length of treatment (Figure S3).

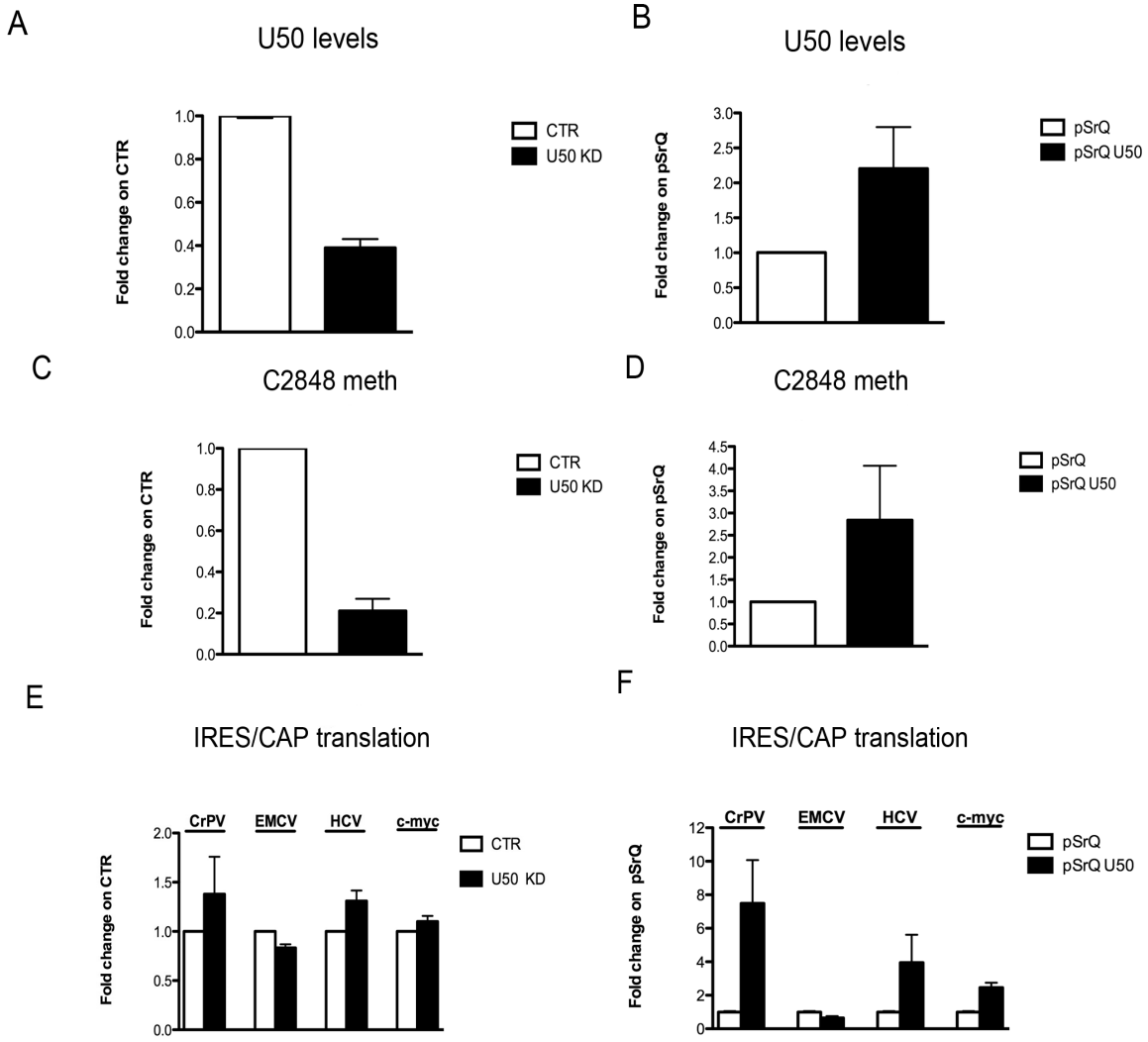
Figure 4. Relationship between U50 and *de novo* rRNA transcription in SW620 cell line. Evaluation of U50 stability performed by measuring U50 and 45S levels in SW620 cells after blockade of Polymerase I with low doses of Act-D. The analyses were performed 20 min, 1 h, 3 h, 8 h, and 24 h after Act-D treatment. The results correspond to means ± S.E.M. of three different experiments. CTR, controls; Act-D, actinomycin-D.



2.4. Effect of U50 Regulation on rRNA Site-Specific Methylation and Ribosome Activity

Prompted by these observations, we investigated the possible effect of U50 modulation on mRNA translation. Indeed, U50 is responsible for the site-specific C2848 methylation on 28s rRNA, and It has recently been demonstrated that incorrect methylation of rRNA is associated with an impaired capacity to initiate translation through Internal Ribosome Entry Site (IRES) [19,25] by influencing the mechanism of 80S complex formation on IRES elements [25]. Actually there are studies reporting that changes in snoRNAs are crucial for ribosome biogenesis or function [26,27], and showing an increase in non-coding host genes either after inhibition of translation and elongation [11,28] or in growth arrest conditions [29]. However, it is not clear whether the modulation of a single snoRNA might also contribute to altering ribosome efficiency. In this paper, we focused both on the role of U50 modulation in C2848 site-specific methylation, and on IRES-mediated translation. Therefore, we semi-quantitatively evaluated the changes in C2848 methylation in PHA stimulated cells and observed a decrease in site-specific C2848 methylation in response to proliferation (Figure S4). To better understand the role of U50 and its influence on methylation and IRES-mediated translation, we regulated U50 levels by specific knockdown (KD) and overexpression. We then evaluated the effect on C2848 of 28S methylation, on the translation of known viral IRESes types (Cricket Paralysis Virus-CrPV, Encephalomyocarditis virus-EMCV and Hepatitis C Virus-HCV), and on a cellular IRES (c-Myc) known to be affected by rRNA methylation [25]. We performed U50 KD on HCT116 cell lines and observed that, 72 h after LNA transfection, we obtained a rather strong U50 KD efficiency associated with a decrease in C2848 methylation (Figure 5A,C). Thus we decided to evaluate IRES-mediated translation after 72 h of U50 KD. We found that U50 KD did not significantly alter ribosome activity (Figure S3, top), although we observed a slight increase in CrPV- and HCV IRESes-mediated translation (Figure 5E). It is likely that we could not appreciate notable differences in ribosome translation activity in the U50 KD model because the experiments were performed in a tumor cell line with already low U50 levels in comparision to untransformed cells.

Figure 5. Effect of U50 modulation on C2848 methylation and IRES-mediated translation. Evaluation of U50 levels (**A** and **B**); C2848 methylation status (**C** and **D**); and IRES-mediated translation (**E** and **F**) in HCT116 cells after 72 h of U50 knockdown (left panel) and U50 overexpression (right panel). The results correspond to means ± S.E.M. of three different experiments. CTR, controls; KD, knockdown; pSrQ, pSIREN-RetroQ vector (empty vector); pSrQ U50, pSIREN-RetroQ-U50 vector; CrPV, CRicket Paralysis Virus; EMCV, EncephaloMyoCarditis Virus; HCV, Hepatitis C Virus.



Furthermore, in order to investigate the effect of U50 overexpression, we used the pSrQ-U50 vector to stably transfect HCT116, SW620, and LoVo cell lines. In addition to HCT116, we chose SW620 and LoVo cells since they displayed the lowest U50 basal level in our cohort of colon cancer lines, as previously reported. We successfully obtained all U50 overexpressing lines, but an increase in C2848 methylation was achieved for the HCT116 line only (Figures 5B,D and S6). In contrast with results after U50 KD, in the presence of higher C2848 methylation levels, we found a general decrease in CAP-mediated translation and an increase in IRES-mediated translation (Figure S5, bottom). These

observations were consistent for CrPV, HCV, and c-Myc IRESes, while EMCV IRES translation was not affected by changes in the methylation of this specific site (Figure 5F). These results might be explained by the augmented ribosome affinity for IRESes secondary structures, which was probably caused by changes in the ribosome structure mediated by C2848 methylation status.

3. Experimental Section

3.1. Patient Materials

Both tumor and adjacent non-tumor (hereinafter referred to as "normal tissue") colon tissues selected from a series of thirty-four consecutive patients who had undergone surgical resection for primary carcinoma at the Surgical Department of the University of Bologna, on the sole basis of frozen tissue availability. For each patient, clinical information was recorded and the corresponding tissue was histologically characterized and processed following standard procedures to define bio-pathological features. Specimen collection and tissue analyses were approved by the Bologna University Ethical Committee on human tissue research. Tissues were preserved at −80 °C until use. A 60 mg piece for each sample was minced in liquid nitrogen and then lysed for total RNA extraction using Tri-reagent solution (Ambion, Life Technologies Corporation, Monza, Italy) according to the manufacturer's instructions.

3.2. Cell Cultures and Treatments

All colon cancer cell lines were grown in their optimal culture medium and conditions at 37 °C in 5% CO_2 according to the American Type Culture Collection instructions.

Lymphocytes were recovered from healthy donors' buffy coats collected by the transfusion center of Sant'Orsola-Malpighi Hospital in Bologna after ficoll gradient and monocyte depletion. Human lymphocytes were grown in RPMI-1640 medium (Sigma, Milan, Italy), supplemented with 10% FBS, 1% of penicillin and streptomycin, 1% of glutamine, and 0.5% of non-essential amino acids (all Sigma). Lymphocyte proliferation assays were performed by adding 10 μg/mL of phytohemoagglutinin (PHA) (Sigma) to the medium.

The inhibition of ribosome biogenesis and mRNA stability assay were performed by exposing cells to Actinomycin D (Act-D) (Sigma) at a concentration of 8 nM and 80 nM, respectively, and assessed at different time points (20 min, 1 h, 3 h, 8 h, and 24 h of treatments). SW620 cells were seeded at 100,000 cells/well in 6-well plates and treated with Act-D, while human lymphocytes were seeded at 1,000,000 cells/mL, PHA stimulated for 72 h (where necessary), and then treated with Act-D. PHA stimulation was evaluated by cell counting and considered effective only when lymphocytes proliferation was increased more than 5 fold compared to control cells.

3.3. Gene Expression Assays

Total RNA was extracted from cell lines using Tri-reagent (Ambion) and 1 μg was reverse-transcribed using the High-Capacity cDNA Archive Kit (Applied Biosystems, Carlsbad, CA, USA). Real-time PCR analysis was carried out in a Gene Amp 7000 Sequence Detection System (Applied Biosystems) using the TaqMan approach for β-glucuronidase (GUS) (Applied Biosystems) and SYBR green approach for all other genes, using specific couples of primers. The housekeeping GUS gene was used to calculate the relative amounts of the studied target genes in all experiments except for those in which Act-D treatment was performed. In the latter case, for each sample we calculated the $2^{\wedge -Ct}$, assumed the values of untreated cells as reference (100%), and then calculated the relative percentage of investigated genes. For each sample, three replicates were analyzed. Analyses in cell lines and tumors were carried out by using aliquots of a single stock of cDNA obtained from HCT116 as internal calibrator. The analysis of U50 levels was performed by qPCR using a primer with a linker sequence attached to a U50-specific sequence for cDNA synthesis and reverse linker specific primer, and U50 forward specific primer for real-time PCR as described in Dong, 2008 [21]. All primers used were reported in Table S2.

3.4. SnoRNA U50 Knockdown and Upregulation

Since it has been shown that RNAi is an unsuitable tool for snoRNA KD [30], we used an oligomediated RNaseH cleavage strategy. A custom-made hybrid DNA-RNA Locked Nucleic Acid (LNA)-antisense specific for U50 (Exiqon, Vedbaek, Denmark) was transfected into cells with Lipofectamine RNAiMAX reagent (Invitrogen, Life Technologies, Carlsbad, CA, USA), while the efficiency of KD was evaluated after 24 h and 72 h. The same amount of a control oligo sequence (TCGAGCGGCCGCCCGGGC) was transfected with lipofectamine in control cells.

For U50 upregulation, we used a retroviral transduction system. We briefly transfected the Phoenix A cell line with 10 μg of pSIREN-RetroQ-U50 plasmid or the pSIRENRetroQ vector control (gifts of Prof. Dong) [21] using Lipofectamine 2000 reagent, and collected the supernatant (containing viruses) after 24 h and 48 h. After that we infected HCT116, SW620, and LoVo cell lines previously seeded in 6-well plates with viruses using standard spinoculation protocols. Seventy-two hours after infection, cell lines were selected for at least 10 days, while adding puromycin into the media at a final concentration of 0.7 μg/mL. The upregulation of U50 was verified by qPCR as described above.

3.5. Semi-Quantitative Site-Specific Methylation Assay

To evaluate site-specific (C2848) methylation levels, we followed a modified previously described method [19]. This modified method consists of the reverse transcription (RT) of total RNA with Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV) by using a 28S specific oligonucleotide (targeting the methylation site downstream, at either low or high dNTPs concentrations, either affecting cDNA synthesis or not, respectively, in the presence of modified sites. The cDNAs

resulting from both RT are evalutated by qPCR; the ratio between high and low dNTPs-derived products semi-quantitatively indicates the modification of the site downstream of the oligonucleotide used. An endogenous housekeeping RNA (GUS) was used to evaluate the efficiency of reverse transcription.

3.6. mRNA Transfection and Internal Ribosome Entry Site (IRES) Translation

Capped mRNA was transcribed from linearized pR-CrPV-IRES-F (gift of Dr. D. Ruggero) [31], pF-EMCV-IRES-R (gift of Prof. A.C. Palmenberg) [32], pR-HCV-IRES-F (gift of Prof R.E. Lloyd) [33], and pR-c-MYC-IRES-F [34], by using the mMessage mMachine T7 or T3 kits (Ambion). Cells were transfected with 0.4 μg RNA/sample using Lipofectamine 2000 (Invitrogen) following the manufacturer's instructions. After 4 h transfection, media were changed and 2 h later cells were harvested and analyzed with a dual-luciferase assay kit (Promega, Madison, WI, USA) according to the manufacturer's instructions.

3.7. Statistical Analysis

All data were analyzed using Prism software, version 5.0a. A paired *t*-test was used for patient data set when comparing normal and correspondent tumoral tissues. A two-sided Student's *t*-test was used for the comparisons between two groups. p values <0.05 were regarded as statistically significant.

4. Conclusions

Recent reports have demonstrated that U50 was downregulated in prostate and breast cancers and in B-cell lymphomas [20,21,23]. In this study we have demonstrated for the first time that snoRNA U50 expression: (i) is also downregulated in colon cancer; (ii) is reduced during cell proliferation; and (iii) its levels are inversely associated with ribosome biogenesis. Furthermore, since U50 mediates site-specific methylation on C2848 of 28S rRNA, we investigated the effect of its modulation on ribosome activity and showed that proliferating cells with low levels of both U50 and C2848 methylation only moderately changed their translational activity, while the U50 and C2848 methylation increments brought about an IRES-mediated translation propensity. At the present time the biological significance of these observations in cancer progression is unclear, and further investigations will aim to investigate which genes are differently expressed and what their roles might be in human cancers and other pathologies.

Acknowledgements

We thank the "Centro Interdipartimentale di ricerche sul cancro-Giorgio Prodi (CIRC)" for supporting AP. We are grateful to Dong for the pSIREN-RetroQ-U50 plasmid. This work was supported by PRIN grant from the Italian Ministry of Education, University and Research n. 20104AE23N_002 to LM and by the Pallotti Legacy for Cancer Research.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Filipowicz, W.; Pogacić, V. Biogenesis of small nucleolar ribonucleoproteins. *Curr. Opin. Cell Biol.* **2002**, *14*, 319–327.
2. Smith, C.M.; Steitz, J.A. Sno storm in the nucleolus: New roles for myriad small RNPs. *Cell* **1997**, *89*, 669–672.
3. Ni, J.; Tien, A.L.; Fournier, M.J. Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* **1997**, *89*, 565–573.
4. Kiss, T. Small nucleolar RNAs: An abundant group of noncoding RNAs with diverse cellular functions. *Cell* **2002**, *109*, 145–148.
5. Kiss-Laszlo, Z.; Henry, Y.; Bachellerie, J.P.; Caizergues-Ferrer, M.; Kiss, T. Site-specific ribose methylation of preribosomal RNA: A novel function for small nucleolar RNAs. *Cell* **1996**, *85*, 1077–1088.
6. Cavaille, J.; Bachellerie, J.P. SnoRNA-guided ribose methylation of rRNA: Structural features of the guide RNA duplex influencing the extent of the reaction. *Nucleic Acids Res.* **1998**, *26*, 1576–1587.
7. Bachellerie, J.P.; Michot, B.; Nicoloso, M.; Balakin, A.; Ni, J.; Fournier, M.J. Antisense snoRNAs: A family of nucleolar RNAs with long complementarities to rRNA. *Trends Biochem. Sci.* **1995**, *20*, 261–264.
8. Kiss, T. Small nucleolar RNA-guided posttranscriptional modification of cellular RNAs. *EMBO J.* **2001**, *20*, 3617–3622.
9. Weinstein, L.B.; Steitz, J.A. Guided tours: From precursor snoRNA to functional snoRNP. *Curr. Opin. Cell Biol.* **1999**, *11*, 378–384.
10. Meyuhas, O.; Avni, D.; Shama, S. Translational Control of Ribosomal Protein mRNAs in Eukaryotes. In *Translational Control*; Hershey, J.W.B., Mathews, M.B., Sonenberg, N., Eds.; Cold Spring Harbor Laboratory Press: New York, NY, USA, 1996; Volume 30, pp. 363–388.
11. Smith, C.M.; Steitz, J.A. Classification of *gas5* as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. *Mol. Cell Biol.* **1998**, *18*, 6897–6909.
12. Decatur, W.A.; Fournier, M.J. rRNA modifications and ribosome function. *Trends Biochem. Sci.* **2002**, *27*, 344–351.
13. Yusupov, M.M.; Yusupova, G.Z.; Baucom, A.; Lieberman, K.; Earnest, T.N.; Cate, J.H.; Noller, H.F. Crystal structure of the ribosome at 5.5 Å resolution. *Science* **2001**, *292*, 883–896.

14. Lane, B.G.; Ofengand, J.; Gray, M.W. Pseudouridine and *O*2'-methylated nucleosides. Significance of their selective occurrence in rRNA domains that function in ribosome-catalyzed synthesis of the peptide bonds in proteins. *Biochimie* **1995**, *77*, 7–15.
15. Green, R.; Noller, H.F. Reconstitution of functional 50S ribosomes from *in vitro* transcripts of *Bacillus stearothermophilus* 23S rRNA. *Biochemistry* **1999**, *38*, 1772–1779.
16. Khaitovich, P.; Tenson, T.; Kloss, P.; Mankin, A.S. Reconstitution of functionally active *Thermus aquaticus* large ribosomal subunits with *in vitro*-transcribed rRNA. *Biochemistry* **1999**, *38*, 1780–1788.
17. Liu, B.; Liang, X.H.; Piekna-Przybylska, D.; Liu, Q.; Fournier, M.J. Mis-targeted methylation in rRNA can severely impair ribosome synthesis and activity. *RNA Biol.* **2008**, *5*, 249–254.
18. Montanaro, L.; Calienni, M.; Bertoni, S.; Rocchi, L.; Sansone, P.; Storci, G.; Santini, D.; Ceccarelli, C.; Taffurelli, M.; Carnicelli, D.; *et al*. Novel dyskerin-mediated mechanism of p53 inactivation through defective mRNA translation. *Cancer Res.* **2010**, *70*, 4767–4777.
19. Belin, S.; Beghin, A.; Solano-Gonzàlez, E.; Bezin, L.; Brunet-Manquat, S.; Textoris, J.; Prats, A.C.; Mertani, H.C.; Dumontet, C.; Diaz, J.J. Dysregulation of ribosome biogenesis and translational capacity is associated with tumor progression of human breast cancer cells. *PLoS One* **2009**, *4*, doi:10.1371/journal.pone.0007147.
20. Tanaka, R.; Satoh, H.; Moriyama, M.; Satoh, K.; Morishita, Y.; Yoshida, S.; Watanabe, T.; Nakamura, Y.; Mori, S. Intronic U50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t(3;6)(q27; q15) of human B-cell lymphoma. *Genes Cells* **2000**, *5*, 277–287.
21. Dong, X.Y.; Rodriguez, C.; Guo, P.; Sun, X.; Talbot, J.T.; Zhou, W.; Petros, J.; Li, Q.; Vessella, R.L.; Kibel, A.S.; *et al*. SnoRNA U50 is a candidate tumor-suppressor gene at 6q14.3 with a mutation associated with clinically significant prostate cancer. *Hum. Mol. Genet*. **2008**, *17*, 1031–1042.
22. Mourtada-Maarabouni, M.; Pickard, M.R.; Hedge, V.L.; Farzaneh, F.; Williams, G.T. GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene* **2009**, *28*, 195–208.
23. Dong, X.Y.; Guo, P.; Boyd, J.; Sun, X.; Li, Q.; Zhou, W.; Dong, J.T. Implication of snoRNA U50 in human breast cancer. *J. Genet .Genomics.* **2009**, *36*, 447–454.
24. Nakamura, Y.; Takahashi, N.; Kakegawa, E.; Yoshida, K.; Ito, Y.; Kayano, H.; Niitsu, N.; Jinnai, I.; Bessho, M. The GAS5 (growth arrest-specific transcript 5) gene fuses to BCL6 as a result of t(1;3)(q25;q27) in a patient with B-cell lymphoma. *Cancer Genet. Cytogenet.* **2008**, *182*, 144–149.
25. Basu, A.; Das, P.; Chaudhuri, S.; Bevilacqua, E.; Andrews, J.; Barik, S.; Hatzoglou, M.; Komar, A.A.; Mazumder, B. Requirement of rRNA methylation for 80S ribosome assembly on a cohort of cellular internal ribosome Entry Sites. *Mol. Cell Biol.* **2011**, *31*, 4482–4499.

26. Nicoloso, M.; Qu, L.H.; Michot, B.; Bachellerie, J.P. Intron-encoded, antisense small nucleolar RNAs: The characterization of nine novel species points to their direct role as guides for the 2'-*O*-ribose methylation of rRNAs. *J. Mol. Biol.* **1996**, *260*, 178–195.
27. Maxwell, E.S.; Fournier, M.J. The small nucleolar RNAs. *Annu. Rev. Biochem.* **1995**, *64*, 897–934.
28. Maquat, L.E. When cells stop making sense: Effects of nonsense codons on RNA metabolism in vertebrate cells. *RNA* **1995**, *1*, 453–465.
29. Kino, T.; Hurt, D.E.; Ichijo, T.; Nader, N.; Chrousos, G.P. Noncoding RNA Gas5 is a growth arrest and starvation-associated repressor of the glucocorticoid receptor. *Sci. Signal*. **2010**, *3*, doi:10.1126/scisignal.2000568.
30. Ploner, A.; Ploner, C.; Lukasser, M.; Niederegger, H.; Hüttenhofer, A. Methodological obstacles in knocking down small noncoding RNAs. *RNA* **2009**, *15*, 1797–1804.
31. Yoon, A.; Peng, G.; Brandenburger, Y.; Zollo, O.; Xu, W.; Rego, E.; Ruggero, D. Impaired control of IRES-mediated translation in X-linked dyskeratosis congenita. *Science* **2006**, *312*, 902–906.
32. Bochkov, Y.A.; Palmenberg, A.C. Translational efficiency of EMCV IRES in bicistronic vectors is dependent upon IRES sequence and gene location. *Biotechniques* **2006**, *41*, 283–284.
33. Van Eden, M.E.; Byrd, M.P.; Sherrill, K.W.; Lloyd, R.E. Demonstrating internal ribosome entry sites in eukaryotic mRNAs using stringent RNA test procedures. *RNA* **2004**, *10*, 720–730.
34. Thoma, C.; Fraterman, S.; Gentzel, M.; Wilm, M.; Hentze, M.W. Translation initiation by the c-Myc mRNA internal ribosome entry sequence and the poly(A) tail. *RNA* **2008**, *14*, 1579–1589.

Reprinted from *IJMS*. Cite as: Hecker, M.; Thamilarasan, M.; Koczan, D.; Schröder, I.; Flechtner, K.; Freiesleben, S.; Füllen, G.; Thiesen, H.-J.; Zettl, U.K. MicroRNA Expression Changes during Interferon-Beta Treatment in the Peripheral Blood of Multiple Sclerosis Patients. *Int. J. Mol. Sci.* **2013**, *14*, 16087-16110.

Article

MicroRNA Expression Changes during Interferon-Beta Treatment in the Peripheral Blood of Multiple Sclerosis Patients

Michael Hecker [1,2,†,*], Madhan Thamilarasan [2,†], Dirk Koczan [3], Ina Schröder [2], Kristin Flechtner [3], Sherry Freiesleben [4], Georg Füllen [4], Hans-Jürgen Thiesen [3] and Uwe Klaus Zettl [2]

[1] Steinbeis Transfer Center for Proteome Analysis, Schillingallee 68, 18057 Rostock, Germany

[2] Department of Neurology, Division of Neuroimmunology, University of Rostock, Gehlsheimer Str. 20, 18147 Rostock, Germany; E-Mails: madhangp@gmail.com (M.T.); ina.schroeder@med.uni-rostock.de (I.S.); uwe.zettl@med.uni-rostock.de (U.K.Z.)

[3] Institute of Immunology, University of Rostock, Schillingallee 70, 18055 Rostock, Germany; E-Mails: dirk.koczan@med.uni-rostock.de (D.K.); kristin.flechtner@googlemail.com (K.F.); hj.thiesen@gmx.de (H.-J.T.)

[4] Institute for Biostatistics and Informatics in Medicine and Ageing Research, University of Rostock, Ernst-Heydemann-Str. 8, 18057 Rostock, Germany; E-Mails: sherry.freiesleben@uni-rostock.de (S.F.); fuellen@uni-rostock.de (G.F.)

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: michael.hecker@rocketmail.com; Tel.: +49-381-494-5891; Fax: +49-381-494-5882.

Received: 8 June 2013; in revised form: 12 July 2013 / Accepted: 26 July 2013 / Published: 5 August 2013

Abstract: MicroRNAs (miRNAs) are small non-coding RNA molecules acting as post-transcriptional regulators of gene expression. They are involved in many biological processes, and their dysregulation is implicated in various diseases, including multiple sclerosis (MS). Interferon-beta (IFN-beta) is widely used as a first-line immunomodulatory treatment of MS patients. Here, we present the first longitudinal study on the miRNA expression changes in response to IFN-beta therapy. Peripheral blood mononuclear cells

(PBMC) were obtained before treatment initiation as well as after two days, four days, and one month, from patients with clinically isolated syndrome (CIS) and patients with relapsing-remitting MS (RRMS). We measured the expression of 651 mature miRNAs and about 19,000 mRNAs in parallel using real-time PCR arrays and Affymetrix microarrays. We observed that the up-regulation of IFN-beta-responsive genes is accompanied by a down-regulation of several miRNAs, including members of the mir-29 family. These differentially expressed miRNAs were found to be associated with apoptotic processes and IFN feedback loops. A network of miRNA-mRNA target interactions was constructed by integrating the information from different databases. Our results suggest that miRNA-mediated regulation plays an important role in the mechanisms of action of IFN-beta, not only in the treatment of MS but also in normal immune responses. miRNA expression levels in the blood may serve as a biomarker of the biological effects of IFN-beta therapy that may predict individual disease activity and progression.

Keywords: interferon-beta; multiple sclerosis; peripheral blood; microRNA; gene expression

1. Introduction

Multiple sclerosis (MS) is a chronic disease of the central nervous system (CNS), which is characterized by multiple discrete areas of inflammatory demyelination, axonal degeneration, and glial scarring. The resulting loss of neurons and axons leads to diverse neurological symptoms, progressive disability, and a significant decrease in quality of life. The disease usually begins in early adulthood, and is more common in females. Different types of MS are distinguished: In about 85% of patients, the disease starts with a single demyelinating episode (clinically isolated syndrome, CIS) and progresses to a relapsing-remitting course (RRMS) with acute exacerbations and periods of remission [1–4].

A number of disease-modifying therapies for MS are available, and they are especially effective when applied in the early stages of the disease [4,5]. Injections of recombinant interferon-beta (IFN-beta) are considered a first-line option in the treatment of RRMS. IFN-beta has been shown to reduce the number of relapses and to suppress the accumulation of new inflammatory lesions in the brain. Three different preparations of IFN-beta are in clinical use. They differ in dose, route, and frequency of IFN-beta administration, but they are comparable regarding clinical efficacy [6].

IFN-beta has broad effects on the gene regulation of blood cells [7–11]. This has been shown by several studies that used microarray technology to analyze the gene expression dynamics in the peripheral blood of MS patients in response to IFN-beta therapy. In this way, more than a hundred genes have consistently been found differentially expressed during treatment [10]. The transcript levels of most of these genes are up-regulated within a few hours after IFN-beta injection, and they return to pre-treatment levels after a few days [12,13]. These IFN-beta-responsive genes are believed to

mediate the beneficial effects of the treatment through immunomodulatory, antiproliferative, and antipathogenic processes [7–9].

While the therapeutic effects on the regulation of mRNAs have been extensively investigated, studies on the regulation of microRNAs (miRNAs) are lacking. miRNAs are a distinct class of small (~22 nt) non-coding RNA molecules [14]. They originate from precursor RNAs (pre-miRNAs) found in longer primary transcripts (pri-miRNAs), which often also contain the exons of an mRNA. Mature miRNAs act as post-transcriptional regulators. They repress gene expression via base-pairing with complementary sequences within the 3' untranslated regions (UTRs) of target mRNAs. This interaction results in gene silencing by translational repression or target degradation. A miRNA can have hundreds of different mRNA targets, and a target might be regulated by a combination of multiple miRNAs [15]. The human genome encodes over 1000 miRNAs [16,17]. miRNAs are thus likely involved in most biological processes, and they play essential roles in the immune system and in the correct function of the CNS [18,19].

Dysregulated expression of miRNAs is associated with pathological conditions, including neurological diseases. Human MS studies showed altered miRNA expression in peripheral blood samples, lymphocyte subpopulations, and active CNS lesions from MS patients [20–22]. Studies with the animal model of MS, experimental autoimmune encephalomyelitis (EAE), also support the involvement of miRNAs in this disease [23,24]. These findings provided important insights into the pathophysiology of MS and opened a new avenue in biomarker research. If miRNA levels in the blood or brain of MS patients correlate with disease stage and progression of disability, they may also support early diagnosis and effective treatment in future [25].

In order to better understand the molecular mechanisms of action of IFN-beta therapy, it is important to investigate the miRNA expression dynamics during therapy. The regulation of miRNAs may contribute to the immunomodulatory and clinical effects of the treatment. Moreover, miRNAs might be markers for characterizing the biological response to IFN-beta. miRNA biomarkers for treatment monitoring could be useful in the individual management of disease activity. However, so far, there is only one study on the expression of miRNAs in MS patients during IFN-beta therapy: Waschbisch *et al.* obtained peripheral blood mononuclear cells (PBMC) from patients with RRMS, and analyzed the expression of five selected miRNAs by real-time PCR [26]. They compared the miRNA levels between treatment-naive patients (n = 36), IFN-beta-treated patients (n = 18), and patients treated with glatiramer acetate (GA, n = 20). As a result, none of the five miRNAs was differentially expressed in IFN-beta-treated patients, but *hsa-miR-146a-5p* and *hsa-miR-142-3p* were expressed at significantly lower levels in GA-treated patients [26]. Other researchers used microarrays to study the expression of hundreds of miRNAs in IFN-stimulated cells. In this way, O'Connell *et al.* observed that *hsa-miR-155-5p* is induced in primary murine macrophages after exposure to IFN-beta for 6 h [27]. Pedersen *et al.* studied the regulation of miRNAs in Huh7 cells and primary hepatocytes, which were stimulated with different concentrations of IFN-beta for up to 48 h [28]. They observed increased and reduced miRNA expression in response to IFN-beta, and showed that some of the

IFN-beta-induced miRNAs mediate antiviral effects against hepatitis C virus. This provides an example of miRNAs as components of the innate immune response.

In this study, we used microarrays to investigate in parallel the expression dynamics of mRNAs and miRNAs in PBMC of patients with CIS or RRMS in response to therapy with subcutaneous (sc.) IFN-beta. The blood samples were obtained longitudinally from six patients at four time points in the early phase of therapy, namely before the first (baseline), second, and third IFN-beta injection as well as after one month of treatment. We then screened for significant changes in miRNA and mRNA expression, and identified several miRNAs as differentially expressed during therapy. Information of different databases was then integrated [22,29] to examine whether the expression of these miRNAs is cell type-specific and correlates with the levels of their target mRNAs. Predicted and experimentally verified miRNA-mRNA interactions were compiled to construct a network of IFN-beta-responsive genes and miRNAs. To our knowledge, this is the first genome-wide miRNA profiling study on the *in vivo* effects of IFN-beta treatment in MS.

2. Results and Discussion

2.1. Study Population

Six female patients of Western European descent, and diagnosed with CIS (n = 2) or RRMS (n = 4), were recruited for this study (Pat1-6, mean age 37.5 years, Table 1). The patients were treatment-naive and started an immunomodulatory therapy with IFN-beta-1b (Betaferon, Bayer HealthCare) administered subcutaneously every other day. In the first weeks, the Betaferon titration pack was used, hence the patients started with a low dose (62.5 μg for the first three injections) that was gradually increased to the full dose (250 μg) after three weeks. All patients were continuously treated with IFN-beta-1b for at least one year. During follow-up, they were monitored for relapses and rated using the Expanded Disability Status Scale (EDSS). The individual disease activity during therapy was relatively low: Four of the patients (Pat1-4) were relapse-free and showed no disability progression within the first year of treatment (Table 1). The two patients with CIS (Pat1 and Pat5) did not convert to clinically definite MS in this period of time.

Table 1. Clinical data and demographic data of the patients.

Patient	Type	Age	Disease duration	EDSS (baseline)	EDSS (1 year)	Relapses (1 year)
Pat1	CIS	28	1	1.0	1.0	0
Pat2	RRMS	38	2	1.5	1.5	0
Pat3	RRMS	22	1	1.5	1.0	0
Pat4	RRMS	50	12	2.5	2.5	0
Pat5	CIS	60	2	1.5	2.5	0
Pat6	RRMS	27	2	2.0	1.0	2

Six female patients were recruited for the main cohort of this study. They were diagnosed with relapsing-remitting MS (RRMS) or clinically isolated syndrome (CIS) suggestive of MS. The age at study onset (in years) and the duration from the diagnosis to the start of IFN-beta-1b sc. therapy (in months) are shown. Additionally, the EDSS scores before treatment initiation (baseline) and after one year, as well as the number of relapses during the first year of clinical follow-up are given in the table.

Note that the patient group included only women. A differential hormonal regulation of immune system genes in blood cells has been described for different phases of the menstrual cycle [30]. Such differences in gene expression may have led to increased variance in the data. However, prior mRNA profiling studies observed no significant gender-specific differences in the gene expression signature in response to IFN-beta therapy [8,31], and this seems to be the case regarding the expression of miRNAs as well (see Section 2.5).

2.2. Parallel Measurement of mRNAs and MicroRNAs in Blood Cells

Patient blood samples were drawn immediately before first IFN-beta injection as well as two days, four days, and one month post therapy initiation. Total RNA of Ficoll-isolated PBMC from each sample was extracted to measure the levels of mRNAs and miRNAs with different platforms. We used TaqMan Array Human MicroRNA cards (Applied Biosystems, Foster City, CA, USA) to quantify the expression of 651 mature miRNAs and Affymetrix HG-U133 Plus 2.0 microarrays (Affymetrix, Santa Clara, CA, USA) to quantify the expression of about 19,000 mRNAs. In this way, we obtained in parallel the mRNA and miRNA expression profiles from six patients (Pat1-6) within the first month of IFN-beta treatment.

The data were preprocessed as described in Section 3.6. Relatively low variation in the transcriptome profiles indicated high data quality, comparable to our previous microarray time course data sets [8,31,32]. In the miRNA data, systematic and stochastic variation was higher. For the TaqMan miRNA B-cards of Pat5, the raw threshold cycle (Ct) values were generally higher (Supplementary File 1) due to an unknown measurement bias. In the PBMC samples of the other five patients, approximately 400 miRNAs could be detected (Ct < 38) with the TaqMan miRNA arrays (Table 2). The raw TaqMan data were transformed to the linear scale, and coefficients of variation (*CV*) were calculated to assess the effects of data normalization. The normalization decreased the average *CV* over all 768 measured assays from 0.953 to 0.894. The *CV* for the housekeeping miRNA *hsa-miR-191-5p* [33] was 0.183. In comparison, assays for the non-coding RNAs *U6*, *U44* and *U48* had *CV*s of >0.35. The *CV* for the housekeeping mRNA GAPDH was 0.071.

Table 2. Numbers of microRNAs detected in the samples with the TaqMan cards set.

Patient	Baseline	~48 h	~96 h	1 month
Pat1	380	374	364	373
Pat2	375	407	392	362
Pat3	427	400	408	407
Pat4	390	387	388	389
Pat5	285	258	266	262
Pat6	431	451	446	391

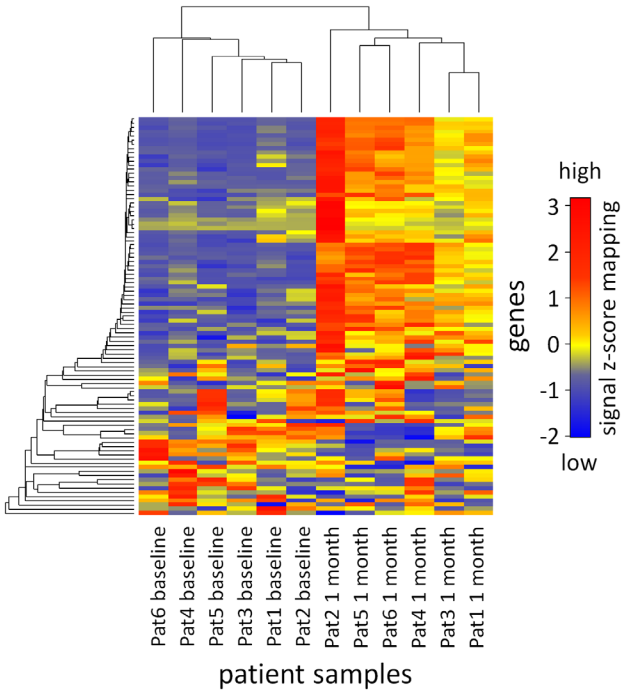
Approximately 400 miRNAs were found to be expressed (raw Ct value < 38) per sample. The numbers are lower for the PBMC samples of patient Pat5, where generally lower miRNA amounts were measured with the B-cards for all four time points (Supplementary File 1).

2.3. Analysis of mRNA Expression Dynamics

We filtered for mRNAs showing strong expression changes in response to IFN-beta therapy by comparing the baseline expression levels with the expression levels at the three time points during treatment. The MAID filtering method [34] was used to analyze the mRNA dynamics. As a result, 14, 34, and 66 genes were found to be expressed at higher or lower levels after two days, four days, and one month, respectively. In total, 95 genes were identified as up-regulated ($n = 75$) or down-regulated ($n = 20$) in the early course of the therapy (Supplementary File 2). The gene expression changes in the first month of subcutaneous IFN-beta-1b treatment are visualized in the heat map in Figure 1.

A permutation test (see Section 3.7) revealed that the number of 95 differentially expressed genes is significantly higher than would be expected by chance. In randomly permuted data sets, 29.6 genes on average were filtered. The number of filtered genes was below 95 in 98.8% of the permutations, which implies an empirical p-value of <0.05, demonstrating that most of the mRNA expression changes that we found are indeed due to the therapy.

Figure 1. Heat map visualization of the mRNA expression changes in response to IFN-beta. Shown are the baseline and one month transcript levels of the 95 genes that were identified as differentially expressed during IFN-beta therapy. The patient samples are represented in the columns, the genes are represented in the rows, and the gene expression levels were centered and scaled in row direction (z-scores). The clustering analysis clearly separated the PBMC samples obtained at baseline and after one month of therapy. The row labels of the heat map (*i.e.*, the respective genes) are given in Supplementary File 2. The upper half of the heat map contains most of the IFN-beta-induced genes.

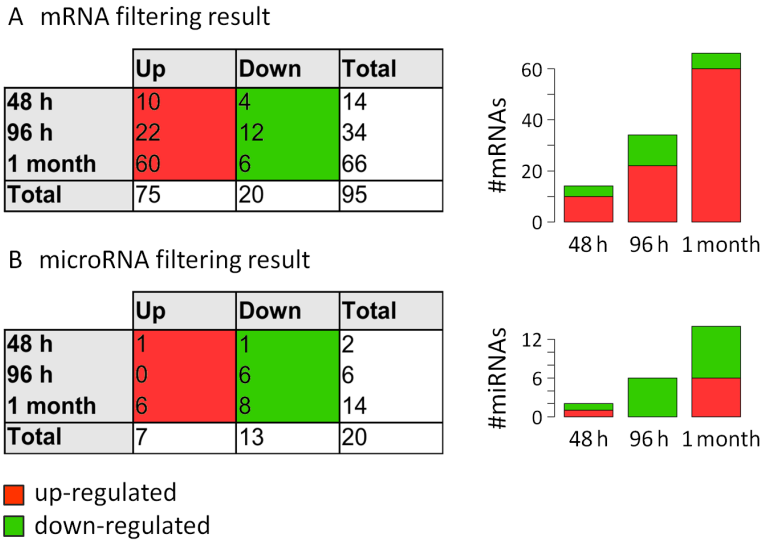


Despite the relatively small patient cohort ($n = 6$), the mRNA results were quite consistent with the literature. In another study, we already analyzed the PBMC gene expression profiles of a larger group of MS patients ($n = 25$) treated with IFN-beta-1b sc. [31]. In total, 63 of the 95 differentially expressed genes were also filtered in the previous study. Recently, we completed a similar microarray study on the effects of IFN-beta-1a sc. [8], in which 49 of the 95 genes were already identified as transcriptionally modulated (Supplementary File 2). Moreover, more up-regulated than down-regulated genes were filtered, which also confirms previous findings [10]. After the first IFN-beta injections, fewer genes were altered in expression than after one month (*cf.* Goertsches *et al.*, 2010 [31]). This can be explained by the fact that the patients started the first week of the therapy with a quarter of the full dose (Betaferon titration pack). Most of the filtered genes are part of an up-regulated type I IFN signature. For instance, *IFI6*, *IFI44L*, and *SIGLEC1* are known type I IFN-induced genes, which were up-regulated at all time points during therapy in comparison to baseline. In contrast, FCER1A was consistently down-regulated in response to therapy as has been described previously as well [8,31,32]. Regarding the functions of these genes, the reader is referred to the literature [7-9].

2.4. Analysis of MicroRNA Expression Dynamics

As for the mRNA data, the MAID filtering method [34] was used to identify miRNAs that are differentially expressed in PBMC within the first month of IFN-beta treatment. When we compared the expression levels at the three time points during therapy with the expression levels at baseline, 20 different miRNAs were filtered. Of these, seven miRNAs appeared as up-regulated and 13 miRNAs appeared as down-regulated in response to the therapy (Figure 2, Table 3 and Supplementary File 3). According to the permutation test, this is significantly more than expected by chance: In only 0.6% of the randomly permuted data sets, 20 (or more) miRNAs were filtered, and only 7.7 miRNAs were filtered on average.

Figure 2. Summary of the filtering of IFN-beta-responsive mRNAs and microRNAs. PBMC expression levels during therapy were compared to pre-treatment levels. The number (#) of differentially expressed mRNAs and miRNAs is depicted in the tables and bar plots. The row "Total" gives the union set over all three time point comparisons. (**A**) In the mRNA data, 95 genes were found to be modulated in expression in response to IFN-beta-1b treatment. As expected, most of them were up-regulated (n = 75) and known type I IFN-induced genes, and the strongest changes were observed at one month versus baseline; (**B**) In the miRNA data, the filtering method identified more down-regulated than up-regulated miRNAs during therapy, again with the strongest effects seen after one month.



Two of the 20 miRNAs (*hsa-miR-149-5p* and *hsa-miR-708-5p*) were filtered at two different time points. For the remaining miRNAs, the expression changes were not very stable in the course of therapy. This may be due to the small number of patients and the fact that the accuracy of miRNA measurements is in general limited. Therefore, our list of 20 miRNAs represents candidates that have to be validated in a larger patient cohort using, e.g., single real-time PCR assays.

Most of the miRNAs (n = 14) were filtered as up-regulated or down-regulated one month after IFN-beta-1b sc. treatment initiation. This corresponds to the results of the gene expression profiling, where the strongest changes in mRNA levels were also observed after one month (Figure 2). Previous studies demonstrated that the majority of IFN-beta-responsive genes can be seen at this time point [8,31]. Therefore, we hypothesize that, similarly, the number of miRNAs that are modulated in expression is not much higher after long-term treatment. Instead, the development of neutralizing antibodies (NAb) to IFN-beta might impair the biological response to the drug in some patients [35]. However, further studies are needed to investigate the long-term regulation of miRNAs and the potential effects of NAb.

Table 3. Details of microRNAs differentially expressed during IFN-beta therapy.

Mature miRNA	Sequence	Expression change	Family	pre-miRNA	pri-miRNA	Location
hsa-let-7a-5p	UGAGGUAGUAGGUUGUAUAGUU	up-regulated	let-7	MIRLET7A1		chr9 (q22.32)
				MIRLET7A2	MIR100HG	chr11 (q24.1)
				MIRLET7A3	MIRLET7BHG	chr22 (q13.31)
hsa-let-7b-5p	UGAGGUAGUAGGUUGUGUGGUU	up-regulated	let-7	MIRLET7B	MIRLET7BHG	chr22 (q13.31)
hsa-miR-16-5p	UAGCAGCACGUAAAUAUUGGCG	up-regulated	mir-15	MIR16-1	DLEU2	chr13 (q14.2)
				MIR16-2	SMC4	chr3 (q25.33)
hsa-miR-27a-5p	AGGGCUUAGCUGCUUGUGAGCA	down-regulated	mir-27	MIR27A		chr19 (p13.13)
hsa-miR-29a-3p	UAGCACCAUCUGAAAUCGGUUA	down-regulated	mir-29	MIR29A		chr7 (q32.3)
hsa-miR-29b-1-5p	GCUGGUUUCAUAUGGUGGUUUAGA	down-regulated	mir-29	MIR29B1		chr7 (q32.3)
hsa-miR-29c-3p	UAGCACCAUUUGAAAUCGGUUA	down-regulated	mir-29	MIR29C		chr1 (q32.2)
hsa-miR-95	UUCAACGGGUAUUUAUUGAGCA	down-regulated	mir-95	MIR95	ABLIM2	chr4 (p16.1)
hsa-miR-149-5p	UCUGGCUCCGUGUCUUCACUCCC	down-regulated	mir-149	MIR149	GPC1	chr2 (q37.3)
hsa-miR-181c-3p	AACCAUCGACCGUUGAGUGGAC	down-regulated	mir-181	MIR181C		chr19 (p13.13)
hsa-miR-193a-3p	AACUGGCCUACAAAGUCCCAGU	down-regulated	mir-193	MIR193A		chr17 (q11.2)
hsa-miR-193a-5p	UGGGUCUUUGCGGGCGAGAUGA	down-regulated	mir-193	MIR193A		chr17 (q11.2)
hsa-miR-342-5p	AGGGGUGCUAUCUGUGAUUGA	up-regulated	mir-342	MIR342	EVL	chr14 (q32.2)
hsa-miR-346	UGUCUGCCCGCAUGCCUGCCUCU	up-regulated	mir-346	MIR346	GRID1	chr10 (q23.2)
hsa-miR-423-5p	UGAGGGGCAGAGAGCGAGACUUU	down-regulated	mir-423	MIR423	NSRP1	chr17 (q11.2)
hsa-miR-518b	CAAAGCGCUCCCCUUUAGAGGU	up-regulated	mir-515	MIR518B		chr19 (q13.42)
hsa-miR-532-5p	CAUGCCUUGAGUGUAGGACCGU	down-regulated	mir-188	MIR532	CLCN5	chrX (p11.23)
hsa-miR-708-5p	AAGGAGCUUACAAUCUAGCUGGG	down-regulated	mir-708	MIR708	TENM4	chr11 (q14.1)
hsa-miR-760	CGGCUCUGGGUCUGUGGGGA	up-regulated	mir-760	MIR760		chr1 (p22.1)
hsa-miR-874	CUGCCCUGGCCCGAGGGACCGA	down-regulated	mir-874	MIR874	KLHL3	chr5 (q31.2)

The table lists the 20 miRNAs found to be expressed at higher or lower levels in the PBMC of patients with CIS or MS in response to IFN-beta therapy. The base sequence, the gene regulatory effect of the treatment ("Expression change"), the miRNA family, the HGNC symbols of the precursor and primary miRNAs, as well as the genomic location are shown. Two of the mature miRNAs (*hsa-let-7a-5p* and *hsa-miR-16-5p*) are processed from more than one precursor miRNA. For 11 miRNAs the pri-miRNA transcript has been annotated. None of these pri-miRNAs appeared in the mRNA filtering result. Precursor miRNAs of *hsa-let-7a-5p* and *hsa-let-7b-5p*, and of *hsa-miR-29a-3p* and *hsa-miR-29b-1-5p* are clustered, *i.e.*, they share their transcription locus.

Apparently, there are more down-regulated than up-regulated miRNAs during therapy, which is the opposite of the mRNA results. This suggests that the induction of IFN-beta-responsive genes is paralleled by a preferential down-regulation of miRNAs, which is plausible given that miRNAs act as gene silencers. Therefore, we analyzed whether the miRNAs indeed participate in the regulation of the mRNA transcripts (see Section 2.7).

2.5. Validation of IFN-beta-Induced MicroRNA Expression Changes

We used Affymetrix miRNA microarrays to replicate the miRNA measurements of the PBMC samples from three patients (Pat1-3) before the start of IFN-beta therapy as well as after one month (see Section 3.4). These microarrays had a lower measurement range than the TaqMan miRNA arrays, though the comparability of the data was acceptable (Spearman's rho = 0.823). The statistical analysis of this additional data set is limited by the small number of patients. However, when we compared the mean miRNA expression levels before and one month after treatment initiation, 13 of the 20 filtered miRNAs showed the same trend of up-regulation or down-regulation (Supplementary File 3). At the significance threshold alpha = 0.10, four miRNAs (*hsa-miR-29a-3p*, *hsa-miR-29c-3p*, *hsa-miR-193a-3p*, and *hsa-miR-532-5p*) were confirmed to be down-regulated during treatment.

For further validation, we selected five of the 20 filtered miRNAs to quantify their expression in PBMC of an independent cohort of 12 patients using TaqMan single-tube assays (Supplementary File 4 and Supplementary File 5). These 12 patients (8 RRMS/4 CIS, 7 females/5 males, mean age 36.2 years) also started a therapy with IFN-beta-1b sc. The PBMC were obtained again in a longitudinal manner before the first drug injection and after one month of treatment. In this data set, *hsa-miR-29a-3p* and *hsa-miR-29c-3p* could be confirmed as differentially expressed in response to IFN-beta therapy (*t*-test *p*-values < 0.001). *hsa-miR-29c-3p* was expressed at lower levels during therapy in comparison to pre-treatment levels in all 12 patients, hence independent of disease stage (CIS or RRMS), age, and gender (Figure 3). Additionally, *hsa-miR-532-5p* was confirmed to be down-regulated (*p*-value = 0.048). *hsa-miR-16-5p* and *hsa-miR-149-5p* showed the same trend of expression change as in the other data sets (TaqMan miRNA arrays and Affymetrix miRNA arrays), but this was not statistically significant (*p*-values > 0.10, Supplementary File 3).

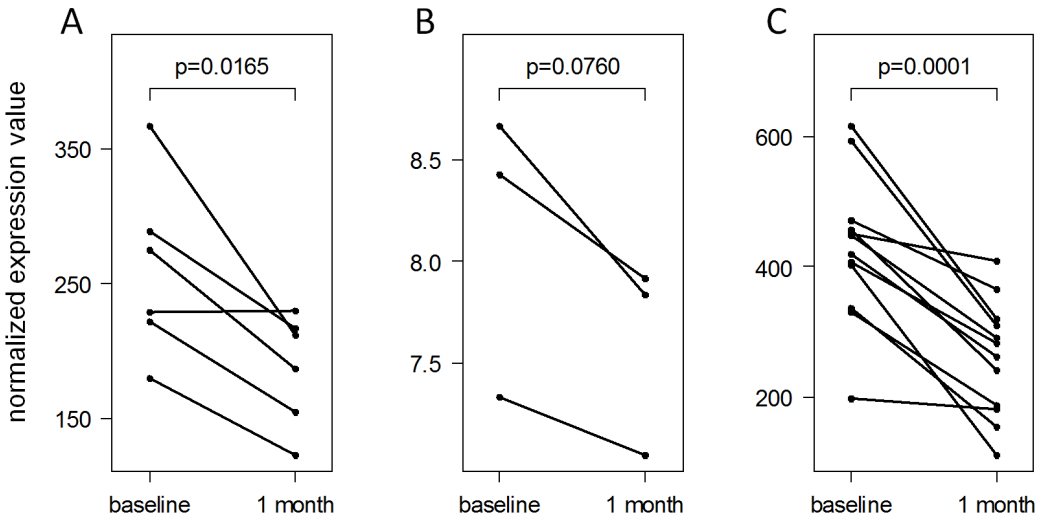
2.6. Functions of the MicroRNAs in the Context of Multiple Sclerosis

The filtered miRNAs (Table 3) affect diverse cellular functions and pathways, and some of them have been implicated in MS. In our data set, *hsa-let-7a-5p* and *hsa-let-7b-5p*, which belong to the let-7 family, were expressed at higher levels during therapy. Lehmann *et al.* showed that let-7 family members can activate TLR7 signaling in macrophages and microglia, thereby inducing neurodegeneration [36]. In CD4+ T cells, let-7 miRNAs reduce the expression of IL10, a cytokine with anti-inflammatory properties [37]. The up-regulation of *hsa-let-7b-5p* by IFN-beta has already been demonstrated *in vitro* in primary macrophages [38]. Interestingly, *hsa-let-7b-5p* is in turn capable of binding the endogenous IFN-beta transcript, forming a negative feedback loop for the regulation of IFN-beta protein [38]. Recently, Gandhi *et al.* found altered *hsa-let-7a-5p* levels in the blood plasma of patients with the secondary-progressive form of MS, but not of patients with RRMS [39].

The observed up-regulation of *hsa-miR-16-5p* in response to IFN-beta therapy may restore the aberrant expression of this miRNA in the disease. Two studies observed reduced levels of *hsa-miR-16-5p* in the blood of RRMS patients. One study compared the expression in PBMC and

CD4+ T cells from untreated patients and healthy donors [40]. The other study showed that *hsa-miR-16-5p* is down-regulated in B cells as well [41].

Figure 3. Down-regulation of *hsa-miR-29c-3p* in response to IFN-beta therapy. The *hsa-miR-29c-3p* expression dynamics within the first month of IFN-beta treatment are presented. (**A**) TaqMan miRNA cards revealed reduced levels of this miRNA in the PBMC of 6 patients (Pat1-6, the main cohort); (**B**) Affymetrix miRNA arrays were then used to replicate the measurement for three of these patients (Pat1-3); (**C**) Finally, the down-regulation of *hsa-miR-29c-3p* was confirmed in an independent group of 12 patients (the validation cohort) using TaqMan single-tube assays. The Affymetrix analysis was based on hybridization of miRNA molecules to probes (probe set "*hsa-miR-29c_st*"), whereas the TaqMan analyses were based on real-time PCR. The TaqMan data are in linear scale, and the Affymetrix data are in log2 scale due to a different data preprocessing.



Three miRNAs of the mir-29 family were down-regulated one month after the start of therapy, and this was confirmed for two of them in an independent validation cohort of 12 patients (Figure 3). *hsa-miR-29a-3p* and *hsa-miR-29c-3p* were both expressed at relatively high levels (Supplementary File 3), and their mature sequences differ only in one base (Table 3). Therefore, it is likely that they play similar roles as post-transcriptional regulators. *hsa-miR-29a-3p* and *hsa-miR-29b-1-5p* belong to the same genomic cluster. Smith *et al.* demonstrated that IFN-gamma enhances the transcription of this *miR-29ab1* cluster, which acts in a negative feedback loop by regulating TBET and IFN-gamma [42]. Additionally, they showed decreased *hsa-miR-29b-3p* levels upon T cell activation in MS patients. This suggests a dysregulation of the feedback loop, which may bias T helper type 1 cell differentiation and may contribute to chronic inflammation [42]. Other studies also provided evidence that the members of the mir-29 family control innate and adaptive immune responses by targeting IFN-gamma [43,44]. The therapeutic down-modulation of mir-29 miRNAs might be mediated by NF-kappaB. The activation of NF-kappaB signaling, via ligation of Toll-like receptors, was shown to inhibit *miR-29ab1*

expression [45]. Functionally, mir-29 promotes apoptosis, whereas repression of mir-29 levels is protective [45]. *hsa-miR-29a-3p* has been further shown to regulate myelin gene expression by Schwann cells [46].

hsa-miR-181c-3p was filtered as down-regulated during IFN-beta therapy. Several studies described the other strand of its pre-miRNA to be dysregulated in MS. Lower levels of *hsa-miR-181c-5p* were measured in PBMC [47] and in MS lesions [48] of patients in comparison to controls. On the other hand, *hsa-miR-181c-5p* seems to be overabundant in the cerebrospinal fluid of patients with MS [49]. However, the biological processes that are influenced by the *hsa-mir-181c* miRNAs and their role in MS therapy remain largely unknown.

The expression of the mir-193 family members *hsa-miR-193a-3p* and *hsa-miR-193a-5p* was repressed during the therapy (Table 3). A study by Lindberg *et al.* demonstrated increased expression of *hsa-miR-193a-5p* in CD4+ T cells of RRMS patients compared to healthy subjects [50]. Otaegui *et al.* confirmed the potential relevance of this miRNA duplex in MS. Based on a co-expression network analysis, they postulated that *hsa-miR-193a-3p* is related to the remission stage of MS [51]. Moreover, the precursor molecule *hsa-mir-193a* was found to modulate apoptotic processes by promoting CASP3 activation induced by TNFSF10 signaling [52]. TNFSF10 (=TRAIL), in turn, is a known IFN-beta-induced gene and was transcriptionally up-regulated in the patients' PBMC (Supplementary File 2). The concomitant regulation of mir-193 miRNAs may thus contribute to the molecular mechanisms of action of IFN-beta.

Another miRNA caught our attention: *hsa-miR-223-3p* appeared to be the highest expressed miRNA in most PBMC samples (mean raw Ct value = 13.6). In a microarray study by Keller *et al.*, elevated levels of this miRNA were measured in the peripheral blood of RRMS patients as compared with healthy controls [53]. A significantly increased expression of *hsa-miR-223-3p* was later confirmed in PBMC from RRMS patients using real-time PCR [54]. Functionally, *hsa-miR-223-3p* modulates inflammatory responses by modulating the NF-kappaB pathway [55].

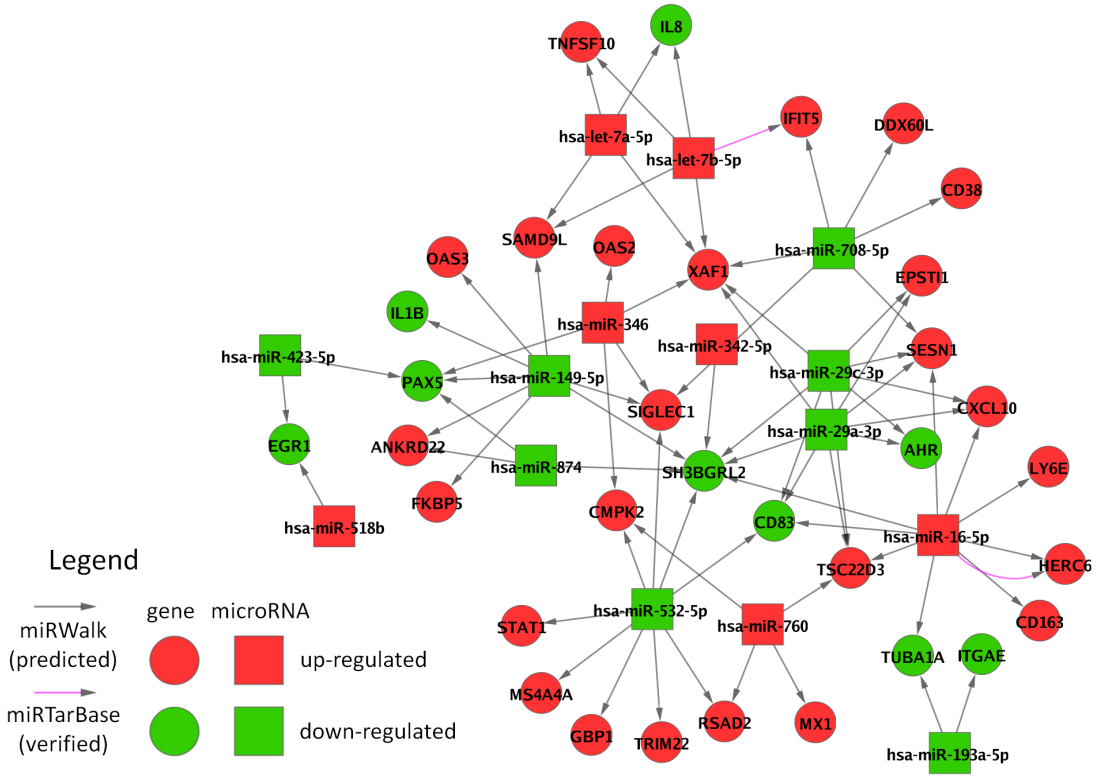
The remaining miRNAs identified in our study have so far not been mentioned in the context of MS, and their functions are poorly understood. miRNAs that have been repeatedly described to be differentially expressed in MS, e.g., *hsa-miR-142-3p*, *hsa-miR-146a-5p*, *hsa-miR-155-5p* and *hsa-miR-326* [22], were not contained in the filtering result, thus the therapy did not normalize their abnormal expression (*cf.* Waschbisch *et al.* [26]). Further studies are needed to decipher the immunological pathways involved, and to better understand the role of miRNA-dependent regulatory mechanisms in the immunopathogenesis of MS.

2.7. Interactions between Filtered MicroRNAs and mRNAs

Interactions between the 20 filtered miRNAs and the 95 filtered mRNAs were derived from two databases providing potential target genes of miRNAs. The miRWalk database [56] was used to obtain miRNA-mRNA interactions consistently predicted by multiple computational algorithms. This resulted in 74 potential interactions for 15 of the 20 filtered miRNAs. The miRTarBase database [57]

was used to extract interactions with experimental evidence in the literature. This resulted in two verified interactions: *hsa-miR-16-5p* was identified as a post-transcriptional regulator of HERC6 in both databases. Moreover, there is an experimentally determined interaction between *hsa-let-7b* and IFIT5. IFIT5 is a known IFN-beta-induced gene. A recent study by Abbas *et al.* characterized IFIT5 as an innate immune effector molecule acting as a sensor of viral single-stranded RNAs [58]. This confers antiviral defense by inhibiting viral replication. However, IFIT5 also recognizes cellular RNAs, including tRNAs [59]. The full network of miRNA-mRNA target interactions is visualized in Figure 4. The most connected regulators in the network were *hsa-miR-16-5p* and *hsa-miR-532-5p* with nine IFN-beta-responsive target genes each. *hsa-miR-29a-3p* and *hsa-miR-29c-3p*, which are closely related to each other, had eight predicted gene targets in common due to their similar mature sequences.

Figure 4. Verified and predicted interactions between IFN-beta-responsive microRNAs and mRNAs. The network of miRNA-target interactions between differentially expressed miRNAs and mRNAs was built using the databases miRWalk and miRTarBase. miRWalk reported results of 10 different prediction algorithms, and we only considered mRNA targets being computationally predicted by at least five of the 10 algorithms. miRTarBase provided experimentally validated miRNA-target interactions. In total, 15 of the filtered miRNAs were linked to 34 of the filtered genes by 74 predicted and two validated interactions. *hsa-miR-29a-3p* and *hsa-miR-29c-3p* (in the center-right) belong to the same miRNA family and are predicted to regulate eight target genes in common. The network is available as a Cytoscape session file (Supplementary File 6).



Down-regulated SH3BGRL2 and up-regulated XAF1 are simultaneously targeted by several ($n > 5$) miRNAs (miRNA target hubs [60]). Both genes have already been identified as differentially expressed in our previous studies on the effects of IFN-beta therapy [8,31] (Supplementary File 2). XAF1 is a critical mediator of IFN-beta-induced apoptosis. Its expression correlates with the cellular sensitivity to the pro-apoptotic actions of TNFSF10 [7,61]. However, while this supports the notion that miRNAs contribute to the mechanisms of action of IFN-beta, it should be noted that the majority of interactions in the network is predicted. For instance, the miRNA-mRNA interaction between *hsa-miR-16-5p* and SESN1 is predicted by nine out of 10 algorithms implemented in the miRWalk database, but has not yet been experimentally demonstrated. Another limitation is that we did not analyze whether the miRNAs bind their target mRNAs at multiple sites. Such a cooperative regulation through repetitive elements in the 3' UTR can increase repression efficacy [62]. Additional studies are needed to validate the miRNA targets, e.g., by luciferase assays.

Opposing effects exist in the network since several mRNAs are targeted by down-regulated and up-regulated miRNAs (Figure 4). Moreover, the network does not include the effects of the transcription factors (TF), which are activated through the IFN-beta signaling pathway and which are known to regulate the expression of most of the genes [31,63]. Therefore, it is difficult to disentangle the effects of miRNA expression changes on the mRNA levels measured during therapy. Further studies are thus required, e.g., specific miRNA transfection experiments examining the impact on potential target genes at both the mRNA and protein level.

Our network analysis was limited to the set of filtered genes, and the many other potential target genes of the filtered miRNAs were out of scope. Moreover, we did not investigate whether the miRNAs target viral RNAs, which is also worth to be explored in detail [28]. Despite these limitations, we conclude that miRNA-mediated regulation plays an important role in the pleiotropic effects of IFN-beta in normal immune responses as well as in the treatment of MS.

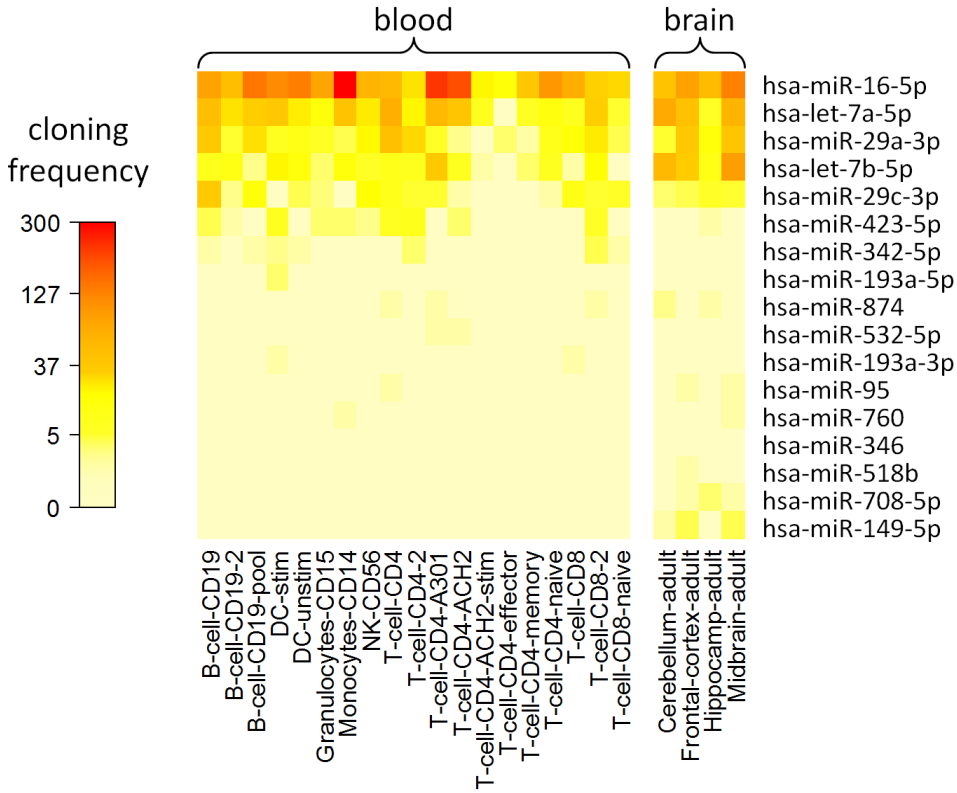
2.8. Cell Type-Specific Expression of IFN-beta-Responsive MicroRNAs

We analyzed whether the miRNA expression changes during therapy affect the gene expression in different cell populations involved in the disease. This was done by comparing the cell type-specific expression of the miRNAs using information from the smirnaDB database [64] (Figure 5). Of the filtered miRNAs, *hsa-miR-16-5p* is highly expressed among diverse blood cells, in particular monocytes and CD4+ T cell lines, whereas *hsa-miR-149-5p* is only detected in brain (Figure 5). Accordingly, in our time course data set, *hsa-miR-16-5p* was expressed at very high levels, and *hsa-miR-149-5p* was expressed at relatively low levels (Supplementary File 3). Several miRNAs, however, were not detected in some cell populations (clone count = 0) due to the limited sensitivity of the measurement technique used to generate these data [65]. More sensitive methods could be used to further analyze the cell type-specific roles of selected miRNAs in MS and therapy, e.g., *hsa-miR-29a-3p* and *hsa-miR-29c-3p*, which were expressed in B cells.

2.9. Final Remarks and Perspectives

The cellular regulation of the miRNAs is still not well understood. Many type I IFN-responsive genes harbor in their promoter region a specific sequence motif, the IFN-stimulated response element (ISRE), which is bound by IFN-activated TFs [7]. However, when we searched miRGen 2.0, a database of TF binding sites for miRNA transcripts [66], there was only one miRNA (*hsa-mir-203a*) that was predicted to have an ISRE located near the transcription start site. It is conceivable that some of the filtered miRNAs are regulated at the RNA processing level rather than at the transcriptional level. Several post-transcriptional mechanisms can affect mature miRNA biogenesis and stability [67]. Recent studies provided evidence that miRNAs can be suppressed by circular RNAs, which act as natural miRNA sponges [68]. On the other hand, the localization of the miRNAs might be altered, e.g., by microparticle shedding [69]. Additional research is needed to elucidate how the activity of miRNAs might be modulated during therapy.

Figure 5. Expression of 17 filtered miRNAs in different cell populations. The heat map visualizes the levels of IFN-beta-responsive miRNAs in 19 blood cell populations and four brain tissues. The data were downloaded from the smirnaDB database [64], which did not contain three of the 20 filtered miRNAs (*hsa-miR-27a-5p*, *hsa-miR-29b-1-5p* and *hsa-miR-181c-3p*). *hsa-let-7a-5p* and *hsa-miR-16-5p* are highly expressed in peripheral blood and brain. *hsa-let-7b-5p* and *hsa-miR-149-5p* are preferentially expressed in brain tissues. Several of the miRNAs (e.g., *hsa-miR-346*) were not detected in certain cell types [65].



Currently, there is a clear lack of studies investigating the changes in miRNA expression during MS therapy. Apart from the study by Waschbisch *et al.* [26] (see Section 1), there is only another study by Sievers *et al.*, who found differentially expressed miRNAs in B cells of patients treated with natalizumab [41]. Future miRNA profiling analyses should use a longitudinal design and address both the short-term and the long-term effects of the available treatments. Such studies may also help to understand why some patients continue to have clinical relapses, disability progression or active lesions despite therapy. Defining the individual response to treatment is difficult, but miRNAs may have the potential to be used as prognostic biomarkers, thereby facilitating improved patient care. The identification of miRNA biomarkers should be supported by functional studies on how miRNAs affect complex biological processes by targeting multiple genes in different cell types.

3. Experimental Section

3.1. Samples

Fifteen milliliters peripheral venous EDTA blood samples were taken from six patients (Table 1) immediately before first (baseline), second, and third IFN-beta injection as well as after one month. For validation, further blood samples were taken from an independent cohort of 12 patients (Supplementary File 4) before and one month after the start of IFN-beta-1b sc. therapy. The samples were always processed within one hour. PBMC were separated by isopycnic centrifugation in Ficoll density gradients, and total RNA enriched with small RNAs was isolated using the mirVana miRNA isolation kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturers' protocols. The study was approved by the University of Rostock's ethics committee and carried out according to the Declaration of Helsinki. Written informed consent was obtained from all patients before study onset.

3.2. Gene Expression Profiling Using Microarrays

To quantify the mRNA levels, total RNA from each of the 24 PBMC samples of the main study cohort was labeled and hybridized to Affymetrix microarrays. Biotinylated cRNA were prepared according to the standard Affymetrix 3' IVT protocol from 200 ng total RNA (Expression Analysis Technical Manual; Affymetrix, Santa Clara, CA, USA). Following fragmentation, 15 µg of cRNA were hybridized for 16 h at 45 °C on Affymetrix GeneChip Human Genome U133 Plus 2.0 Arrays. The microarrays were washed and stained in the Affymetrix Fluidics Station 450, and scanned using the Affymetrix GeneChip Scanner 3000 (Affymetrix, Santa Clara, CA, USA).

3.3. MicroRNA Expression Analysis Using Real-Time PCR

To quantify the miRNA expression levels, we used the TaqMan Array Human MicroRNA A + B Cards Set v2.0 (Applied Biosystems, Foster City, CA, USA), which consists of two 384-well plates with TaqMan assay reagents. These plates contain 720 assays to measure 651 different human

miRNAs. Moreover, there are 30 assays for positive controls, 2 assays for negative controls, and 16 assays were discarded as they link to dead miRNA entries in the miRBase database (release 17) [17]. Total RNA (120 ng) from each sample (n = 24) was reverse transcribed to cDNA using Megaplex RT Primers in combination with Megaplex PreAmp Primers (Life Technologies, Carlsbad, CA, USA). The real-time PCR measurements were then performed with predesigned primers and TaqMan probes with 45 cycles according to the manufacturer's instructions in a 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA, USA). Raw Ct values were computed using the SDS 2.3 and RQ Manager 1.2 software (Applied Biosystems, Foster City, CA, USA), and undetermined data were set to Ct = 45.

3.4. Validation MicroRNA Analysis Using Microarrays

For validation, we replicated the miRNA expression measurements of the PBMC from three patients (Pat1-3) at two time points (before the first IFN-beta injection as well as after one month). Total RNA of these six samples was labeled and hybridized to Affymetrix GeneChip miRNA 2.0 arrays. Biotinylated RNA was prepared using the FlashTag Biotin HSR RNA labeling kit according to the standard Affymetrix protocol from 600 ng total RNA (Expression Analysis Technical Manual; Affymetrix). Following fragmentation, the biotin-labeled RNA was hybridized for 16 h at 45 °C on Affymetrix miRNA 2.0 arrays. The microarrays were washed and stained in the Affymetrix Fluidics Station 450, and scanned using the Affymetrix GeneChip Scanner 3000.

3.5. Validation MicroRNA Analysis Using Real-Time PCR

To verify the results in an independent cohort of patients, we measured 5 of the 20 filtered miRNAs in PBMC samples from 12 additional patients (Supplementary File 5). The blood samples were obtained before and after one month of IFN-beta-1b treatment. The miRNAs were selected based on a combination of different criteria, e.g., change of expression after one month according to both the TaqMan miRNA array data and the Affymetrix miRNA microarray data. The validation experiment was performed using TaqMan single-tube assays for *hsa-miR-16-5p* (Assay ID 000391), *hsa-miR-29a-3p* (Assay ID 002112), *hsa-miR-29c-3p* (Assay ID 000587), *hsa-miR-149-5p* (Assay ID 002255), and *hsa-miR-532-5p* (Assay ID 001518). Additionally, the housekeeping miRNA *hsa-miR-191-5p* (Assay ID 002299) was measured for normalization [33]. For each assay, 10 ng of total RNA from each sample (n = 24) were used to convert an individual miRNA to cDNA using an RT primer specific for the miRNA of interest (Applied Biosystems). The real-time PCR quantitation was performed in triplicates with predesigned primers and TaqMan probes according to the TaqMan MicroRNA Assay protocol with 45 cycles in a 7900HT Sequence Detection System (Applied Biosystems). An equivalent of 0.5 ng total RNA was used to obtain a single data point. Raw Ct values were computed using the SDS 2.3 and RQ Manager 1.2 software (Applied Biosystems).

3.6. Expression Data Preprocessing

In the case of the Affymetrix U133 Plus 2.0 gene expression microarrays, the raw probe-level signals were converted to expression values (signal intensities) using the MAS5.0 algorithm and custom GeneAnnot-based chip definition files (version 2.2.0) [70]. Data normalization was performed by loess normalization using the R package "affy". Each Affymetrix GeneChip yielded mRNA levels of 19,204 human genes.

In case of the TaqMan miRNA cards, we first set the detection limit at Ct = 38 [71], and converted the raw Ct values to the linear scale using the equation $2^{-Ct} \times 10^{-9}$. After this step, a Ct value of 38 corresponds to an expression signal of 0.004, and a Ct value of 20 corresponds to an expression signal of 953. Systematic differences in the time course data were then corrected by loess normalization. This was done separately for each patient ($n = 6$) and each card (A and B). Finally, to reduce variation in the expression signals between the patients, we scaled the data of card A and of card B so that the respective 95% quantile was the same for each patient.

In case of the Affymetrix miRNA microarrays, the raw signals were converted to expression values using the RMA algorithm with quantile normalization. For each chip, this resulted in log-transformed expression levels for 20,706 probe sets interrogating small non-coding RNA transcripts. For each of the 20 filtered miRNAs (Table 3), there was one designated probe set.

In case of the TaqMan single-tube real-time PCR miRNA assays, we set the detection limit at Ct = 38 [71], calculated the mean Ct value of each triplicate, converted the raw Ct values to the linear scale using the equation 2^{-Ct}, normalized the results to the expression values of the housekeeping miRNA *hsa-miR-191-5p* [33], and scaled these ratios by a factor of 1000 for convenience.

The non-normalized and the normalized expression data of the six patients receiving IFN-beta therapy are available in the Gene Expression Omnibus (GEO) database via the SuperSeries record GSE46293. This GEO entry links to all data from the Affymetrix U133 Plus 2.0 microarrays, the TaqMan MicroRNA Cards Sets v2.0, and the Affymetrix miRNA 2.0 microarrays.

3.7. Filtering of Differentially Expressed mRNAs and MicroRNAs

We filtered the Affymetrix U133 Plus 2.0 and TaqMan miRNA cards data sets for IFN-beta-responsive genes and miRNAs by comparing the PBMC expression levels immediately before treatment initiation (baseline) with the expression levels two days, four days, and one month after the start of IFN-beta therapy. Up-regulation and down-regulation of genes and miRNAs were quantified using signal intensity-dependent fold-changes (MAID-scores) as described in our previous studies [8,34] (see also http://www.hki-jena.de/index.php/0/2/490). MAID-scores represent adjusted fold-changes, where a higher fold-change (*i.e.*, relative change in expression) is required for genes expressed at low levels than for genes expressed at high levels. This is realized by an exponential function (MAID regression curve) that isfitted to the signal intensity -dependent variation in the data. For each time point comparison and each type of array, we computed the MAID-score for all patients

and for all measured genes and miRNAs. We then selected the genes and miRNAs being up-regulated (MAID-score above the cutoff C) or being down-regulated (MAID-score $< -C$) in at least four of the six patients. We chose $C = 2$ for the gene expression data set, and $C = 1$ for the miRNA expression data set. For the latter, a lower MAID-score cutoff was chosen because the larger variation in the miRNA data already leads to a higher MAID regression curve.

To provide an estimate of the number of genes and miRNAs passing the MAID filtering by chance, a permutation test was performed. The data sets were permuted 1000 times by randomly rearranging the temporal sequence of the data of each patient. The same filtering criteria as described above were then applied to each permutation.

As an alternative filtering criterion, we statistically compared the PBMC expression levels of mRNAs and miRNAs before and during treatment using paired *t*-tests (Supplementary File 2 and Supplementary File 3). However, considering the relatively small number of patients in our study, the MAID filtering method is thought to be more robust to the variation in the data.

3.8. Visualization of the mRNA Expression Data

A heat map was used to visualize the expression changes of the 95 filtered mRNAs (Figure 1). The heat map displays the respective data of all 6 patients before the start of therapy and after one month. To reorder the rows and columns of the data matrix, hierarchical clustering was performed with the single linkage method and Pearson's correlation coefficient as a measure of similarity. For visualization purposes, the expression values (signal intensities) were centered and scaled row-wise (resulting in z-scores) with the standard R heat map function.

3.9. Evaluation of the MicroRNA Validation Data

Spearman's rank correlation coefficient (rho) was calculated to evaluate whether the TaqMan miRNA array data correlate with the Affymetrix miRNA microarray data (120 data pairs: 20 miRNAs, three patients, and two time points). Paired *t*-tests were computed for assessing the difference in expression after one month of IFN-beta-1b sc. therapy versus baseline in the preprocessed and normalized data of the Affymetrix microarrays (Pat1-3) and of the TaqMan single-tube assays ($n = 12$ additional patients).

3.10. Interaction Network Analysis

We studied the regulatory interactions between the miRNAs and their target mRNAs by integrating the information from two different databases. Experimentally verified and computationally predicted target genes of the 20 IFN-beta-responsive miRNAs were extracted from the databases miRTarBase (version 3.5) [57] and miRWalk (April 2013) [56], respectively. miRWalk contained predictions for all 20 filtered miRNAs, and miRTarBase contained interactions for 9 of the 20 miRNAs. For the prediction of targets with miRWalk, we applied the option of the web server to run the calculations with 10 different prediction algorithms on 3' UTRs of all human genes, and then gathered only the

miRNA-mRNA interactions that were predicted by at least 5 of the 10 algorithms. Finally, interactions being associated with the 95 filtered genes were visualized as a network using the Cytoscape software (version 2.8.0) [72].

3.11. MicroRNA Expression in Different Cell Populations

To investigate the expression of the filtered miRNAs in different peripheral blood cell types and brain regions, we used the smirnaDB database, which provides expression levels of 692 human miRNAs for 170 cell populations and tissues [64]. This miRNA expression atlas is based on sequence analysis of small RNA clone libraries [65]. The relative cloning frequencies of miRNAs represent a measure of miRNA expression. However, in this data set, many miRNAs were identified at very low clone counts (*cf.* Landgraf *et al.* [65]). The data for 19 blood cell types (including three CD4+ T cell lines) and four brain tissues were downloaded from smirnaDB and visualized as a heat map in the R software environment.

4. Conclusions

To our knowledge, this is the first longitudinal genome-wide study examining the *in vivo* effects of IFN-beta treatment on miRNA expression in blood cells of patients with CIS or RRMS. The strongest changes in mRNA and miRNA expression were detected one month after the start of IFN-beta-1b sc. treatment. We observed that the induction of IFN-beta-responsive genes is paralleled by a preferential down-regulation of miRNAs. This suggests that the regulation of miRNAs contributes to the molecular mechanisms of action of IFN-beta in protective immune responses as well as in MS therapy. We confirmed the down-regulation of *hsa-miR-29a-3p*, *hsa-miR-29c-3p*, and *hsa-miR-532-5p* in an independent cohort of patients. We further analyzed the interactions between differentially expressed miRNAs and mRNAs. The largest number of predicted interactions to IFN-responsive genes was found for *hsa-miR-532-5p* and *hsa-miR-16-5p*. Up-regulated *hsa-miR-16-5p* was expressed at very high levels in different cell types of the blood, in particular monocytes. However, unraveling the complex gene regulatory interactions between TFs, miRNAs and genes remains a big challenge for the future. Functionally, some of the 20 filtered miRNAs (e.g., members of the mir-29 family) are associated with apoptosis and are involved in IFN signaling feedback loops. miRNA expression profiles in blood cells may provide biomarkers for monitoring the biological response to therapy to predict individual disease activity and progression. They may also help to better understand the pathogenetic mechanisms and to optimize the treatment of MS. Our results provide a rationale for subsequent studies in larger MS cohorts.

Acknowledgments

We thank our colleagues Ildikó Tóth, Brigitte Paap, Gabriele Gillwaldt and Caroline Siebrecht for laboratory assistance and clinical documentation. We thank Charlotte Angerstein for helpful discussions

and comments on microRNA databases. We are grateful to study nurse Christa Tiffert for her excellent coordination of patient care and sample collection. We thank the United Europeans for the development of PHArmacogenomics in Multiple Sclerosis (UEPHA*MS) consortium for their funding.

The experiments were partially funded by Bayer HealthCare. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of Interest

U.K. Zettl received research support as well as speaking fees from Bayer HealthCare, Biogen Idec, Merck Serono, Novartis, Sanofi, Almirall and Teva. M. Hecker received speaking fees from Bayer HealthCare, Novartis and Teva. M. Thamilarasan, D. Koczan, I. Schröder, K. Flechtner, S. Freiesleben, G. Füllen and H.-J. Thiesen declare no potential conflicts of interest.

Supplementary Information

Supplementary File 1 (TIFF image): Distribution of the raw Ct values measured with the TaqMan miRNA arrays. The data of the A-cards (A) and the B-cards (B) are shown for all six patients and all four time points. The detection limit was set at Ct = 38.

Supplementary File 2 (XLS Excel spreadsheet): mRNA filtering result. This table provides the 95 genes that were up-regulated or down-regulated in the PBMC of the patients two days, four days or one month after the initiation of subcutaneous IFN-beta-1b therapy. Different types of information are given for each gene, e.g., gene symbol, official full name, Entrez Gene identifier, and the MAID filtering results.

Supplementary File 3 (XLS Excel spreadsheet): microRNA filtering result and validation. In comparison to pre-treatment levels, 20 miRNAs were found to be expressed at higher or lower levels in PBMC during IFN-beta treatment. The table provides, e.g., the TaqMan Detector identifiers, the MAID filtering outputs as well as the results from the validation analyses.

Supplementary File 4 (XLS Excel spreadsheet): Clinical data and demographic data of the patients in the validation cohort. Twelve patients were recruited to confirm the observed expression changes of five selected miRNAs within the first month of IFN-beta-1b sc. treatment.

Supplementary File 5 (XLS Excel spreadsheet): Validation real-time PCR data set. TaqMan single-tube assays were used to quantify the expression of five selected miRNAs before the start of IFN-beta therapy and after one month in PBMC of an independent cohort of 12 patients (Supplementary File 4). This table contains the raw Ct values of these five miRNAs and the housekeeping miRNA *hsa-miR-191-5p* [33]. The measurements were done in triplicates.

Supplementary File 6 (CYS Cytoscape session file): miRNA-mRNA interaction network. This Cytoscape file (http://www.cytoscape.org) [72] contains computationally predicted and experimentally determined interactions between IFN-beta-responsive miRNAs and mRNAs. The interactions were

obtained from the databases miRWalk [56] and miRTarBase [57]. A visualization of the network is shown in Figure 4.

References

1. Compston, A.; Coles, A. Multiple sclerosis. *Lancet* **2008**, *372*, 1502–1517.
2. Miller, D.H.; Chard, D.T.; Ciccarelli, O. Clinically isolated syndromes. *Lancet Neurol.* **2012**, *11*, 157–169.
3. Polman, C.H.; Reingold, S.C.; Banwell, B.; Clanet, M.; Cohen, J.A.; Filippi, M.; Fujihara, K.; Havrdova, E.; Hutchinson, M.; Kappos, L.; *et al.* Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann. Neurol.* **2011**, *69*, 292–302.
4. Vosoughi, R.; Freedman, M.S. Therapy of MS. *Clin. Neurol. Neurosurg.* **2010**, *112*, 365–385.
5. Buck, D.; Hemmer, B. Treatment of multiple sclerosis: Current concepts and future perspectives. *J. Neurol.* **2011**, *258*, 1747–1762.
6. Limmroth, V.; Putzki, N.; Kachuck, N.J. The interferon beta therapies for treatment of relapsing-remitting multiple sclerosis: Are they equally efficacious? A comparative review of open-label studies evaluating the efficacy, safety, or dosing of different interferon beta formulations alone or in combination. *Ther. Adv. Neurol. Disord.* **2011**, *4*, 281–296.
7. Borden, E.C.; Sen, G.C.; Uze, G.; Silverman, R.H.; Ransohoff, R.M.; Foster, G.R.; Stark, G.R. Interferons at age 50: Past, current and future impact on biomedicine. *Nat. Rev. Drug Discov.* **2007**, *6*, 975–990.
8. Hecker, M.; Hartmann, C.; Kandulski, O.; Paap, B.K.; Koczan, D.; Thiesen, H.J.; Zettl, U.K. Interferon-beta therapy in multiple sclerosis: The short-term and long-term effects on the patients' individual gene expression in peripheral blood. *Mol. Neurobiol.* **2013**, doi:10.1007/s12035-013-8463-1.
9. Dhib-Jalbut, S.; Marks, S. Interferon-beta mechanisms of action in multiple sclerosis. *Neurology* **2010**, *74*, S17–S24.
10. Goertsches, R.H.; Hecker, M.; Zettl, U.K. Monitoring of multiple sclerosis immunotherapy: From single candidates to biomarker networks. *J. Neurol.* **2008**, *255*, 48–57.
11. Goertsches, R.H.; Zettl, U.K.; Hecker, M. Sieving treatment biomarkers from blood gene-expression profiles: A pharmacogenomic update on two types of multiple sclerosis therapy. *Pharmacogenomics* **2011**, *12*, 423–432.
12. Reder, A.T.; Velichko, S.; Yamaguchi, K.D.; Hamamcioglu, K.; Ku, K.; Beekman, J.; Wagner, T.C.; Perez, H.D.; Salamon, H.; Croze, E. IFN-beta1b induces transient and variable gene expression in relapsing-remitting multiple sclerosis patients independent of neutralizing antibodies or changes in IFN receptor RNA expression. *J. Interferon Cytokine Res.* **2008**, *28*, 317–331.

13. Gilli, F.; Marnetto, F.; Caldano, M.; Sala, A.; Malucchi, S.; di Sapio, A.; Capobianco, M.; Bertolotto, A. Biological responsiveness to first injections of interferon-beta in patients with multiple sclerosis. *J. Neuroimmunol.* **2005**, *158*, 195–203.
14. Bartel, D.P. MicroRNAs: Target recognition and regulatory functions. *Cell* **2009**, *136*, 215–233.
15. Friedman, R.C.; Farh, K.K.; Burge, C.B.; Bartel, D.P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **2009**, *19*, 92–105.
16. Bentwich, I.; Avniel, A.; Karov, Y.; Aharonov, R.; Gilad, S.; Barad, O.; Barzilai, A.; Einat, P.; Einav, U.; Meiri, E.; *et al.* Identification of hundreds of conserved and nonconserved human microRNAs. *Nat. Genet.* **2005**, *37*, 766–770.
17. Kozomara, A.; Griffiths-Jones, S. miRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **2011**, *39*, D152–D157.
18. Xiao, C.; Rajewsky, K. MicroRNA control in the immune system: Basic principles. *Cell* **2009**, *136*, 26–36.
19. Im, H.I.; Kenny, P.J. MicroRNAs in neuronal function and dysfunction. *Trends Neurosci.* **2012**, *35*, 325–334.
20. Tufekci, K.U.; Oner, M.G.; Genc, S.; Genc, K. MicroRNAs and multiple sclerosis. *Autoimmune Dis.* **2010**, *2011*, 807426.
21. Guerau-de-Arellano, M.; Alder, H.; Ozer, H.G.; Lovett-Racke, A.; Racke, M.K. miRNA profiling for biomarker discovery in multiple sclerosis: From microarray to deep sequencing. *J. Neuroimmunol.* **2012**, *248*, 32–39.
22. Angerstein, C.; Hecker, M.; Paap, B.K.; Koczan, D.; Thamilarasan, M.; Thiesen, H.J.; Zettl, U.K. Integration of MicroRNA databases to study microRNAs associated with multiple sclerosis. *Mol. Neurobiol.* **2012**, *45*, 520–535.
23. Thamilarasan, M.; Koczan, D.; Hecker, M.; Paap, B.; Zettl, U.K. MicroRNAs in multiple sclerosis and experimental autoimmune encephalomyelitis. *Autoimmun. Rev.* **2012**, *11*, 174–179.
24. Mycko, M.P.; Cichalewska, M.; Machlanska, A.; Cwiklinska, H.; Mariasiewicz, M.; Selmaj, K.W. MicroRNA-301a regulation of a T-helper 17 immune response controls autoimmune demyelination. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E1248–E1257.
25. Paap, B.K.; Hecker, M.; Koczan, D.; Zettl, U.K. Molecular biomarkers in multiple sclerosis. *J. Clin. Cell. Immunol.* **2013**, *S10*, doi:10.4172/2155-9899.S10-009.
26. Waschbisch, A.; Atiya, M.; Linker, R.A.; Potapov, S.; Schwab, S.; Derfuss, T. Glatiramer acetate treatment normalizes deregulated microRNA expression in relapsing remitting multiple sclerosis. *PLoS One* **2011**, *6*, e24604.
27. O'Connell, R.M.; Taganov, K.D.; Boldin, M.P.; Cheng, G.; Baltimore, D. MicroRNA-155 is induced during the macrophage inflammatory response. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1604–1609.

28. Pedersen, I.M.; Cheng, G.; Wieland, S.; Volinia, S.; Croce, C.M.; Chisari, F.V.; David, M. Interferon modulation of cellular microRNAs as an antiviral mechanism. *Nature* **2007**, *449*, 919–922.
29. Schmitz, U.; Wolkenhauer, O. Web resources for microRNA research. *Adv. Exp. Med. Biol.* **2013**, *774*, 225–250.
30. Dosiou, C.; Lathi, R.B.; Tulac, S.; Huang, S.T.; Giudice, L.C. Interferon-related and other immune genes are downregulated in peripheral blood leukocytes in the luteal phase of the menstrual cycle. *J. Clin. Endocrinol. Metab.* **2004**, *89*, 2501–2504.
31. Goertsches, R.H.; Hecker, M.; Koczan, D.; Serrano-Fernandez, P.; Moeller, S.; Thiesen, H.J.; Zettl, U.K. Long-term genome-wide blood RNA expression profiles yield novel molecular response candidates for IFN-beta-1b treatment in relapsing remitting MS. *Pharmacogenomics* **2010**, *11*, 147–161.
32. Hecker, M.; Goertsches, R.H.; Fatum, C.; Koczan, D.; Thiesen, H.J.; Guthke, R.; Zettl, U.K. Network analysis of transcriptional regulation in response to intramuscular interferon-beta-1a multiple sclerosis treatment. *Pharmacogenomics J.* **2012**, *12*, 134–146.
33. Peltier, H.J.; Latham, G.J. Normalization of microRNA expression levels in quantitative RT-PCR assays: Identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA* **2008**, *14*, 844–852.
34. Hecker, M.; Goertsches, R.H.; Engelmann, R.; Thiesen, H.J.; Guthke, R. Integrative modeling of transcriptional regulation in response to antirheumatic therapy. *BMC Bioinforma.* **2009**, *10*, 262.
35. Killestein, J.; Polman, C.H. Determinants of interferon beta efficacy in patients with multiple sclerosis. *Nat. Rev. Neurol.* **2011**, *7*, 221–228.
36. Lehmann, S.M.; Krüger, C.; Park, B.; Derkow, K.; Rosenberger, K.; Baumgart, J.; Trimbuch, T.; Eom, G.; Hinz, M.; Kaul, D.; *et al.* An unconventional role for miRNA: Let-7 activates Toll-like receptor 7 and causes neurodegeneration. *Nat. Neurosci.* **2012**, *15*, 827–835.
37. Swaminathan, S.; Suzuki, K.; Seddiki, N.; Kaplan, W.; Cowley, M.J.; Hood, C.L.; Clancy, J.L.; Murray, D.D.; Méndez, C.; Gelgor, L.; *et al.* Differential regulation of the Let-7 family of microRNAs in CD4+ T cells alters IL-10 expression. *J. Immunol.* **2012**, *188*, 6238–6246.
38. Witwer, K.W.; Sisk, J.M.; Gama, L.; Clements, J.E. MicroRNA regulation of IFN-beta protein expression: Rapid and sensitive modulation of the innate immune response. *J. Immunol.* **2010**, *184*, 2369–2376.
39. Gandhi, R.; Healy, B.; Gholipour, T.; Egorova, S.; Musallam, A.; Shuja, M.; Nejad, P.; Patel, B.; Hei, H.; Khoury, S.; *et al.* Circulating microRNAs as biomarkers for disease staging in multiple sclerosis. *Ann. Neurol.* **2013**, *73*, 729–740.
40. Lorenzi, J.C.; Brum, D.G.; Zanette, D.L.; de Paula Alves Souza, A.; Barbuzano, F.G.; Dos Santos, A.C.; Barreira, A.A.; da Silva, W.A. *miR-15a* and *16-1* are downregulated in CD4+ T cells of multiple sclerosis relapsing patients. *Int. J. Neurosci.* **2012**, *122*, 466–471.

41. Sievers, C.; Meira, M.; Hoffmann, F.; Fontoura, P.; Kappos, L.; Lindberg, R.L. Altered microRNA expression in B lymphocytes in multiple sclerosis: Towards a better understanding of treatment effects. *Clin. Immunol.* **2012**, *144*, 70–79.
42. Smith, K.M.; Guerau-de-Arellano, M.; Costinean, S.; Williams, J.L.; Bottoni, A.; Mavrikis Cox, G.; Satoskar, A.R.; Croce, C.M.; Racke, M.K.; Lovett-Racke, A.E.; *et al.* miR-29ab1 deficiency identifies a negative feedback loop controlling Th1 bias that is dysregulated in multiple sclerosis. *J. Immunol.* **2012**, *189*, 1567–1576.
43. Ma, F.; Xu, S.; Liu, X.; Zhang, Q.; Xu, X.; Liu, M.; Hua, M.; Li, N.; Yao, H.; Cao, X. The microRNA *miR-29* controls innate and adaptive immune responses to intracellular bacterial infection by targeting interferon-gamma. *Nat. Immunol.* **2011**, *12*, 861–869.
44. Steiner, D.F.; Thomas, M.F.; Hu, J.K.; Yang, Z.; Babiarz, J.E.; Allen, C.D.; Matloubian, M.; Blelloch, R.; Ansel, K.M. *MicroRNA-29* regulates T-box transcription factors and interferon-gamma production in helper T cells. *Immunity* **2011**, *35*, 169–181.
45. Mott, J.L.; Kurita, S.; Cazanave, S.C.; Bronk, S.F.; Werneburg, N.W.; Fernandez-Zapico, M.E. Transcriptional suppression of *mir-29b-1/mir-29a* promoter by c-Myc, hedgehog, and NF-kappaB. *J. Cell Biochem.* **2010**, *110*, 1155–1164.
46. Verrier, J.D.; Lau, P.; Hudson, L.; Murashov, A.K.; Renne, R.; Notterpek, L. Peripheral myelin protein 22 is regulated post-transcriptionally by *miRNA-29a*. *Glia* **2009**, *57*, 1265–1279.
47. Martinelli-Boneschi, F.; Fenoglio, C.; Brambilla, P.; Sorosina, M.; Giacalone, G.; Esposito, F.; Serpente, M.; Cantoni, C.; Ridolfi, E.; Rodegher, M.; *et al.* MicroRNA and mRNA expression profile screening in multiple sclerosis patients to unravel novel pathogenic steps and identify potential biomarkers. *Neurosci. Lett.* **2012**, *508*, 4–8.
48. Junker, A.; Krumbholz, M.; Eisele, S.; Mohan, H.; Augstein, F.; Bittner, R.; Lassmann, H.; Wekerle, H.; Hohlfeld, R.; Meinl, E. MicroRNA profiling of multiple sclerosis lesions identifies modulators of the regulatory protein CD47. *Brain* **2009**, *132*, 3342–3352.
49. Haghikia, A.; Haghikia, A.; Hellwig, K.; Baraniskin, A.; Holzmann, A.; Décard, B.F.; Thum, T.; Gold, R. Regulated microRNAs in the CSF of patients with multiple sclerosis: A case-control study. *Neurology* **2012**, *79*, 2166–2170.
50. Lindberg, R.L.; Hoffmann, F.; Mehling, M.; Kuhle, J.; Kappos, L. Altered expression of *miR-17-5p* in CD4+ lymphocytes of relapsing-remitting multiple sclerosis patients. *Eur. J. Immunol.* **2010**, *40*, 888–898.
51. Otaegui, D.; Baranzini, S.E.; Armañanzas, R.; Calvo, B.; Muñoz-Culla, M.; Khankhanian, P.; Inza, I.; Lozano, J.A.; Castillo-Triviño, T.; Asensio, A.; *et al.* Differential micro RNA expression in PBMC from multiple sclerosis patients. *PLoS One* **2009**, *4*, e6309.
52. Ovcharenko, D.; Kelnar, K.; Johnson, C.; Leng, N.; Brown, D. Genome-scale microRNA and small interfering RNA screens identify small RNA modulators of TRAIL-induced apoptosis pathway. *Cancer Res.* **2007**, *67*, 10782–10788.

53. Keller, A.; Leidinger, P.; Lange, J.; Borries, A.; Schroers, H.; Scheffler, M.; Lenhof, H.P.; Ruprecht, K.; Meese, E. Multiple sclerosis: MicroRNA expression profiles accurately differentiate patients with relapsing-remitting disease from healthy controls. *PLoS One* **2009**, *4*, e7440.
54. Ridolfi, E.; Fenoglio, C.; Cantoni, C.; Calvi, A.; de Riz, M.; Pietroboni, A.; Villa, C.; Serpente, M.; Bonsi, R.; Vercellino, M.; *et al.* Expression and genetic analysis of microRNAs involved in multiple sclerosis. *Int. J. Mol. Sci.* **2013**, *14*, 4375–4384.
55. Li, T.; Morgan, M.J.; Choksi, S.; Zhang, Y.; Kim, Y.S.; Liu, Z.G. MicroRNAs modulate the noncanonical transcription factor NF-kappaB pathway by regulating expression of the kinase IKKalpha during macrophage differentiation. *Nat. Immunol.* **2010**, *11*, 799–805.
56. Dweep, H.; Sticht, C.; Pandey, P.; Gretz, N. miRWalk—Database: Prediction of possible miRNA binding sites by "walking" the genes of three genomes. *J. Biomed. Inform.* **2011**, *44*, 839–847.
57. Hsu, S.D.; Lin, F.M.; Wu, W.Y.; Liang, C.; Huang, W.C.; Chan, W.L.; Tsai, W.T.; Chen, G.Z.; Lee, C.J.; Chiu, C.M.; *et al.* miRTarBase: A database curates experimentally validated microRNA-target interactions. *Nucleic Acids Res.* **2011**, *39*, D163–D169.
58. Abbas, Y.M.; Pichlmair, A.; Górna, M.W.; Superti-Furga, G.; Nagar, B. Structural basis for viral 5'-PPP-RNA recognition by human IFIT proteins. *Nature* **2013**, *494*, 60–64.
59. Katibah, G.E.; Lee, H.J.; Huizar, J.P.; Vogan, J.M.; Alber, T.; Collins, K. tRNA binding, structure, and localization of the human interferon-induced protein IFIT5. *Mol. Cell* **2013**, *49*, 743–750.
60. Lai, X.; Schmitz, U.; Gupta, S.K.; Bhattacharya, A.; Kunz, M.; Wolkenhauer, O.; Vera, J. Computational analysis of target hub gene repression regulated by multiple and cooperative miRNAs. *Nucleic Acids Res.* **2012**, *40*, 8818–8834.
61. Leaman, D.W.; Chawla-Sarkar, M.; Vyas, K.; Reheman, M.; Tamai, K.; Toji, S.; Borden, E.C. Identification of X-linked inhibitor of apoptosis-associated factor-1 as an interferon-stimulated gene that augments TRAIL Apo2L-induced apoptosis. *J. Biol. Chem.* **2002**, *277*, 28504–28511.
62. Elefant, N.; Berger, A.; Shein, H.; Hofree, M.; Margalit, H.; Altuvia, Y. RepTar: A database of predicted cellular targets of host and viral miRNAs. *Nucleic Acids Res.* **2011**, *39*, D188–D194.
63. Hundeshagen, A.; Hecker, M.; Paap, B.K.; Angerstein, C.; Kandulski, O.; Fatum, C.; Hartmann, C.; Koczan, D.; Thiesen, H.J.; Zettl, U.K. Elevated type I interferon-like activity in a subset of multiple sclerosis patients: Molecular basis and clinical relevance. *J. Neuroinflammation* **2012**, *9*, 140.
64. Hausser, J.; Berninger, P.; Rodak, C.; Jantscher, Y.; Wirth, S.; Zavolan, M. MirZ: An integrated microRNA expression atlas and target prediction resource. *Nucleic Acids Res.* **2009**, *37*, W266–W272.
65. Landgraf, P.; Rusu, M.; Sheridan, R.; Sewer, A.; Iovino, N.; Aravin, A.; Pfeffer, S.; Rice, A.; Kamphorst, A.O.; Landthaler, M.; *et al.* A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **2007**, *129*, 1401–1414.
66. Alexiou, P.; Vergoulis, T.; Gleditzsch, M.; Prekas, G.; Dalamagas, T.; Megraw, M.; Grosse, I.; Sellis, T.; Hatzigeorgiou, A.G. miRGen 2.0: A database of microRNA genomic information and regulation. *Nucleic Acids Res.* **2010**, *38*, D137–D141.

67. Kai, Z.S.; Pasquinelli, A.E. MicroRNA assassins: Factors that regulate the disappearance of miRNAs. *Nat. Struct. Mol. Biol.* **2010**, *17*, 5–10.
68. Hentze, M.W.; Preiss, T. Circular RNAs: Splicing's enigma variations. *EMBO J.* **2013**, *32*, 923–925.
69. Valadi, H.; Ekström, K.; Bossios, A.; Sjöstrand, M.; Lee, J.J.; Lötvall, J.O. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **2007**, *9*, 654–659.
70. Ferrari, F.; Bortoluzzi, S.; Coppe, A.; Sirota, A.; Safran, M.; Shmoish, M.; Ferrari, S.; Lancet, D.; Danieli, G.A.; Bicciato, S. Novel definition files for human GeneChips based on GeneAnnot. *BMC Bioinf.* **2007**, *8*, 446.
71. Burns, M.; Valdivia, H. Modelling the limit of detection in real-time quantitative PCR. *Eur. Food Res. Technol.* **2008**, *226*, 1513–1524.
72. Smoot, M.E.; Ono, K.; Ruscheinski, J.; Wang, P.L.; Ideker, T. Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics* **2011**, *27*, 431–432.

Reprinted from *IJMS*. Cite as: Curtale, G.; Citarella, F. Dynamic Nature of Noncoding RNA Regulation of Adaptive Immune Response. *Int. J. Mol. Sci.* **2013**, *14*, 17347-17377.

Review

Dynamic Nature of Noncoding RNA Regulation of Adaptive Immune Response

Graziella Curtale [1,2,*] and Franca Citarella [3]

[1] Department of Medical Biotechnologies and Translational Medicine, University of Milan, Rozzano 20089, Italy

[2] Humanitas Clinical and Research Center, Rozzano 20089, Italy

[3] Department of Cellular Biotechnology and Hematology, Sapienza University of Rome, Rome 00185, Italy; E-Mail: citarella@bce.uniroma1.it

* Author to whom correspondence should be addressed; E-Mail: graziella.curtale@humanitasresearch.it; Tel.: +39-028-224-5146; Fax: +39-028-224-5088.

Received: 1 July 2013; in revised form: 30 July 2013 / Accepted: 12 August 2013 / Published: 22 August 2013

Abstract: Immune response plays a fundamental role in protecting the organism from infections; however, dysregulation often occurs and can be detrimental for the organism, leading to a variety of immune-mediated diseases. Recently our understanding of the molecular and cellular networks regulating the immune response, and, in particular, adaptive immunity, has improved dramatically. For many years, much of the focus has been on the study of protein regulators; nevertheless, recent evidence points to a fundamental role for specific classes of noncoding RNAs (ncRNAs) in regulating development, activation and homeostasis of the immune system. Although microRNAs (miRNAs) are the most comprehensive and well-studied, a number of reports suggest the exciting possibility that long ncRNAs (lncRNAs) could mediate host response and immune function. Finally, evidence is also accumulating that suggests a role for miRNAs and other small ncRNAs in autocrine, paracrine and exocrine signaling events, thus highlighting an elaborate network of regulatory interactions mediated by different classes of ncRNAs during immune response. This review will explore the multifaceted roles of ncRNAs in the adaptive immune response. In particular, we will focus on the well-established role of miRNAs and on the emerging role of lncRNAs and circulating ncRNAs, which all make

indispensable contributions to the understanding of the multilayered modulation of the adaptive immune response.

Keywords: noncoding RNAs; adaptive immunity; lymphocytes; epigenetic

1. Introduction

Adaptive immune response, generated by clonal selection of antigen-specific lymphocytes, provides an extremely versatile mechanism of host organism defense and increases its protection against subsequent exposure to the same antigen and/or reinfection with the same pathogen. Cells involved in the adaptive immune response are characterized by temporal development in primary lymphoid organs and functional differentiation in secondary lymphoid organs, phenomena whose successful regulation relies on a multilayered system of control [1,2]. In fact, it is now clear that the process of transcriptional regulation, carried out by general and specific transcription factors, must be framed in a picture in which the epigenetic control and the post-transcriptional control, which are also the result of non-protein coding RNA (ncRNA) activities, contribute in a fundamental manner. A growing number of observations indicate that different genome-encoded functional RNA species orchestrate the development and the differentiation state of complex tissues and specific cell types [3,4]. With regard to the adaptive immune response, while the regulative roles of several microRNAs (miRNAs) are well-established [5], recent, yet emerging, evidence demonstrates the relevance of a new heterogeneous repertoire of functional noncoding RNAs [6], which can be collectively referred to as long noncoding RNAs (lncRNAs).

The first and most generic association of miRNAs with the immune system was due to the detection of different miRNA expression profiles either in normal lymphocytes or lymphocytes derived from patients affected by immunological malignancies [7]. As a logical consequence of the identification of specific miRNA expression signatures, the next step has been the investigation of a possible causal relationship between miRNAs and the disease. Despite several *in vitro* studies identifying specific miRNA target genes involved in different aspects of cell biology, the first strong genetic evidence of the pivotal role of miRNAs in immune biology came from the analysis of mice knock-out for genes that are involved in the biogenesis of mature miRNAs, namely Dicer and Drosha. These studies established that miRNAs, as a whole, control the development and the homeostasis of the immune system [8–10]. Successively, it was established that the expression of several miRNAs is restricted to hematopoietic cells and is associated with specific developmental stages or functional differentiation of lymphocytes [11]. Moreover, it has been shown that miRNA concentration is dynamically modulated during the immune response, due to the existence of feed-back loops in which miRNAs control their own expression and, in so doing, contribute to the homeostasis of the immune system and

the maintenance of cell identity [12,13]. In this context, uncovering the regulation of miRNA expression in the immune system became a major issue.

Genes coding for miRNAs are mostly transcribed by RNA polymerase II and, therefore, are subjected to the regulation of transcription factors, whose repertoire is different in different cells and whose activation can be regulated by extracellular stimuli. This, then, makes way for the transcriptional regulation of miRNAs by co-activators and co-repressors, including chromatin remodeling factors and other epigenetic factors. In addition, recent studies indicate that the expression of mature miRNAs in the cell can be controlled at a post-transcriptional level through several mechanisms, such as processing, subcellular localization and decoy. The latter mechanisms call into question lncRNA functions.

Concerning lncRNAs, global genome profiling studies uncovered that the expression of many of them is lymphocyte-specific and change dynamically during T-cell differentiation, suggesting that, also, these functional RNAs may contribute to the development and homeostasis of the immune system [14,15]. This new evidence, together with the observation that cells are able to exchange information not only via hormones, but also via genetic material, is dictating the need to build a model of multilayered control of the development and the differentiation of the immune system cells, as well as of the regulation of the immune response. Perturbations in any component of this regulatory system will be associated with the onset of immune diseases.

This review will explore the proposed multifaceted roles of these classes of ncRNAs in adaptive immunity.

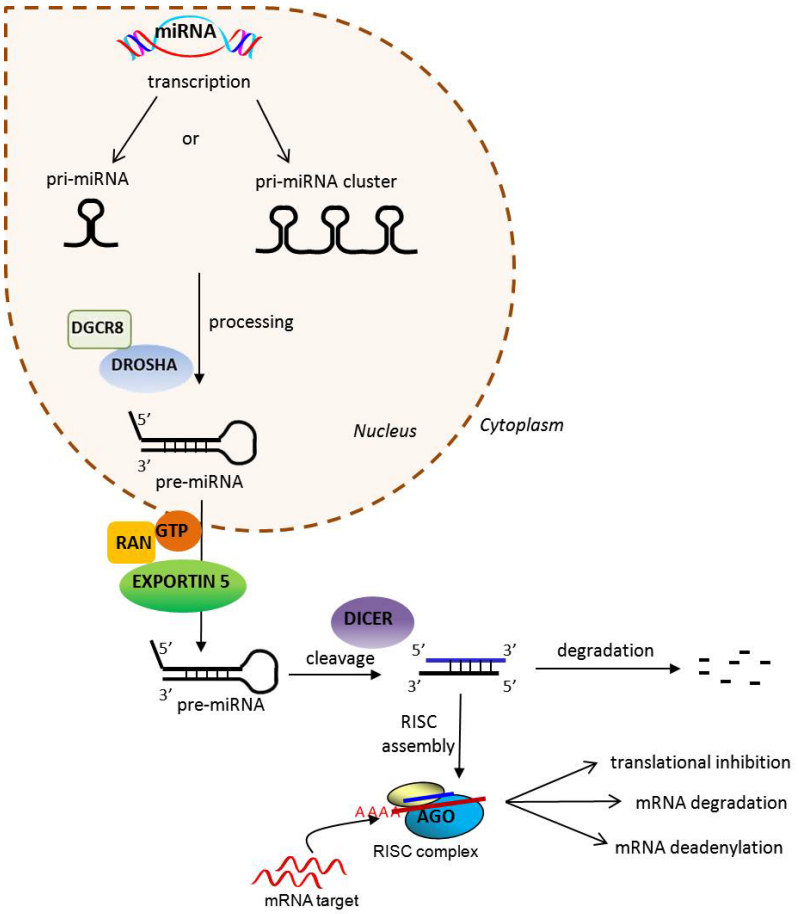
2. MicroRNAs

Over the past decades, it has been estimated that only 2% of the mammalian genome encodes for proteins, whereas the vast majority of it is extensively transcribed to give rise to a large transcriptome of ncRNA species [16]. This pervasive transcription, which has been recently unraveled, thanks to the new improved RNA-sequencing technologies, challenges our traditional understanding of RNAs as simple intermediates between genes and proteins. Extensive studies demonstrated that ncRNAs are functional RNA species, involved in the modulation of primary biological functions, and miRNAs, that are, to date, the most characterized class of small ncRNAs, have been described as fundamental to drive the development and differentiation of immune cells [17–20] .

miRNAs are small, single-stranded endogenous ncRNAs, of about 22 nucleotides (nt), many of which are highly conserved through evolution. They are encoded by genes that produce transcripts for either single or multiple miRNAs or by intronic sequences of protein coding genes. Their expression is highly regulated and can be tissue- and time-specific. Following transcription, the primary RNA transcripts (pri-miRNAs), which form a distinctive hairpin structure, are processed by a complex formed by ribonuclease III-type Drosha and DiGeorge syndrome critical region gene 8 (Drosha/DGCR8) and, then, exported from the nucleus into the cytoplasm as pre-miRNA [21]. In the cytoplasm, pre-miRNAs are recognized and cleaved by another endoribonuclease, Dicer, which

generates a 22 nt double-stranded miRNA. The double stranded miRNA is unwound, and one strand, the guide strand, is loaded into the RNA-induced silencing complex (RISC). Within the RISC, the mature miRNA interacts with Argonaute (Ago) proteins, driving Ago and other associated factors to partially complementary target sites, mainly located in the 3' UTR (3' untranslated region) of the target messenger RNAs (mRNAs) [21–23]. The resulting effect of miRNA-target mRNA interaction inside the RISC complex is the translational inhibition, degradation or deadenylation of the target mRNAs [24,25] (Figure 1).

Figure 1. Key steps in microRNA (miRNA) biogenesis and activity. miRNAs originate from the nucleus as pri-miRNA precursor molecules, organized as single transcriptional units or as a cluster of miRNAs, co-transcribed as a polycistronic transcript. They are processed by the RNAse III-type enzyme, Drosha, in association with the RNA-binding protein, DGCR8, into smaller precursor miRNAs (pre-miRNAs), then exported to cytoplasm, where they are cleaved by Dicer to their mature form of 22 nt double stranded miRNA. The guide strand of the mature miRNA is incorporated into the miRNA-induced silencing complex (miRISC), where it binds to target mRNA by partial complementarity with its 3'UTR. This results in translational inhibition, mRNA degradation or mRNA deadenylation of the recognized miRNA target. Ago, Argonaute.



MiRNAs have been demonstrated to affect the expression of numerous mRNAs within a cell, as well mRNAs can be targets of different miRNAs that act cooperatively. Although the effect of miRNAs on single targets can be relatively modest, the coordinated effect on multiple components of the same pathway adds strength and dynamicity to miRNA-mediated regulation of a biological process.

From the perspective of adaptive immunity, miRNAs have been extensively studied both in physiological and pathological contexts; they have been shown to regulate cell development and differentiation decisions playing a role in lineage commitment and in maintaining lineage stability. These effects are achieved by regulating either, positively or negatively, several signaling pathways.

2.1. miRNAs in T-Cell Development

Lymphoid progenitors developed from hematopoietic stem cells in the bone marrow migrate to the thymus to complete their antigen-independent maturation into functional T-cells. Thymic microenvironment directs T-lymphocyte maturation, characterized by the expression of well-defined cell-surface markers: CD4 and CD8, with thymocytes first starting as $CD4^{-}CD8^{-}$ double negative (DN), then becoming $CD4^{+}CD8^{+}$ double positive (DP) and, lastly, maturing into single positive (SP) $CD4^{+}$ or $CD8^{+}$ T-cells [26].

T-cell-specific deletions of Dicer revealed that the miRNA pathway is required for the normal development of T-cells, as well as for the differentiation of effector T-cell subsets. Indeed, conditional deletions of Dicer in the T-cell lineage resulted in a decrease of T-lymphocytes in thymus and the periphery [8,9,13].

Expression profiling studies, using oligonucleotide arrays (miRNA microarray analysis), small RNA cloning, real-time PCR (polymerase chain reaction) and Next-Generation Sequencing, have shown that miRNA expression changes dynamically during T-cell development and differentiation, implying that the expression of individual miRNAs is associated with different stages of development and different T-cell subsets [27,28]. It was noticed that early B- and T-cell precursors are more closely related to each other than to splenic B-cells and naive T-cells, respectively, and that miRNAs highly expressed in bone marrow progenitor populations were downregulated in thymocytes and other mature lineages. These observations suggested that miRNAs shared among early precursors conceivably regulate their self-renewal and maintenance of an undifferentiated state [29,30]. For instance, one of the most expressed miRNAs in hematopoietic stem cells (HSCs) is miR-125b [29,30], whose levels drop significantly in committed progenitors, indicating that it possibly regulates hematopoiesis at the stem-cell level. This hypothesis is supported by the demonstration that miR-125b increases the size of the hematopoietic stem cell compartment in mice, targeting several component of the pro-apoptotic pathway, as well as genes coding molecules involved in both B- and T-cell differentiation [29,31–33].

The expression of two miRNAs, miR-181a and miR-150, has been shown to be dynamically regulated during thymocyte development.

MiR-181a is transiently upregulated at the late DN to DP stages of T-cell development, while in SP and mature T-cells, its expression is very low [34]. This pattern of expression indicates that

miR-181a plays an important role in the development of T-cells in the thymus. Enforced expression of miR-181a in mature T-lymphocytes results in increased sensitivity to peptide antigens, whereas the inhibition of miR-181a in immature DP cells attenuates their sensitivity and impairs both positive and negative selection [35]. miR-181a modulates T-cell receptor (TCR) signaling by targeting several phosphatases that restrain ERK activity. In light of this evidence, it has been proposed that miR-181a may function as an intrinsic antigen-sensitivity "rheostat", operating during T-cell development, able to quantitatively modulate T-cell sensitivity to antigen, setting a TCR signaling threshold for proper agonist selection [35]. Along this line, expression of miR-181a is altered during autoimmune disease and significantly downregulated in pediatric Systemic Lupus Erythematosus (SLE) patients [36].

Concerning miR-150, it is expressed at high levels in mature naive B- and T-cells and strongly downregulated in their precursors and upon activation. This leads to the conclusion that a common set of miRNAs could be employed in B- and T lineages to regulate similar effector functions, such as tissue homing and cytokine production [37–40]. In particular, miR-150 is strongly upregulated during human T-cell maturation; its expression is low in DN T-cells, increases in DP and $CD8^+$ cells and is higher in $CD4^+$ cells [20]. miR-150 upregulation in this context could be important to downregulate genes associated with an immature phenotype, such as NOTCH 3, which has been identified as a target of miR-150 [20]. The observation that downregulation of miR-150 confers a growth advantage to transformed lymphocytes corroborated the hypothesis that one of the functions of miR-150 in lymphocyte development is to hamper proliferation during specific stages of maturation to allow for differentiation. Importantly, a proper immune response also requires downregulation of miR-150 after activation of mature B- and T-cells in order to release them from growth arrest and permit rapid proliferation to occur [37].

It is worth mentioning that, besides changes in the expression of specific miRNAs, the dynamic regulation of miRNA expression relies also on precursor processing, which appears to vary during T-cell development. Kirigin *et al*. have reported a significant degree of polymorphisms in mature miRNA ends at the late DN to DP stages, suggesting that a stage-specific processing of miRNA precursors may concurrently affect target recognition, thus contributing to proteome changes associated with T-cell development [29].

2.2. miRNAs in T-Cell Function

Antigen-specific naive T-cells, upon activation, proliferate and differentiate into effector T-cells, specialized in either cytotoxic function or cytokine production. Upon antigen recognition, naive $CD4^+$ T-cells proliferate and differentiate into distinct types of memory and effector T-cell subsets, that home to different tissue and produce specific types of cytokines. Among $CD4^+$ effector T-cells, T-helper 1 (Th1) cells are critical for host defense against intracellular pathogens and secrete interferon-γ (IFN-γ), whereas T-helper 2 (Th2) cells control parasitic infections and secrete interleukin 4 (IL-4), IL-5 and IL-13, important for mucosal barrier function, as well as for the induction of B-cell proliferation. T-helper 17 (Th-17) cells promote the response against extracellular pathogens by producing IL-17

and are also involved in autoimmunity. An additional effector population, more recently described, includes Th9 and Th22 cells, respectively, producing IL-9 and IL-22, whose *in vivo* role has not been clearly elucidated, yet [41]. It has been noted that during T-cell differentiation, the global level of miRNA expression inversely correlates with the activation status of the cells being highest in quiescent naive and memory T-cells and sharply decreasing in proliferating and functionally differentiated effector cells. This occurrence suggests that miRNAs may play a general role in stabilizing the expression of genes involved in the maintenance of a naive state. Accordingly, it has been recently observed that Ago proteins are post-transcriptionally downregulated by ubiquitination upon T-cell activation [42]. Ago proteins' downregulation, together with miRISC-accelerated turnover, could contribute to a rapid reprogramming of miRNA repertoire in favor of miRNAs whose expression is induced during T-cell activation, such as miR-182 [43].

The human $CD4^+$ T lymphocytes-specific miRNA signature changes during differentiation from naive to memory cells. In particular, miR-125b, which is highly expressed in human naive $CD4^+$ T-cells and is downregulated during differentiation toward effector T-cells [44], may contribute to maintenance of the naive state of primary human T-cells. In human, miR-125b directly regulates the expression of genes encoding molecules required for the differentiation of different effector T-lymphocytes, such as IFN-γ, peculiar to the Th1 subset, IL-2 receptor-β (IL2RB), which contributes to memory cell identity, IL-10 receptor-α (IL10RA) and the transcriptional repressor, Blimp-1 (PRDM19), which promote T-cell terminal differentiation. Thereby, activation-induced miR-125b down-modulation is associated with increased expression of these genes and acquisition of an effector memory phenotype by $CD4^+$ T-cells [44].

A key feature that characterizes the transition from a naive to an activated status of lymphocytes is the active proliferation, triggered by antigen recognition, which leads to the activation of the TCR signaling pathway. Once T-helper (Th) lymphocytes are induced to proliferate, their clonal expansion becomes independent of further antigen stimulation and is fostered by autocrine or paracrine IL-2 production. In this scenario, miR-182, whose expression is induced by IL-2, has been described as an important modulator of Th cell population, by promoting clonal expansion through post-transcriptional inhibition of forkhead box O1 (Foxo1), a suppressor of proliferation expressed in resting Th lymphocytes [42]. This finding is consistent with the present assumption that miRNAs induced during T-cell activation may contribute to downregulating the expression of genes involved in the maintenance of the naive state to allow for effector cell proliferation and differentiation. Moreover, it has been shown that activated proliferating T-cells express mRNAs with shortened 3' UTR regions and, consequently, fewer miRNA target sites, thus exploiting another mechanism to regulate immune system homeostasis [45].

Among the other miRNAs whose expression is enhanced upon T-cell activation, we will discuss miR-155, miR-146a and cluster miR-17-92 (Table 1).

Table 1. miRNAs in adaptive immunity. HSC, hematopoietic stem cell; DN, double negative; DP, double positive; SP, single positive.

	MicroRNA	Expression	Targets	Function	Related alterations	References
T-cell development	miR-125b	↑ in HSC	Bak	Regulation of HSC compartment size	HSC exhaustion	[29–33]
	miR-181	↑ in DN and DP cells; ↓ in SP and mature T-cells	Dusp5, Dusp6, Shp2, Ptpn22	B-cell lineage differentiation	↑ in SLE	[34–36]
	miR-150	↑ DP CD8+ cells ↓ DN CD8+ cells	Notch3	T-cell development	↓in T-CLL	[20]
	MicroRNA	**Expression**	**Targets**	**Function**	**Related alterations**	**References**
B-cell development	miR-17-92	↑ in progenitor B-cells; ↓ in mature B-cells	Pten, Bim	Defective central memory development	↑ B-cell lymphomas	[46,47]
	miR-150	↑ resting B-cells	cMyb	Impaired B1 cell maturation and Ab response	↓ B-cell lymphomas	[37,38]
	miR-34a	↓ pro-B lymphocytes	FoxP1	Required for pro-B to pre-B-cell transition	↑ B-cell lymphomas	[48]
	miR-155	↑ activated mature B-cells	Ship, C/EbpB, Hdac4	Impaired germinal center B-cell response	↑ diffuse large B-cell lymphoma	[13,49,50]
	miR15a-16	↑ CD5+ B-cells	Bcl2	Defect in apoptosis	↓ in B-CLL	[51]
T-cell function	miR-125b	↑ naïve CD4+ T-cells	Ifng, IL10Ra, IL2Rb, Prdm1	Differentiation of effector T-cells	↓ in SLE	[44]
	miR-182	↑ activated T-cells	Foxo1	Altered T-cell induced inflammation	Altered Treg mediated control of Th2 response	[43]
	miR155	↑ activated CD4+ T-cells	cMaf	Th lineage decisions: KO mice shows ↑ Th2 cells and ↓ Th1 and Th17 cells	↑ in SLE and RA	[13,52]
	miR-146a	↑ CD4+ and CD8+ memory cells	Irak1, Traf6, Fadd	Modulation of IL-2 production and AICD	↑ in RA and ↓ in SLE	[18,52–55]
	miR-17-92	↑ CD4+ and CD8+ memory cells	Pten, Bim	Defect in apoptosis	Lymphoproliferation	[46,56]
Treg cells	miR-10a	↑ Treg	Bcl6, Ncor2	Altered Th17 differentiation	Breakdown of peripheral tolerance	[57]
	miR-155	↑ Treg	Socs1	Altered Treg proliferation and homeostasis	Defect in Treg cell-mediated tolerance	[58–60]
	miR-146a	↑ Treg	Stat1	Altered Treg number, compromised suppressor activity	Autoimmunity due to altered Treg-mediated control of Th1 response	[61]

miR-155, which maps within an exon of the noncoding RNA, Bic [13], has been proven to regulate several aspects of the immune system [62]. miR-155 expression increases upon T-cell activation and controls the differentiation of $CD4^+$ T-cells into different Th cell subsets [11,13], as well as the development of regulatory T (Treg)-cells [58,59]. In miR-155 deficient mice, $CD4^+$ T-cells preferentially differentiate toward the Th2 subtype. Accordingly, *in vivo* expression data from miR-155-deficient mice and *in vitro* reporter assays indicate that musculoaponeurotic fibrosarcoma oncogene homolog (c-Maf), a potent trans-activator of the IL-4 gene promoter, is a target of miR-155 [13]. Therefore, miR-155 may promote differentiation toward a Th1 phenotype by limiting the expression of IL-4, a cytokine whose expression characterizes the Th2 phenotype.

In addition to its function in innate immunity [12,17,63,64], miR-146a is also known to be implicated in the adaptive immune response. In mice, miR-146a expression is higher in Th1 cells and lower in Th2 cells when compared to naive T-cells, suggesting that miR-146a may be involved in the lineage differentiation of T-lymphocytes [11]. In human T-cells, miR-146a is induced upon TCR stimulation and is highly expressed in central memory T-cells [18]. Consistent with this expression profile, miR-146a has been shown to modulate activation-induced cell death (AICD) by targeting the Fas-associated death domain (FADD) and to impair both activator protein 1 (AP-1) activity and interleukin-2 (IL-2) production induced by TCR engagement [18]. Moreover, the observation of increased miR-146a levels in both synovial tissues and in peripheral blood mononuclear cells (PBMCs) of rheumatoid arthritis patients suggest that miR-146a may promote the survival of self-reactive T-cells in autoimmune diseases [52–55]. Finally, in miR-146a-deficient mice, it has been demonstrated that miR-146a contributes to proper resolution of T-cell response, being a part of the negative feedback loop that modulates TCR signaling to the nuclear factor κ-light chain enhancer of activated B-cells (NF-κB). Indeed, NF-κB, a transcription factor activated upon TCR stimulation, induces the expression of miR-146a, which, in turn, downregulates NF-κB activation through repression of NF-κB activators, TNF receptor associated factor 6 (TRAF6) and interleukin-1 receptor-associated kinase 1 (IRAK1) [12,64]. NF-κB binds to the human miR-17-92 promoter, indicating that, also, the miR-17-92 cluster may be one of the intracellular signaling components that promotes proliferation and effector differentiation in response to antigen stimulation. miR-17-92 expression is induced upon T-cell activation during viral infection, while it is downregulated after clonal expansion and during memory cell differentiation [56]. Notably, failure in miR-17-92 downregulation leads to defective central memory cell development [56]. In mice, ectopic expression of the miR-17-92 cluster in the lymphocyte compartment induces a severe lymphoproliferative disease characterized by an expansion of almost all lymphocyte populations, especially the $CD4^+$ T subset [46]. Moreover, B- and T-lymphocytes from transgenic mice show increased proliferation and survival after activation *in vitro* [46]. The phenotype associated with miR17-92 cluster overexpression can be, at least in part, the result of the downregulation of two direct targets of the miR17-92 cluster: bcl-2 interacting protein (Bim), a pro-apoptotic protein belonging to the Bcl2 family, and the tumor suppressor phosphatase and tensin homolog (PTEN).

Studies comparing naive, effector and memory $CD8^+$ T-cells show that a small set of miRNAs is downregulated in effector T-cells compared to naive cells, but also that their expression tends to come back in memory T-cells [52]. Activated T-cells committed to a central memory fate activate a program that enables the acquisition of a central memory phenotype through downregulation of miR-155 and upregulation of miR-150, whose expression is suppressed upon T-cell activation (Table 1). This opposite regulation of miR-150 and miR-155 with respect to the activation state appears coherent, as an effort of the emerging central memory T-cells to distance themselves from the activated state and acquire a maintenance state [39].

2.3. miRNAs in Treg Cells

Treg cells produce inhibitory cytokines, such as IL-10 and transforming growth factor-β (TGF-β), which help to limit immune responses and prevent autoimmunity by suppressing T-lymphocyte activity. Treg cells are critical for the maintenance of immune cell homeostasis, as evidenced by the lethal condition consequent to their ablation in mice [65]. miRNAs have been established to be essential for Treg cell function: indeed, conditional deletion of Drosha or Dicer in forkhead box protein $P3^+$ ($Foxp3^+$) Treg cells induces a fatal, early-onset autoimmune pathology indistinguishable from that observed in Foxp3-knockout mice devoid of Treg cells [10,66,67]. The loss of suppressor function observed in Dicer-deficient Treg cells is most likely due to the deficiency of miRNAs that are over-represented in these cells. Among them, miR-10a, miR-155 and miR-146a have been shown to contribute to distinct aspects of Treg cells homeostasis and function (Table 1). miRNAs confer robustness to differentiation processes by tuning gene expression networks; therefore, they may help to preserve certain phenotypes by targeting transcriptional factor pathways that promote alternative fates. An interesting recent study [57] has shown that miR-10a contributes to the stability of the Treg phenotype and limits the conversion of Treg cells into T-follicular helper (TFH) by targeting the transcriptional repressor B-cell leukemia/lymphoma 6 (Bcl-6). miR-10a is expressed at high levels in naturally occurring Treg (nTreg) cells and is induced by TGF-β and retinoic acid (RA) during inducible Treg (iTreg) cell differentiation from $CD4^+$ T-cells. Actually, TGF-β induces retinoic acid receptor α (RARα) expression, which, upon stimulation by RA, activates miR-10a expression. miR-10a directly targets not only Bcl-6, a master positive regulator of TFH differentiation, but also nuclear receptor corepressor 2 (Ncor2), a corepressor of RARα and, in so doing, activates a positive feedback loop that amplifies its own induction by RA. On the other hand, the miR-10a-induced downregulation of Bcl-6 [57], a negative regulator of T-cell-specific T-box transcription factor (Tbet), elicits higher levels of Tbet, which is known to inhibit Th17 differentiation. Therefore, miR-10a modulates T helper cell stability by restraining, either directly or indirectly, multiple differentiation pathways.

miR-155 not only favors T-cell differentiation toward a Th1 phenotype, but also influences Treg cell homeostasis. miR-155-deficient mice have reduced numbers of Treg cells, both in the thymus and periphery; however, miR-155-deficient Treg cells maintain their suppressor activity [60]. Remarkably,

Foxp3, the key transcription factor controlling Treg cell development and function, upregulates the expression of miR-155, which, in turn, positively regulates Treg cell differentiation by targeting the suppressor of cytokine signaling 1 (Socs1) [60]. Accordingly, T-cell-specific deletion of Socs1 results in an increase in the proportion and absolute numbers of Treg cells in the thymus without affecting their suppressive function [59,60,68]. Therefore, miR-155 stabilizes the transcriptional program induced by Foxp3 by repressing genes whose expression counteracts Treg cell differentiation. In contrast to miR-155-deficient mice that exhibited a decreased number of Treg cells in thymus and in the periphery, miR-146a-deficient mice showed an increased number of Treg cells in the periphery; however, miR-146a-deficient Treg cell suppressor activity was seriously compromised [61]. Fatal immune-mediated lesions observed in mice in the presence of miR-146a-deficient Treg cells were accompanied by sharply augmented Th1 responses and were dependent upon increased amounts of IFNγ. Li-Fan Lu *et al*. demonstrated that miR-146a ensures Treg cell-mediated regulation of Th1 responses, in a significant way, through targeting signal transducer and activator of transcription (STAT1), a transcription factor that regulates IFNγ expression [61].

2.4. miRNAs in B-Cells

In the adaptive immune response, B-cells are responsible for antibody-mediated response. B-cell development begins in primary lymphoid tissue, with subsequent functional maturation in secondary lymphoid tissue. It is a continuum of stages defined by the expression and re-arrangement of functional B-cell receptor (BCR)/immunoglobulin (Ig) genes [2]. The first developmental stage exhibiting commitment to the B-cell lineage is called pro-B and is characterized by rearrangement of the Ig heavy chain. In the pro-B stage, Iga (CD79a) and Igβ (CD79b) are expressed at the cell surface. Rearrangements at the immunoglobulin locus result in the generation and surface expression of the pre-B-cell-receptor (pre-BCR), composed of an Igμ heavy chain and surrogate light chains and, finally, a mature BCR, capable of binding antigen [2]. Mature B-cells that move into the periphery can be activated by antigen, undergo clonal expansion and become an antibody-secreting plasma cell or a memory B-cell, which will respond more quickly to a second exposure to antigen. Unlike conventional B-cells (B-2 cells), which originate from adult bone marrow progenitors, another population of B-lymphocytes, called B-1, appear during fetal life and express surface IgM, but little or no IgD. In mouse, B1 cells can be further subdivided into B-1a ($CD5^+$) and B-1b ($CD5^-$) subtypes [69]. B1 lymphocytes do not undergo much class switch, thus producing antibodies with low affinity and having been implicated in the pathogenesis of autoimmune diseases and many chronic leukemias.

Conditional ablation of Dicer or Ago2 in early B-cell progenitors results in an almost complete block of B-cell development at the pro-B to pre-B transition, which underscores the relevance of miRNA function in B-cell development [70]. Moreover, Dicer-deficient B-cells show increased expression of terminal deoxynucleotidyl transferase (TDT) and increased antibody diversification [47,70,71]. Several specific miRNAs have been proven to control B-cell development, which is impaired in Dicer-deficient mice (Table 1). The miR-17-92 cluster is highly expressed in

progenitor B-cells, and its expression decreases during B-cell maturation, suggesting that it is a positive regulator of B-cell differentiation [47]. In miR-17-92-deficient mice, transition from pro-B-cells to pre-B-cells is compromised, and increased apoptosis occurs, which correlates with higher expression of the pro-apoptotic protein, Bim, a target of the miR-17-92 cluster [46]. Accordingly, overexpression of the miR-17-92 cluster, often detected in human B-cell lymphomas, possibly facilitates transformation by inappropriate repression of the same pro-apoptotic gene, Bim, and the tumor suppressor, PTEN [56]. Another positive regulator of B-cell differentiation is miR-181a, whose ectopic expression in hematopoietic stem cells leads to an increased production of B-lineage cells [72]. Conversely, the overexpression of miR-150 in hematopoietic progenitor/stem cells significantly reduces the mature B-cell population by targeting c-Myb, a transcription factor critically important in B-cell development. MiR-150 is selectively expressed in mature, resting B- and T-cells, but not in their progenitors [9,38]. The expression pattern of c-Myb in B-cells inversely correlates with that of miR-150: it is highly expressed in progenitor cells and downregulated in mature B-lymphocytes. Premature downregulation of c-Myb by ectopic miR-150 expression triggers apoptosis during the pro-B stage, determining the observed phenotype. These findings substantiate the view that higher levels of miR-150 are required during the pre-B to mature B-cell transition to downregulate c-Myb expression and guarantee normal B-cell development [37,38].

The constitutive expression of miR-34a, whose expression is downregulated at the pro-B to pre-B transition, determines a block of the early B-cell development with an accumulation of pro-B-cells and a reduction in pre-B-cells and mature B-cells [48]. These effects have been proven to be the result of targeting Foxp1, a transcription factor required for early B-cell development. Interestingly, the investigators highlight that the effect of miR-34a on the B-cell developmental pathway [48] is consistent with previously reported abnormalities associated with a deficiency of p53, namely, an increased number of pre-B-cells, as well as mature B-cells, with the latter finding being also a consequence of the loss of miR-34a function. This evidence implies a connection between p53 and Foxp1 through the action of miR-34a, which has been demonstrated to be a direct transcriptional target of p53 and to participate in the control of inappropriate cell proliferation [73,74].

miRNAs also regulate B-cell development in the periphery (Table 1). Indeed, Dicer ablation in mature B-cells results in augmented B-cells in the marginal zone and diminished follicular B-cells. miR-155, whose levels are low in progenitor and mature B-cells and are induced upon B-cell receptor ligation, has been implicated in the differentiation of normal activated B-cells. miR-155-deficient mice show multiple defects in the germinal center (GC) response, such as a decreased number of GC B-cells, decreased IgG production, and decreased affinity maturation. Using mouse models, miR-155 has been demonstrated to affect the regulation of the GC response through modulation of cytokine production [13,49] and by direct post-transcriptional regulation of the transcription factor, Pu.1 [75], and activation-induced cytidine deaminase (AID) [50,76].

2.5. miRNAs and the Anti-Viral Immune Response

The ability of small noncoding RNAs to act as immune sentinels was initially described in plants, nematodes, fungi and insects [22,55,64,77] in which the RNAi-mediated recognition of viral RNAs has evolved as a defense mechanism against RNA viruses. In mammals, another class of small RNA, the piwi RNAs (piRNAs), has been shown to suppress transposable element (TE) expression and mobilization through their silencing [78]. Recently, human cellular miRNAs have also been described as intracellular immune mediators, adopted by the organism to fight viruses. The short seed sequence and the tolerance for mismatches represent peculiar properties of miRNAs, which confer them the capability to recognize and downregulate viral mRNAs, required for successful replication and host infection. In mammals, the repression of Dicer and, hence, miRNAs production, confers increased susceptibility to human immunodeficiency virus type 1 (HIV-1), influenza A virus and vesicular stomatitis virus (VSV) [79–81]. In particular, the anti-HIV effects of Dicer are mediated by two cellular miRNAs, miR-17-5p and miR-20a, which target histone acetylase p300/CBP-associated factor (PCAF), an important cofactor of the HIV protein, Tat [79]. Two other miRNAs, miR-93 and miR-24, instead, contribute to the host defense against the rhabdovirus vesicular stomatitis virus (VSV), by targeting complementary sites in the VSV genome [81]]. On the other hand, viruses evolved several mechanisms to evade host defense, including quenching of the miRNA machinery [79], and, also, co-opted miRNA genes to suppress the host antiviral response; therefore, further studies are required to fully understand the role of miRNAs in the complex interaction between viruses and their mammalian hosts.

2.6. miRNAs and Autoimmunity

Given the pivotal role of miRNAs in the modulation of development, maturation and function of lymphocytes in normal immune response, it is not surprising that they have been also involved in the pathogenesis of several immune disorders. An integral aspect of miRNA-mediated regulation is their ability to function over a narrow range of concentration during normal immune responses; they induce subtle variations in the expression of key proteins that, if inappropriate, may lead to an overt pathological phenotype. Many immune diseases arise from the altered balance of gene expression patterns rather than from heavy changes at the levels of protein-coding regions of the genome. As in the case of multiple polymorphisms, which individually have a small effect [82], miRNAs exert their regulatory role through the subtle individual modulation of multiple targets involved in a common signaling pathway rather than through a strong repression of unique targets, thus contributing to lowering the threshold for the onset of an autoimmune response. Notably, up to one half of immune genes are predicted to be under the direct regulation of miRNAs; therefore, mutations that directly affect miRNA expression or alter the miRNA biogenesis pathway may have a causative effect during disease initiation and progression, owing to inappropriate repression or de-repression of key protein targets.

The tight regulation of immune cell activation and the proper selection of a functional, non-self-reactive repertoire of specific antigen receptors is at the base of the fine balance between a normal immune response and the development of autoimmunity. Because of the miRNA ability to regulate the survival and death of lymphocytes [18,51], control over their expression is essential to prevent adaptive immune cells from dysregulated proliferation and activity leading to leukemia and autoimmunity [83,84]. Several miRNAs exhibit a different expression profile in patients affected by autoimmune diseases compared to normal subjects [84–86], suggesting a causative role for the pathogenesis of human autoimmune diseases, including rheumatoid arthritis (RA), psoriasis and SLE. The expression of some miRNAs, like miR-146a and miR-155, appears to be constantly dysregulated in the aforementioned autoimmune diseases.

SLE is an autoimmune disease characterized by the loss of immune tolerance, resulting in activation and expansion of autoreactive $CD4^+$ T-helper cells [87]. In MRL/lpr mice, a murine model of SLE characterized by the accumulation of $CD4^+CD25^+Foxp3^+$ Treg cells in lymphoid organs [88], the onset of the autoimmune disease correlates with a reduced functional capacity of Treg cells that show a distinct phenotype (*i.e.*, increased CD69 and reduced CD62L expression). CD62L, an L-selectin, which determines the homing properties of T-lymphocytes, is a target of miR-155, one of the miRNAs upregulated in SLE. These data indicate that miR-155 dysregulated expression may directly contribute to the accumulation of Treg cells in the lymphoid organs of MRL/lpr mice [89]. On the contrary, miR-146a expression is reduced in the PBMCs of SLE patients. MiR-146a downregulation inversely correlates with clinical disease activity and with interferon (IFN) scores and may be responsible for the IFN overproduction observed in SLE. This phenotype is consistent with miR-146a-mediated repression of the type I IFN pathway, through the direct targeting of STAT-1 and IFN regulatory factor 5 (IRF5) [90].

In RA patients, upregulation of miR-146a and miR-155 has been found in PBMCs, as well as in the synovial tissue [52–54]. miR-146a also is highly expressed in $CD4^+$ T lymphocytes from RA patients and exhibits a strong correlation with increased levels of TNFα [86]. Interestingly, T-cells isolated from joint tissue and synovial fluid of RA patients show an activated and memory phenotype and are resistant to apoptosis, although the high levels of pro-apoptotic factors, like Fas-ligand (FasL), TNFα and TNF-related apoptosis-inducing ligand (TRAIL). Given that our study [18] and others [54,91] demonstrated an anti-apoptotic role of miR-146a in T-lymphocytes and other cell types, the increased survival of T-lymphocytes in RA patients could be in part explained by the enhanced expression of miR-146a. Another anti-apoptotic miRNA that has been related to the pathogenesis of autoimmune disorders is miR-17-92. Its overexpression results in a decreased activation-induced cell death, sustained lymphoproliferation and a marked presence of serum autoantibodies and lymphoid infiltrates [46]. Although the development of T-lymphocytes does appear normal, the relevant increase in the number of mature $CD4^+$ T-cells and their activated profile strongly suggest a failure of peripheral tolerance. The molecular mechanisms contributing to this immune disease-related phenotype of miR-17-92 transgenic mice are based on the downregulation of the tumor suppressor, PTEN, and the pro-apoptotic,

Bim [46]. An interesting study demonstrating the involvement of miRNAs in the possible alteration of autoimmune-related signaling pathways was performed in 2007 in sanroque mice, a murine model resembling a lupus-like autoimmune syndrome [92]. These mice are homozygotes for an M199R mutation in the ROQ domain of Roquin, a protein that binds to mRNAs and targets gene expression post-transcriptionally. The evident accumulation of autoimmune lymphocytes in this murine model depends, at least in part, on the failure to repress the expression of the co-stimulatory molecule inducible co-stimulatory molecule (ICOS), particularly in naive T-cells. Roquin normally limits ICOS expression by promoting the degradation of ICOS mRNA, by a mechanism that involves the recognition of a conserved miRNA binding sequence within the 3' UTR. It has been reported that miR-101 is required for the Roquin-mediated degradation of ICOS mRNA [92], and the alteration of the seed site for miR-101 in the 3' UTR of ICOS disrupts the inhibitory activity by Roquin. These results demonstrate a critical miRNA-mediated regulatory pathway that prevents lymphocyte accumulation and autoimmunity. Further evidence supporting a role of miRNAs in autoimmunity has been provided by Ebert *et al*., who demonstrated the relevance of miR-181a expression for the elimination of self-reactive thymocytes; indeed, the inhibition of miR-181a expression during thymic development converts endogenous positively selecting peptides into autoantigens [93].

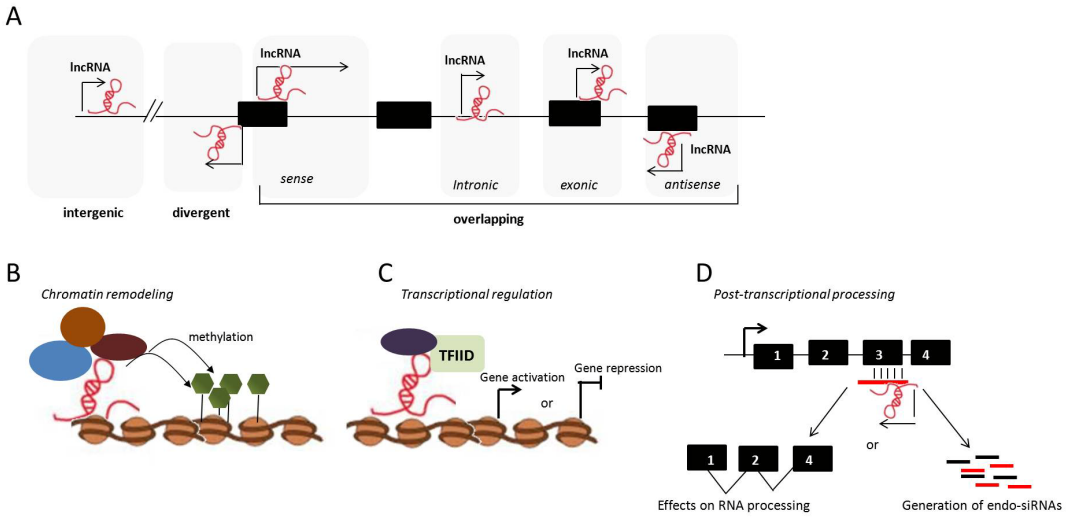
Along with the miRNA-mediated deregulation of disease susceptibility genes, few studies also detected autoantibodies targeting key components of the RNAi/miRNA machinery, including Argonaute proteins (referred to as anti-Su antibodies) and Dicer, in patient blood. The anti-Su autoimmune sera recognize several members of the Ago protein family, which are known to associate into RISC and which have a high degree of sequence identity. Despite that the first characterization of autoantigens recognized by anti-Su autoantibodies [94] had been provided in 1984, the disease specificity of anti-Su antibodies has been addressed in only a few studies, so far. They were initially reported to be specifically associated with SLE [94], but later studies showed that anti-Su antibodies are, instead, found in a variety of systemic rheumatic diseases [95]. The existence of an autoimmune response directed against macromolecular complexes involved in the post-transcriptional regulation of gene expression may indicate the involvement of the miRNA biogenesis pathway in the production of autoantibodies, further supporting the intriguing hypothesis that dysfunction of the RNAi/miRNA machinery lies at the core of the pathogenesis of autoimmune diseases.

3. Long Noncoding RNAs

At least 80% of mammalian genome transcription results in the generation of lncRNAs. It has been estimated that 10,000 to 200,000 different types of lncRNAs are transcribed in the human genome, but functional and molecular mechanisms have been elucidated only for a few [96]. Given their unexpected abundance, lncRNAs were initially considered as spurious transcriptional noise, without biological function, and this assumption was corroborated by the observation that many of the transcripts are expressed at very low levels and exhibit no sequence conservation [97,98].

Nevertheless, it is now apparent that lncRNAs are functional, and many of them are differentially expressed in particular developmental stages.

Figure 2. Long non coding RNAS (lncRNAs): mechanism of action and genomic organization. Illustrative representation of lncRNAs classification relative to their genomic position or mechanism of action. (**A**) lncRNAs are generally classified according to their proximity to protein coding genes in the genome: intergenic lncRNAs are distant from protein coding regions; divergent lncRNAs are located on the opposite strand closed by the transcription starting site of a protein coding gene; intragenic lncRNAs overlap protein coding genes and can be sense, antisense, intronic or exonic; (**B**) lncRNAs can recruit chromatin remodeling complexes to specific genomic loci, to epigenetically mark the region for gene silencing; (**C**) lncRNAs can permit the recruitment of transcription factors activating gene expression or can interfere with the transcription machinery, occluding the access of transcription factors and, thereby, silencing the gene; (**D**) lncRNAs that are antisense to protein coding genes may regulate the splicing or induce the degradation of their corresponding mRNA transcripts.



LncRNAs comprise many different transcripts, ranging from several hundreds to ten-thousand nucleotides and can be classified according to their genomic position as intergenic or intragenic lncRNAs (Figure 2A). Intergenic lncRNAs are located distant from protein coding regions, whereas divergent lncRNAs, often referred to as bidirectional lncRNAs, are located on the opposite strand from a coding gene whose transcription is initiated in close proximity. Intragenic (overlapping) lncRNAs are transcribed along one or multiple exons (exonic lncRNA) or introns (intronic lncRNA) of another transcript on the same or opposite strand and can then be spliced into a mature RNA (Figure2A).

Unlike miRNAs, lncRNAs possess heterogeneous molecular mechanisms of action, making them master regulators of intracellular networks and pathways in both physiology and diseases [96]. Some mammalian lncRNAs have two unique relevant properties. The first one is the ability to target a single

location through the use of a large sequence space. While transcription factors are effectively recruited to the DNA promoter region by recognizing short DNA motifs, which typically occur thousands of times in the genome, lncRNAs, like Tsix and RepA/Xist, are able to deliver epigenetic complexes to a unique site, thus providing a spatio-temporal regulatory specificity not achievable with proteins and small RNAs [63,99]. A second characteristic of some lncRNA-mediated regulation is the "allelic memory". Proteins loose the memory of their transcriptional origin when the mRNA is shuttled to the cytoplasm to be translated; lncRNAs, instead, remain tethered to the site of transcription and can, therefore, have allele-specific action. Differently from miRNAs, which are mainly post-transcriptional repressors, lncRNAs have a full range of functions, being able to act as scaffold RNAs, co-activators or co-repressors of trans-acting RNAs and epigenetic regulators of protein-coding gene expression. In particular, they can mediate epigenetic changes by recruiting chromatin remodeling complexes to specific genomic loci, thus altering the chromatin status at gene promoters to affect gene expression [100] (Figure 2B). Another major mechanism of lncRNA action is transcriptional regulation, exerted through specific interaction with transcription factors or other protein components of the transcriptional machinery. LncRNAs may sequester transcription factors in the cytoplasm, preventing their translocation into the nucleus [101] or occlude a transcription factor binding site, thereby promoting gene repression [102]. LncRNAs can also promote gene expression, through interactions with components of the core transcriptional machinery, thus enhancing the specific binding of transcription factors to promoter regions (Figure 2C). Moreover, lncRNAs have been implicated in posttranscriptional regulation, either by the direct regulation of alternative splicing through base-pairing interactions or by gene silencing via the generation of endo-siRNAs and the consequent degradation of the targeted transcript (Figure 2D).

3.1. lncRNAs and the Adaptive Immune Response

In contrast to the wealth of recent evidence demonstrating lncRNA involvement in cancer, at present, a role of lncRNAs in immune system regulation is attested by only a handful of mechanistically distinct examples. The importance of lncRNAs in the immune system has been recently highlighted by a report that identified and characterized a set of lncRNAs preferentially expressed in naive and memory mammalian $CD8^+$ T-lymphocytes [14,92]. Over 1000 lncRNAs have been detected and characterized in humans and mice, many of which expressed in a stage- or cell-specific manner. In particular, the authors identified 96 lymphoid-specific lncRNAs. Interestingly, 29 of these transcripts were specific for $CD8^+$ T lymphocytes, 21 lncRNAs were significantly modulated during the T-cell differentiation process and 81 were regulated during effector T-cell activation [14]. These expression patterns strongly suggest a possible role of lncRNA in the modulation of the key cellular processes of T-lymphocyte biology. Remarkably, lncRNAs, as well as the recently discovered large intergenic noncoding RNAs (lincRNAs), have been shown to regulate the expression of adjacent protein-coding genes. Actually, several of the lncRNAs identified in the mentioned study are associated with protein-coding genes that exert key functional roles in T-cells. For

example, an lncRNA, AK009498, overlaps the transcription start site of caspase-8 (FLICE)-like inhibitory protein (Flip), a gene fundamental for the survival and development of T-lymphocytes. Flip is upregulated during the transition to an activated or a memory phenotype, whereas the host lncRNA is downregulated, thus suggesting a potential negative function of AK009498 [14] (Table 2).

Table 2. lncRNAs in adaptive immunity.

	lncRNA	Genomic coordinates	Lenght	Expression	Characteristics	Proposed function	References
has lncRNAs	**Tmevpg1 (NeST)**	chr12: 68, 383, 225-68, 415, 107	32 kb	Activated $CD8^+$ T cells $CD4^+$ Th1 and Th17 cells	Its expression is regulated by Tbet and STAT4	Epigenetic control of Ifnγ locus	[103–105]
	NRON	chr9: 129, 170, 054-129, 172, 783	2.7 kb	$CD4^+$ and $CD8^+$ T cells	Sequence omology conserved back to chicken	Modulator of NFAT nuclear trafficking	[101]
mmu lncRNAs	**AK009498**	chr1: 58, 711, 508-58, 713, 886	710 bp	$CD4^+$ and $CD8^+$ effector and memory cells	Overlaps flip transcription start site	Potential negative regulator of Flip	[14]
	Lef1as (AK029296)	chr3: 131, 109, 026-131, 112, 090	3 kb	naive $CD8^+$ T cells	Located antisense to Lef1 mRNA	Possible role in suppressionLef1	[14]
	AK053349	chr14: 115, 044, 497-115, 046, 726	2.2 kb	effector $CD8^+$ T cells	Partially overlaps miR-17-92 cluster	unknown	[14]
	AK020764	chr11: 87, 755, 577-87, 757, 267	1.6 kb	effector $CD8^+$ T cells	It hosts in its first intron miR-142	Possible functional link with miR-142	[14]

Nettoie Theiler's Pas Salmonella (NeST), also called Theiler's Murine Encephalitis Virus Possible Gene1 (TMEVPG1), a nuclear RNA initially described as involved in the control of viral load during persistent infection, is the first lncRNA of the immune system that has been proven to regulate the expression of a master cytokine, such as IFN-γ [103]. This lncRNA is expressed selectively in Th1 cells relative to Th2 and Th17 cells, and its expression depends on Th1-specific transcription factors, STAT4 and Tbet [104]. Moreover, NeST is located adjacent to the IFNγ locus in both humans and mice and is encoded in antisense respect to the IFNγ gene (Table 2). Recent evidence demonstrated that NeST acts in trans and positively regulates the IFNγ gene at the transcriptional level, by

interacting with chromatin-modifying complexes. More precisely, it binds WDR5, a member of the histone H3 lysine 4 methyltransferase complex, and promotes permissive methylation marks on histone H3 at the IFNγ locus. The effect of such epigenetic control is the modulation of the host response to viral and bacterial pathogens, thus suggesting a role for lncRNAs as functional links between immune regulation and susceptibility to infections [105]. Indeed, alteration of NeST expression could contribute to differences in T-lymphocyte response and disease susceptibility. The discovery of functional intergenic lncRNAs, whose expression is critical for proper gene regulation, may explain the relevance of some intergenic regions as heritable causes of human diseases.

Another interesting feature of lncRNAs is that a small number of lncRNA genes host in their sequence small RNAs and may be processed into, and exert their effects via, smaller functional ncRNAs. Pang *et al*. identified 18 lncRNAs expressed in murine $CD8^+$ T-lymphocytes overlapping with annotated miRNAs and 21 that harbor internally small nucleolar RNAs (snoRNAs) [14]. In this context, AK053349, one of the most highly upregulated lncRNA during effector T-cell activation, partially overlaps miR-17-92, a cluster of miRNAs, often amplified in lymphomas, whose enforced expression in mouse lymphocytes results in lymphoproliferative disease and autoimmunity [46]. Furthermore, miR-142-5p and miR-142-3p, which have been reported as highly expressed in naive T-cells and are involved in the modulation of Foxp3 in Treg lymphocytes, are hosted in the first intron of a lncRNA (AK020764), strongly upregulated in $CD8^+$ T-cells [14] (Table 2). The colocalization could imply a functional link, but it has not yet been verified.

Finally, lncRNAs may influence cell regulatory networks not only at the transcriptional level, but also by controlling protein trafficking, as it has been shown for the noncoding repressor of NFAT (NRON), a negative regulator of the transcription nuclear factor of activated T-cells (NFAT). NRON is an lncRNA enriched in lymphoid tissues, such as the thymus, the spleen and the lymph nodes. This lncRNA has a cytosolic localization and has been proven to interfere with NFAT function by interacting with several proteins (Table 2). One of these proteins is importin beta1 (KPNB1), which, in the presence of NRON, sequesters NFAT in the cytosol. In addition, the interactions of NRON with two other proteins, the serine/threonine protein, phosphatase 2A (PP2A), and the IQ motif containing GTPase activating protein 1 (IQGAP1), may also be significant [101]. These examples underscore the possibility of limitless mechanisms of gene regulation based on cooperation between ncRNAs and proteins.

3.2. lncRNAs as Biomarkers in Autoimmunity and Leukemia

Genome-wide analysis identified a number of highly conserved transcripts, some of which do not encode for proteins and are, therefore, considered non-genic. Among these, the ultra-conserved regions (UCRs) are noncoding genomic sequences of over 200 bases, located in both intra- and inter-genic regions, and are strongly conserved between human, mouse and rat genomes. The unexpected degree of conservation of UCRs among distant species suggests that they may have essential functional importance for the ontogeny and phylogeny of mammals and vertebrates, in general. A large

proportion of transcripts derived from these UCRs have significant RNA secondary structures and are components of clusters containing other sequences with functional noncoding significance [106]. Notably, the highest fraction of transcribed UCRs was found in B-lymphocytes, and distinct UCR signatures have been described as associated with human leukemias and carcinomas. In particular, Calin *et al*. reported a differential expression of 19 UCRs (eight up- and 11 down-regulated) in chronic lymphatic leukemia (CLL) with respect to normal hematopoietic tissues [106,107]. The authors also identified a cluster of seven UCRs (uc. 347 to uc. 353) located within the cancer-associated genomic region (CAGR), and two of them—uc. 349A(P) and uc. 352(N)—are among the transcribed UCRs (T-UCRs) differentially expressed between normal and malignant B-CLL $CD5^+$ cells [106]. Furthermore, the authors correlated a signature of five T-UCRs, namely, uc. 269A(N), uc. 160(N), uc. 215(N), uc. 346A(P) and uc. 348(N), with different CLL prognosis groups, consistent with the expression of 70-kDa zeta-associated protein (ZAP-70), an established prognostic marker for CLL. These data suggest that T-UCRs could be candidate genes for cancer susceptibility. Another interesting aspect emerging from this study is the finding that the expression levels of the five T-UCRs described above negatively correlated with a miRNA expression signature reported in CLL [107]. This evidence represents just an example of how lncRNAs and small RNA biology are strictly connected: a significant fraction of lncRNAs seem to act as precursors for small RNA species, in particular for miRNAs. These findings raise the possibility of complex regulatory mechanisms between lncRNAs and small RNAs, suggesting a possible interplay between the different classes of noncoding RNAs.

Given the emerging role of lncRNAs in biomolecular regulatory interactions within cells and their relevance in the etiology of human disease and cancer, it becomes highly pertinent and imperative to perform an exhaustive evaluation of their functionalities in order to shed light on the specific pathways and regulatory circuits in which they are involved.

4. Secreted Noncoding RNAs

The introduction of Next-Generation Sequencing (NGS) techniques led to the identification of a large number of unexpected ncRNAs, not only at the level of different human tissues, but also in extracellular fluids (e.g., serum, plasma, urine, milk and saliva) [108–112]. The study of circulating RNA populations has almost exclusively focused on miRNAs, most likely due to the availability of array hybridization techniques to detect these small RNAs and, also, because of their recognized role in several biological processes. Alterations in the level and composition of these extracellular miRNA populations are strictly related to different pathologies, including cancer [20,108], diabetes [108] and tissue injury [113]. Extracellular ncRNAs circulate in human plasma within non-vesicular Ago2 ribonucleoprotein complexes, which confer them stability [114]. Alternatively, they can be stably carried in body fluids, as packaged within membranous vesicles (including exosomes, shedding vesicles and apoptotic bodies) [115–117] and spread signals that affect neighboring or distant cells. These different types of microvesicles protect ncRNAs from degradation during systemic transport and, also, are responsible for the specific delivery of them to recipient cells. Exosomes are small

vesicles ranging from 30 to 100 nm in size, originated from endosomes. They consist of a lipid bilayer membrane surrounding a small amount of cytosol. The formation and release of exosomes are tightly regulated by multiple signaling mechanisms. Different stimuli can influence exosome secretion, including bacterial lipopolysaccharide on macrophages and dendritic cells [118]. Shedding vesicles are much larger than exosomes and are heterogeneous in size, ranging from 100 nm to 1 μm [119]. The presence of a ncRNA pool has been reported in exosomes and shedding vesicles derived from a variety of sources, including mast cells [115] blood [120], stem cells [121] and adipocytes [122]. The existence of different secretion mechanisms suggests that specific ncRNA populations may be delivered from different cell types and, therefore, have different fates and functions. It has been speculated that differences between vesicle-enclosed and Ago2-associated ncRNAs may reflect cell type-specific ncRNA release mechanisms. Additional studies are needed to unravel the pathways responsible for ncRNA loading and secretion and to better define the biological function of the different circulating ncRNA species.

Role of Secreted ncRNAs in Immunological Processes

The first indication of the horizontal transfer of nuclei acids in mammalian cells reported the release of miRNAs and mRNAs by mast cells and their functional transfer to vesicle-targeted cells [115]. This evidence, together with intensive studies describing the delivery of ncRNA-loaded exosomes by immune cells [123,124], strongly suggests that secreted RNAs could represent a new intricate level of cellular communication and regulation of immunological processes. To mount an effective immune response, different immune cells need to exchange information and develop highly specialized structures, called “cell synapses”. During the formation of immunological synapse (IS) between T-lymphocytes and antigen presenting cells (APCs), membrane and transmembrane-associated molecules are rearranged into a highly organized structure at the T-cell-APC contact site [125]. The formation of IS is consequent to antigen recognition and allows the initiation and tuning of T-lymphocyte activation. Interestingly, it has been reported that miRNAs are exchanged during cognate immune interactions, and this delivery is strictly dependent on the formation of IS [124]. These specialized structures promote the unidirectional transfer of miRNA-loaded exosomes from T-lymphocytes to APCs, and this kind of genetic communication is antigen-driven and CD63-dependent. Synaptically transferred miRNAs are functional in recipient cells, as demonstrated by Mittelburn *et al*. [124], showing that antigen-dependent transfer of miR-335 from T-lymphocytes to APC resulted in the targeting of specific genes, involved in the modulation of immune response. Notably, not all miRNAs can be incorporated into exosomes, and their presence into delivered microvesicles is a selective process. The miRNA composition in circulating RNAs does not merely reflect the cellular miRNA pool: certain miRNAs are equally abundant between cellular and extracellular environment; others (like miR-760, miR-362, miR-654-5p and miR-671-5p) are significantly more abundant in exosomal fractions derived from all cell types; and finally, another set (for example, miR-21-3p, miR-101 and miR-32) are, instead, more represented in cells than in

exosomes [124]. Moreover, it has been also observed that there is a higher similarity between miRNAs composition in exosomes of different cellular origin than between shuttle RNA and their corresponding cellular miRNAs [123,124]. These data demonstrate that specific miRNA populations are selectively sorted into shuttle RNA, strongly suggesting that cells specifically release these RNA molecules to affect the function of target cells. However, to date, the precise mechanism controlling this selection is still unclear. Together with synaptic-dependent shuttling of miRNA from T-cells to APC, another important source of RNA-loaded exosomes is represented by dendritic cells (DCs), which have a role in the activation of T-lymphocytes. In that context, among the most representative miRNAs vehiculated by exosomes, miR-223 and miR-93 have validated target genes that play important roles in immunity. In particular, miR-223 downregulates myocyte enhancer factor 2C (MEF2C), involved in the transcription of interleukins, whereas miR-93 targets STAT3, thus affecting T-cell activation processes [123]. More interestingly, profiling of shuttle RNA isolated from DC-derived exosomes also revealed the presence of many small noncoding RNA species, other than miRNAs, that could act as regulatory RNAs, in particular, snoRNAs. Altogether, this first piece of evidence suggests a wide range of biological effects that could be mediated by shuttle RNA, and the recent findings that the shuttle RNA population is not restricted to miRNAs, but includes other ncRNAs reciprocally exchanged between immune cells, provide a further level of cellular inter-communication that contributes to affecting the activation and the efficacy of immunological response.

5. Concluding Remarks

The proper functioning of the immune system is accomplished through several steps of development and differentiation that must be strictly regulated to respond, at the right time and with the appropriate effector cells, to different insults. Thanks to the advancement of technologies that allow us to have a global picture of which proteins and functional RNAs are expressed in a cell at a defined time, we have now the instruments to dissect the complexity of this regulation, which is time- and cell-specific, allowing the immune system to dynamically adapt to the many challenges it faces. The crosstalk between signaling pathways and epigenetic processes during development and cell differentiation is still poorly understood. Changes in these networks might lead to deregulation of gene expression, which ultimately results in diseases, e.g., autoimmunity and cancer.

It is conceivable that transcriptional control, mediated by epigenetic modifications, is preferentially responsible for the developmental decision, while posttranscriptional control plays a major role during activation and differentiation. This concept is particularly relevant in the immunology field, where recent advances in the “-omics” techniques have revealed a multilayered network composed of protein-coding RNAs, as well as of ncRNAs.

Although the miRNA and lncRNA pathways were initially believed to be independent and distinct, the lines distinguishing them continue to fade. There is now compelling evidence that, despite their differences, these distinct ncRNA pathways are interconnected, interact, compete and rely on each

other at several levels as they regulate genes and protect the genome from external and internal threats. However, there are still major gaps in our understanding of their function at the molecular level. It is becoming evident that miRNAs, more than being molecules that act at the post-transcriptional level to switch on and off the expression of specific genes at a precise stage, have a more intricate role. On the one hand, being part of regulatory loops, miRNAs may contribute to the maintenance of a certain state of the cell; on the other hand, because of their transient rising, induced by external stimuli and activation of specific signaling pathways, miRNAs help to rapidly remodel the proteome of a cell, rendering it functional for the needs of the organism. Furthermore, we have now to consider the primary contribution of lncRNAs, regulatory molecules that have the prerogative to join some functional peculiarity of proteins and the capability to recognize unique loci in the genome. LncRNAs are responsible for transcriptional control of gene expression, both by site-specific recruitment of chromatin remodeling complexes (epigenetic changes) and by favoring or inhibiting the recruitment of transcription factors on defined promoters. These mechanisms may affect the expression timing of proteins, as well as of miRNAs, thus contributing to shape both the proteome and the miRNome of a developing cell. However, lncRNAs may exert their control also post-transcriptionally by driving molecular complexes responsible for editing, splicing, transport and degradation towards mRNA molecules. At the post-transcriptional level, lncRNAs will work alongside with miRNAs, thus contributing to establish a more flexible and fine-tuned system of regulation. This functional link, evident in adaptive immunity system, reveals a more general mechanism for the control of patterning and lineage commitment. Understanding how ncRNAs are regulated and exert their function in the context of the immune system will encourage further advancements on studies in this field and the exploitation of ncRNAs as possible candidates for future therapies.

Acknowledgments

We thank C. Cicchini and A. Marchetti for critical reading of this manuscript.

This study was supported by Ministero dell'Istruzione dell'Università e della Ricerca PRIN Research Grants 2002061255.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Lim, P.S.; Li, J.; Holloway, A.F.; Rao, S. Epigenetic regulation of inducible gene expression in the immune system. *Immunology* **2013**, *139*, 285–293.
2. Blom, B.; Spits, H. Development of human lymphoid cells. *Annu. Rev. Immunol.* **2006**, *24*, 287–320.

3. Sheik Mohamed, J.; Gaughwin, P.M.; Lim, B.; Robson, P.; Lipovich, L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *RNA* **2010**, *16*, 324–337.
4. Orom, U.A.; Derrien, T.; Beringer, M.; GumiReddy, K.; Gardini, A.; Bussotti, G.; Lai, F.; Zytnicki, M.; Notredame, C.; Huang, Q.; *et al*. Long noncoding RNAs with enhancer-like function in human cells. *Cell* **2010**, *143*, 46–58.
5. O'Connell, R.M.; Rao, D.S.; Chaudhuri, A.A.; Baltimore, D. Physiological and pathological roles for microRNAs in the immune system. *Nat. Rev. Immunol.* **2010**, *10*, 111–122.
6. Mercer, T.R.; Dinger, M.E.; Mattick, J.S. Long non-coding RNAs: Insights into functions. *Nat. Rev. Genet.* **2009**, *10*, 155–159.
7. Calin, G.A.; Croce, C.M. microRNA signatures in human cancers. *Nat. Rev. Cancer* **2006**, *6*, 857–866.
8. Cobb, B.S.; Nesterova, T.B.; Thompson, E.; Hertweck, A.; O'Connor, E.; Godwin, J.; Wilson, C.B.; Brockdorff, N.; Fisher, A.G.; Smale, S.T.; *et al*. T-cell lineage choice and differentiation in the absence of the RNase III enzyme Dicer. *J. Exp. Med.* **2005**, *201*, 1367–1373.
9. Muljo, S.A.; Ansel, K.M.; Kanellopoulou, C.; Livingston, D.M.; Rao, A.; Rajewsky, K. Aberrant T-cell differentiation in the absence of Dicer. *J. Exp. Med.* **2005**, *202*, 261–269.
10. Chong, M.M.; Rasmussen, J.P.; Rudensky, A.Y.; Littman, D.R. The RNAseIII enzyme Drosha is critical in T-cells for preventing lethal inflammatory disease. *J. Exp. Med.* **2008**, *205*, 2005–2017.
11. Monticelli, S.; Ansel, K.M.; Xiao, C.; Socci, N.D.; Krichevsky, A.M.; Thai, T.H.; Rajewsky, N.; Marks, D.S.; Sander, C.; Rajewsky, K.; *et al*. microRNA profiling of the murine hematopoietic system. *Genome Biol.* **2005**, *6*, R71.
12. Taganov, K.D.; Boldin, M.P.; Chang, K.J.; Baltimore, D. NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 12481–12486.
13. Rodriguez, A.; Vigorito, E.; Clare, S.; Warren, M.V.; Couttet, P.; Soond, D.R.; van Dongen, S.; Grocock, R.J.; Das, P.P.; Miska, E.A.; *et al*. Requirement of bic/microRNA-155 for normal immune function. *Science* **2007**, *316*, 608–611.
14. Pang, K.C.; Dinger, M.E.; Mercer, T.R.; Malquori, L.; Grimmond, S.M.; Chen, W.; Mattick, J.S. Genome-wide identification of long noncoding RNAs in CD8+ T-cells. *J. Immunol.* **2009**, *182*, 7738–7748.
15. Pagani, M.; Rossetti, G.; Panzeri, I.; de Candia, P.; Bonnal, R.J.; Rossi, R.L.; Geginat, J.; Abrignani, S. Role of microRNAs and long-non-coding RNAs in CD4(+) T-cell differentiation. *Immunol. Rev.* **2013**, *253*, 82–96.

16. Hangauer, M.J.; Vaughn, I.W.; McManus, M.T. Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLoS Genet.* **2013**, *9*, e1003569.
17. Curtale, G.; MiRolo, M.; Renzi, T.A.; Rossato, M.; Bazzoni, F.; Locati, M. Negative regulation of Toll-like receptor 4 signaling by IL-10-dependent microRNA-146b. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1499–11504.
18. Curtale, G.; Citarella, F.; Carissimi, C.; Goldoni, M.; Carucci, N.; Fulci, V.; Franceschini, D.; Meloni, F.; Barnaba, V.; Macino, G. An emerging player in the adaptive immune response: microRNA-146a is a modulator of IL-2 expression and activation-induced cell death in T lymphocytes. *Blood* **2010**, *115*, 265–273.
19. Rossato, M.; Curtale, G.; Tamassia, N.; Castellucci, M.; Mori, L.; Gasperini, S.; Mariotti, B.; de Luca, M.; MiRolo, M.; Cassatella, M.A.; *et al*. IL-10-induced microRNA-187 negatively regulates TNF-alpha, IL-6, and IL-12p40 production in TLR4-stimulated monocytes. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E3101–E3110.
20. Ghisi, M.; Corradin, A.; Basso, K.; Frasson, C.; Serafin, V.; Mukherjee, S.; Mussolin, L.; Ruggero, K.; Bonanno, L.; Guffanti, A.; *et al*. Modulation of microRNA expression in human T-cell development: targeting of NOTCH3 by miR-150. *Blood* **2011**, *117*, 7053–7062.
21. Lee, Y.; Ahn, C.; Han, J.; Choi, H.; Kim, J.; Yim, J.; Lee, J.; Provost, P.; Radmark, O.; Kim, S.; *et al*. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **2003**, *425*, 415–419.
22. Hutvagner, G.; McLachlan, J.; Pasquinelli, A.E.; Balint, E.; Tuschl, T.; Zamore, P.D. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **2001**, *293*, 834–838.
23. Matranga, C.; Tomari, Y.; Shin, C.; Bartel, D.P.; Zamore, P.D. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **2005**, *123*, 607–620.
24. Lewis, B.P.; Shih, I.H.; Jones-Rhoades, M.W.; Bartel, D.P.; Burge, C.B. Prediction of mammalian microRNA targets. *Cell* **2003**, *115*, 787–798.
25. Nilsen, T.W. Mechanisms of microRNA-mediated gene regulation in animal cells. *Trends Genet.* **2007**, *23*, 243–249.
26. Kisielow, P.; von Boehmer, H. Development and selection of T-cells: Facts and puzzles. *Adv. Immunol.* **1995**, *58*, 87–209.
27. Wu, H.; Neilson, J.R.; Kumar, P.; Manocha, M.; Shankar, P.; Sharp, P.A.; Manjunath, N. miRNA profiling of naive, effector and memory CD8 T-cells. *PLoS One* **2007**, *2*, e1020.
28. Neilson, J.R.; Zheng, G.X.; Burge, C.B.; Sharp, P.A. Dynamic regulation of miRNA expression in ordered stages of cellular development. *Genes Dev.* **2007**, *21*, 578–589.

29. Kirigin, F.F.; Lindstedt, K.; Sellars, M.; Ciofani, M.; Low, S.L.; Jones, L.; Bell, F.; Pauli, F.; Bonneau, R.; Myers, R.M.; *et al*. Dynamic microRNA gene transcription and processing during T-cell development. *J. Immunol.* **2012**, *188*, 3257–3267.
30. O'Connell, R.M.; Chaudhuri, A.A.; Rao, D.S.; Gibson, W.S.; Balazs, A.B.; Baltimore, D. MicroRNAs enriched in hematopoietic stem cells differentially regulate long-term hematopoietic output. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14235–14240.
31. Ooi, A.G.; Sahoo, D.; Adorno, M.; Wang, Y.; Weissman, I.L.; Park, C.Y. microRNA-125b expands hematopoietic stem cells and enriches for the lymphoid-balanced and lymphoid-biased subsets. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 21505–21510.
32. Shi, X.B.; Xue, L.; Yang, J.; Ma, A.H.; Zhao, J.; Xu, M.; Tepper, C.G.; Evans, C.P.; Kung, H.J.; deVere White, R.W. An androgen-regulated miRNA suppresses Bak1 expression and induces androgen-independent growth of prostate cancer cells. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19983–19988.
33. Surdziel, E.; Cabanski, M.; Dallmann, I.; Lyszkiewicz, M.; Krueger, A.; Ganser, A.; Scherr, M.; Eder, M. Enforced expression of miR-125b affects myelopoiesis by targeting multiple signaling pathways. *Blood* **2011**, *117*, 4338–4348.
34. Henao-Mejia, J.; Williams, A.; Goff, L.A.; Staron, M.; Licona-Limon, P.; Kaech, S.M.; Nakayama, M.; Rinn, J.L.; Flavell, R.A. The microRNA miR-181 is a critical cellular metabolic rheostat essential for NKT-cell ontogenesis and lymphocyte development and homeostasis. *Immunity* **2013**, *38*, 984–997.
35. Li, Q.J.; Chau, J.; Ebert, P.J.; Sylvester, G.; Min, H.; Liu, G.; Braich, R.; Manoharan, M.; Soutschek, J.; Skare, P.; *et al*. miR-181a is an intrinsic modulator of T-cell sensitivity and selection. *Cell* **2007**, *129*, 147–161.
36. Lashine, Y.A.; Seoudi, A.M.; Salah, S.; Abdelaziz, A.I. Expression signature of microRNA-181-a reveals its crucial role in the pathogenesis of paediatric systemic lupus erythematosus. *Clin. Exp. Rheumatol.* **2011**, *29*, 351–357.
37. Zhou, B.; Wang, S.; Mayr, C.; Bartel, D.P.; Lodish, H.F. miR-150, a microRNA expressed in mature B and T-cells, blocks early B-cell development when expressed prematurely. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7080–7085.
38. Xiao, C.; Calado, D.P.; Galler, G.; Thai, T.H.; Patterson, H.C.; Wang, J.; Rajewsky, N.; Bender, T.P.; Rajewsky, K. MiR-150 controls B-cell differentiation by targeting the transcription factor c-Myb. *Cell* **2007**, *131*, 146–159.
39. Almanza, G.; Fernandez, A.; Volinia, S.; Cortez-Gonzalez, X.; Croce, C.M.; Zanetti, M. Selected microRNAs define cell fate determination of murine central memory CD8 T-cells. *PLoS One* **2010**, *5*, e11243.
40. Lin, Y.C.; Kuo, M.W.; Yu, J.; Kuo, H.H.; Lin, R.J.; Lo, W.L.; Yu, A.L. c-Myb is an evolutionary conserved miR-150 target and miR-150/c-Myb interaction is important for embryonic development. *Mol. Biol. Evol.* **2008**, *25*, 2189–2198.

41. Zhu, J.; Paul, W.E. Peripheral CD4+ T-cell differentiation regulated by networks of cytokines and transcription factors. *Immunol. Rev.* **2010**, *238*, 247–262.
42. Bronevetsky, Y.; Villarino, A.V.; Eisley, C.J.; Barbeau, R.; Barczak, A.J.; Heinz, G.A.; Kremmer, E.; Heissmeyer, V.; McManus, M.T.; Erle, D.J.; *et al*. T-cell activation induces proteasomal degradation of Argonaute and rapid remodeling of the microRNA repertoire. *J. Exp. Med.* **2013**, *210*, 417–432.
43. Stittrich, A.B.; Haftmann, C.; Sgouroudis, E.; Kuhl, A.A.; Hegazy, A.N.; Panse, I.; Riedel, R.; Flossdorf, M.; Dong, J.; Fuhrmann, F.; *et al*. The microRNA miR-182 is induced by IL-2 and promotes clonal expansion of activated helper T lymphocytes. *Nat. Immunol.* **2010**, *11*, 1057–1062.
44. Rossi, R.L.; Rossetti, G.; Wenandy, L.; Curti, S.; Ripamonti, A.; Bonnal, R.J.; Birolo, R.S.; Moro, M.; Crosti, M.C.; Gruarin, P.; *et al*. Distinct microRNA signatures in human lymphocyte subsets and enforcement of the naive state in CD4+ T-cells by the microRNA miR-125b. *Nat. Immunol.* **2011**, *12*, 796–803.
45. Sandberg, R.; Neilson, J.R.; Sarma, A.; Sharp, P.A.; Burge, C.B. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **2008**, *320*, 1643–1647.
46. Xiao, C.; Srinivasan, L.; Calado, D.P.; Patterson, H.C.; Zhang, B.; Wang, J.; Henderson, J.M.; Kutok, J.L.; Rajewsky, K. Lymphoproliferative disease and autoimmunity in mice with increased miR-17–92 expression in lymphocytes. *Nat. Immunol.* **2008**, *9*, 405–414.
47. Mendell, J.T. miRiad roles for the miR-17–92 cluster in development and disease. *Cell* **2008**, *133*, 217–222.
48. Rao, D.S.; O'Connell, R.M.; Chaudhuri, A.A.; Garcia-Flores, Y.; Geiger, T.L.; Baltimore, D. microRNA-34a perturbs B lymphocyte development by repressing the forkhead box transcription factor Foxp1. *Immunity* **2010**, *33*, 48–59.
49. Thai, T.H.; Calado, D.P.; Casola, S.; Ansel, K.M.; Xiao, C.; Xue, Y.; Murphy, A.; Frendewey, D.; Valenzuela, D.; Kutok, J.L.; *et al*. Regulation of the germinal center response by microRNA-155. *Science* **2007**, *316*, 604–608.
50. Teng, G.; Hakimpour, P.; Landgraf, P.; Rice, A.; Tuschl, T.; Casellas, R.; Papavasiliou, F.N. microRNA-155 is a negative regulator of activation-induced cytidine deaminase. *Immunity* **2008**, *28*, 621–629.
51. Cimmino, A.; Calin, G.A.; Fabbri, M.; Iorio, M.V.; Ferracin, M.; Shimizu, M.; Wojcik, S.E.; Aqeilan, R.I.; Zupo, S.; Dono, M.; *et al*. miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13944–13949.
52. Stanczyk, J.; Pedrioli, D.M.; Brentano, F.; Sanchez-Pernaute, O.; Kolling, C.; Gay, R.E.; Detmar, M.; Gay, S.; Kyburz, D. Altered expression of microRNA in synovial fibroblasts and synovial tissue in rheumatoid arthritis. *Arthritis Rheum.* **2008**, *58*, 1001–1009.

53. Nakasa, T.; Miyaki, S.; Okubo, A.; Hashimoto, M.; Nishida, K.; Ochi, M.; Asahara, H. Expression of microRNA-146 in rheumatoid arthritis synovial tissue. *Arthritis Rheum.* **2008**, *58*, 1284–1292.
54. Pauley, K.M.; Satoh, M.; Chan, A.L.; Bubb, M.R.; Reeves, W.H.; Chan, E.K. Upregulated miR-146a expression in peripheral blood mononuclear cells from rheumatoid arthritis patients. *Arthritis Res. Ther.* **2008**, *10*, R101.
55. Nagata, S. Human autoimmune lymphoproliferative syndrome, a defect in the apoptosis-inducing Fas receptor: A lesson from the mouse model. *J. Hum. Genet.* **1998**, *43*, 2–8.
56. Wu, T.; Wieland, A.; Araki, K.; Davis, C.W.; Ye, L.; Hale, J.S.; Ahmed, R. Temporal expression of microRNA cluster miR-17–92 regulates effector and memory CD8+ T-cell differentiation. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9965–9970.
57. Takahashi, H.; Kanno, T.; Nakayamada, S.; Hirahara, K.; Sciume, G.; Muljo, S.A.; Kuchen, S.; Casellas, R.; Wei, L.; Kanno, Y.; *et al*. TGF-beta and retinoic acid induce the microRNA miR-10a, which targets Bcl-6 and constrains the plasticity of helper T-cells. *Nat. Immunol.* **2012**, *13*, 587–595.
58. Kohlhaas, S.; Garden, O.A.; Scudamore, C.; Turner, M.; Okkenhaug, K.; Vigorito, E. Cutting edge: The Foxp3 target miR-155 contributes to the development of regulatory T-cells. *J. Immunol.* **2009**, *182*, 2578–2582.
59. Lu, L.F.; Thai, T.H.; Calado, D.P.; Chaudhry, A.; Kubo, M.; Tanaka, K.; Loeb, G.B.; Lee, H.; Yoshimura, A.; Rajewsky, K.; *et al*. Foxp3-dependent microRNA155 confers competitive fitness to regulatory T-cells by targeting SOCS1 protein. *Immunity* **2009**, *30*, 80–91.
60. Yao, R.; Ma, Y.L.; Liang, W.; Li, H.H.; Ma, Z.J.; Yu, X.; Liao, Y.H. microRNA-155 modulates Treg and Th17 cells differentiation and Th17 cell function by targeting SOCS1. *PLoS One* **2012**, *7*, e46082.
61. Lu, L.F.; Boldin, M.P.; Chaudhry, A.; Lin, L.L.; Taganov, K.D.; Hanada, T.; Yoshimura, A.; Baltimore, D.; Rudensky, A.Y. Function of miR-146a in controlling Treg cell-mediated regulation of Th1 responses. *Cell* **2010**, *142*, 914–929.
62. Tsai, C.Y.; Allie, S.R.; Zhang, W.J.; Usherwood, E.J. microRNA miR-155 affects antiviral effector and effector memory CD8 T-cell differentiation. *J. Virol.* **2013**, *87*, 2348–2351.
63. Brown, C.J.; Hendrich, B.D.; Rupert, J.L.; Lafreniere, R.G.; Xing, Y.; Lawrence, J.; Willard, H.F. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **1992**, *71*, 527–542.
64. Yang, L.; Boldin, M.P.; Yu, Y.; Liu, C.S.; Ea, C.K.; Ramakrishnan, P.; Taganov, K.D.; Zhao, J.L.; Baltimore, D. miR-146a controls the resolution of T-cell responses in mice. *J. Exp. Med.* **2012**, *209*, 1655–1670.
65. Hobeika, A.C.; Morse, M.A.; Osada, T.; Peplinski, S.; Lyerly, H.K.; Clay, T.M. Depletion of human regulatory T-cells. *Methods Mol. Biol.* **2011**, *707*, 219–231.

66. Zhou, X.; Jeker, L.T.; Fife, B.T.; Zhu, S.; Anderson, M.S.; McManus, M.T.; Bluestone, J.A. Selective miRNA disruption in T reg cells leads to uncontrolled autoimmunity. *J. Exp. Med.***2008**, *205*, 1983–1991.
67. Liston, A.; Lu, L.F.; O'Carroll, D.; Tarakhovsky, A.; Rudensky, A.Y. Dicer-dependent microRNA pathway safeguards regulatory T-cell function. *J. Exp. Med.* **2008**, *205*, 1993–2004.
68. Zhan, Y.; Davey, G.M.; Graham, K.L.; Kiu, H.; Dudek, N.L.; Kay, T.; Lew, A.M. SOCS1 negatively regulates the production of Foxp3+ CD4+ T-cells in the thymus. *Immunol. Cell Biol.* **2009**, *87*, 473–480.
69. Berland, R.; Wortis, H.H. Origins and functions of B-1 cells with notes on the role of CD5. *Annu. Rev. Immunol.* **2002**, *20*, 253–300.
70. Xu, S.; Guo, K.; Zeng, Q.; Huo, J.; Lam, K.P. The RNase III enzyme Dicer is essential for germinal center B-cell formation. *Blood* **2012**, *119*, 767–776.
71. Koralov, S.B.; Muljo, S.A.; Galler, G.R.; Krek, A.; Chakraborty, T.; Kanellopoulou, C.; Jensen, K.; Cobb, B.S.; Merkenschlager, M.; Rajewsky, N.; *et al*. Dicer ablation affects antibody diversity and cell survival in the B lymphocyte lineage. *Cell* **2008**, *132*, 860–874.
72. Chen, C.Z.; Li, L.; Lodish, H.F.; Bartel, D.P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* **2004**, *303*, 83–86.
73. Chang, T.C.; Wentzel, E.A.; Kent, O.A.; Ramachandran, K.; Mullendore, M.; Lee, K.H.; Feldmann, G.; Yamakuchi, M.; Ferlito, M.; Lowenstein, C.J.; *et al*. Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. *Mol. Cell* **2007**, *26*, 745–752.
74. Bommer, G.T.; Gerin, I.; Feng, Y.; Kaczorowski, A.J.; Kuick, R.; Love, R.E.; Zhai, Y.; Giordano, T.J.; Qin, Z.S.; Moore, B.B.; *et al*. p53-mediated activation of miRNA34 candidate tumor-suppressor genes. *Curr. Biol.* **2007**, *17*, 1298–1307.
75. Vigorito, E.; Perks, K.L.; Abreu-Goodger, C.; Bunting, S.; Xiang, Z.; Kohlhaas, S.; Das, P.P.; Miska, E.A.; Rodriguez, A.; Bradley, A.; *et al*. microRNA-155 regulates the generation of immunoglobulin class-switched plasma cells. *Immunity* **2007**, *27*, 847–859.
76. Dorsett, Y.; McBride, K.M.; Jankovic, M.; Gazumyan, A.; Thai, T.H.; Robbiani, D.F.; di Virgilio, M.; Reina San-Martin, B.; Heidkamp, G.; Schwickert, T.A.; *et al*. MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. *Immunity* **2008**, *28*, 630–638.
77. Li, H.; Li, W.X.; Ding, S.W. Induction and suppression of RNA silencing by an animal virus. *Science* **2002**, *296*, 1319–1321.
78. Watanabe, T.; Tomizawa, S.; Mitsuya, K.; Totoki, Y.; Yamamoto, Y.; Kuramochi-Miyagawa, S.; Iida, N.; Hoki, Y.; Murphy, P.J.; Toyoda, A.; *et al*. Role for piRNAs and noncoding RNA in *de novo* DNA methylation of the imprinted mouse Rasgrf1 locus. *Science* **2011**, *332*, 848–852.

79. Triboulet, R.; Mari, B.; Lin, Y.L.; Chable-Bessia, C.; Bennasser, Y.; Lebrigand, K.; Cardinaud, B.; Maurin, T.; Barbry, P.; Baillat, V.; *et al*. Suppression of microRNA-silencing pathway by HIV-1 during virus replication. *Science* **2007**, *315*, 1579–1582.
80. Matskevich, A.A.; Moelling, K. Dicer is involved in protection against influenza A virus infection. *J. Gen. Virol.* **2007**, *88*, 2627–2635.
81. Otsuka, M.; Jing, Q.; Georgel, P.; New, L.; Chen, J.; Mols, J.; Kang, Y.J.; Jiang, Z.; Du, X.; Cook, R.; *et al*. Hypersusceptibility to vesicular stomatitis virus infection in Dicer1-deficient mice is due to impaired miR24 and miR93 expression. *Immunity* **2007**, *27*, 123–134.
82. Cookson, W.; Liang, L.; Abecasis, G.; Moffatt, M.; Lathrop, M. Mapping complex disease traits with global gene expression. *Nat. Rev. Genet.* **2009**, *10*, 184–194.
83. Pauley, K.M.; Cha, S.; Chan, E.K. microRNA in autoimmunity and autoimmune diseases. *J. Autoimmun.* **2009**, *32*, 189–194.
84. Furer, V.; Greenberg, J.D.; Attur, M.; Abramson, S.B.; Pillinger, M.H. The role of microRNA in rheumatoid arthritis and other autoimmune diseases. *Clin. Immunol.* **2010**, *136*, 1–15.
85. Iborra, M.; Bernuzzi, F.; Invernizzi, P.; Danese, S. microRNAs in autoimmunity and inflammatory bowel disease: crucial regulators in immune response. *Autoimmun. Rev.* **2012**, *11*, 305–314.
86. Li, J.; Wan, Y.; Guo, Q.; Zou, L.; Zhang, J.; Fang, Y.; Fu, X.; Liu, H.; Lu, L.; Wu, Y. Altered microRNA expression profile with miR-146a upregulation in CD4+ T-cells from patients with rheumatoid arthritis. *Arthritis Res. Ther.* **2010**, *12*, R81.
87. Nambiar, M.P.; Krishnan, S.; Tsokos, G.C. T-cell signaling abnormalities in human systemic lupus erythematosus. *Methods Mol. Med.* **2004**, *102*, 31–47.
88. Zhang, D.; Yang, W.; Degauque, N.; Tian, Y.; Mikita, A.; Zheng, X.X. New differentiation pathway for double-negative regulatory T-cells that regulates the magnitude of immune responses. *Blood* **2007**, *109*, 4071–4079.
89. Divekar, A.A.; Dubey, S.; Gangalum, P.R.; Singh, R.R. Dicer insufficiency and microRNA-155 overexpression in lupus regulatory T-cells: An apparent paradox in the setting of an inflammatory milieu. *J. Immunol.* **2011**, *186*, 924–930.
90. Tang, Y.; Luo, X.; Cui, H.; Ni, X.; Yuan, M.; Guo, Y.; Huang, X.; Zhou, H.; de Vries, N.; Tak, P.P.; *et al*. microRNA-146A contributes to abnormal activation of the type I interferon pathway in human lupus by targeting the key signaling proteins. *Arthritis Rheumat.* **2009**, *60*, 1065–1075.
91. Suzuki, Y.; Kim, H.W.; Ashraf, M.; Haider, H. Diazoxide potentiates mesenchymal stem cell survival via NF-kappaB-dependent miR-146a expression by targeting Fas. *Am. J. Physiol. Heart Circ. Physiol.* **2010**, *299*, H1077–H1082.
92. Yu, D.; Tan, A.H.; Hu, X.; Athanasopoulos, V.; Simpson, N.; Silva, D.G.; Hutloff, A.; Giles, K.M.; Leedman, P.J.; Lam, K.P.; *et al*. Roquin represses autoimmunity by limiting inducible T-cell co-stimulator messenger RNA. *Nature* **2007**, *450*, 299–303.

93. Ebert, P.J.; Jiang, S.; Xie, J.; Li, Q.J.; Davis, M.M. An endogenous positively selecting peptide enhances mature T-cell responses and becomes an autoantigen in the absence of microRNA miR-181a. *Nat. Immunol.* **2009**, *10*, 1162–1169.
94. Treadwell, E.L.; Alspaugh, M.A.; Sharp, G.C. Characterization of a new antigen-antibody system (Su) in patients with systemic lupus erythematosus. *Arthritis Rheumat.* **1984**, *27*, 1263–1271.
95. Satoh, M.; Langdon, J.J.; Chou, C.H.; McCauliffe, D.P.; Treadwell, E.L.; Ogasawara, T.; Hirakata, M.; Suwa, A.; Cohen, P.L.; Eisenberg, R.A.; *et al*. Characterization of the Su antigen, a macromolecular complex of 100/102 and 200-kDa proteins recognized by autoantibodies in systemic rheumatic diseases. *Clin. Immunol. Immunopathol.* **1994**, *73*, 132–141.
96. Lipovich, L.; Johnson, R.; Lin, C.Y. MacroRNA underdogs in a microRNA world: Evolutionary, regulatory, and biomedical significance of mammalian long non-protein-coding RNA. *Biochim. Biophys. Acta* **2010**, *1799*, 597–615.
97. Cabili, M.N.; Trapnell, C.; Goff, L.; Koziol, M.; Tazon-Vega, B.; Regev, A.; Rinn, J.L. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **2011**, *25*, 1915–1927.
98. Clark, M.B.; Johnston, R.L.; Inostroza-Ponta, M.; Fox, A.H.; Fortini, E.; Moscato, P.; Dinger, M.E.; Mattick, J.S. Genome-wide analysis of long noncoding RNA stability. *Genome Res.* **2012**, *22*, 885–898.
99. Zhao, J.; Sun, B.K.; Erwin, J.A.; Song, J.J.; Lee, J.T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **2008**, *322*, 750–756.
100. Lee, J.T. Epigenetic regulation by long noncoding RNAs. *Science* **2012**, *338*, 1435–1439.
101. Willingham, A.T.; Orth, A.P.; Batalov, S.; Peters, E.C.; Wen, B.G.; Aza-Blanc, P.; Hogenesch, J.B.; Schultz, P.G. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **2005**, *309*, 1570–1573.
102. Martianov, I.; Ramadass, A.; Serra Barros, A.; Chow, N.; Akoulitchev, A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* **2007**, *445*, 666–670.
103. Vigneau, S.; Rohrlich, P.S.; Brahic, M.; Bureau, J.F. Tmevpg1, a candidate gene for the control of Theiler's virus persistence, could be implicated in the regulation of gamma interferon. *J. Virol.* **2003**, *77*, 5632–5638.
104. Collier, S.P.; Collins, P.L.; Williams, C.L.; Boothby, M.R.; Aune, T.M. Cutting edge: Influence of Tmevpg1, a long intergenic noncoding RNA, on the expression of Ifng by Th1 cells. *J. Immunol.* **2012**, *189*, 2084–2088.
105. Gomez, J.A.; Wapinski, O.L.; Yang, Y.W.; Bureau, J.F.; Gopinath, S.; Monack, D.M.; Chang, H.Y.; Brahic, M.; Kirkegaard, K. The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon-gamma locus. *Cell* **2013**, *152*, 743–754.

106. Calin, G.A.; Liu, C.G.; Ferracin, M.; Hyslop, T.; Spizzo, R.; Sevignani, C.; Fabbri, M.; Cimmino, A.; Lee, E.J.; Wojcik, S.E.; *et al*. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. *Cancer Cell* **2007**, *12*, 215–229.
107. Calin, G.A.; Ferracin, M.; Cimmino, A.; di Leva, G.; Shimizu, M.; Wojcik, S.E.; Iorio, M.V.; Visone, R.; Sever, N.I.; Fabbri, M.; *et al*. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N. Engl. J. Med.* **2005**, *353*, 1793–1801.
108. Chen, X.; Ba, Y.; Ma, L.; Cai, X.; Yin, Y.; Wang, K.; Guo, J.; Zhang, Y.; Chen, J.; Guo, X.; *et al.* Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res.* **2008**, *18*, 997–1006.
109. Lawrie, C.H.; Gal, S.; Dunlop, H.M.; Pushkaran, B.; Liggins, A.P.; Pulford, K.; Banham, A.H.; Pezzella, F.; Boultwood, J.; Wainscoat, J.S.; *et al*. Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. *Br. J. Haematol.* **2008**, *141*, 672–675.
110. Park, N.J.; Zhou, H.; Elashoff, D.; Henson, B.S.; Kastratovic, D.A.; Abemayor, E.; Wong, D.T. Salivary microRNA: Discovery, characterization, and clinical utility for oral cancer detection. *Clin. Cancer Res.* **2009**, *15*, 5473–5477.
111. Kosaka, N.; Izumi, H.; Sekine, K.; Ochiya, T. microRNA as a new immune-regulatory agent in breast milk. *Silence* **2010**, *1*, 7.
112. Mitchell, P.S.; Parkin, R.K.; Kroh, E.M.; Fritz, B.R.; Wyman, S.K.; Pogosova-Agadjanyan, E.L.; Peterson, A.; Noteboom, J.; O'Briant, K.C.; Allen, A.; *et al*. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10513–10518.
113. Ji, X.; Takahashi, R.; Hiura, Y.; Hirokawa, G.; Fukushima, Y.; Iwai, N. Plasma miR-208 as a biomarker of myocardial injury. *Clin. Chem.* **2009**, *55*, 1944–1949.
114. Arroyo, J.D.; Chevillet, J.R.; Kroh, E.M.; Ruf, I.K.; Pritchard, C.C.; Gibson, D.F.; Mitchell, P.S.; Bennett, C.F.; Pogosova-Agadjanyan, E.L.; Stirewalt, D.L.; *et al*. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 5003–5008.
115. Valadi, H.; Ekstrom, K.; Bossios, A.; Sjostrand, M.; Lee, J.J.; Lotvall, J.O. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **2007**, *9*, 654–659.
116. Zernecke, A.; Bidzhekov, K.; Noels, H.; Shagdarsuren, E.; Gan, L.; Denecke, B.; Hristov, M.; Koppel, T.; Jahantigh, M.N.; Lutgens, E.; *et al*. Delivery of microRNA-126 by apoptotic bodies induces CXCL12-dependent vascular protection. *Sci. Signal.* **2009**, *2*, ra81.
117. Zhang, Y.; Liu, D.; Chen, X.; Li, J.; Li, L.; Bian, Z.; Sun, F.; Lu, J.; Yin, Y.; Cai, X.; *et al.* Secreted monocytic miR-150 enhances targeted endothelial cell migration. *Mol. Cell* **2010**, *39*, 133–144.

118. Bhatnagar, S.; Schorey, J.S. Exosomes released from infected macrophages contain *Mycobacterium avium* glycopeptidolipids and are proinflammatory. *J. Biol. Chem.* **2007**, *282*, 25779–25789.
119. Cocucci, E.; Racchetti, G.; Meldolesi, J. Shedding microvesicles: Artefacts no more. *Trends Cell Biol.* **2009**, *19*, 43–51.
120. Hunter, M.P.; Ismail, N.; Zhang, X.; Aguda, B.D.; Lee, E.J.; Yu, L.; Xiao, T.; Schafer, J.; Lee, M.L.; Schmittgen, T.D.; *et al*. Detection of microRNA expression in human peripheral blood microvesicles. *PLoS One* **2008**, *3*, e3694.
121. Collino, F.; Deregibus, M.C.; Bruno, S.; Sterpone, L.; Aghemo, G.; Viltono, L.; Tetta, C.; Camussi, G. Microvesicles derived from adult human bone marrow and tissue specific mesenchymal stem cells shuttle selected pattern of miRNAs. *PLoS One* **2010**, *5*, e11803.
122. Muller, G.; Schneider, M.; Biemer-Daub, G.; Wied, S. Microvesicles released from rat adipocytes and harboring glycosylphosphatidylinositol-anchored proteins transfer RNA stimulating lipid synthesis. *Cell. Signal.* **2011**, *23*, 1207–1223.
123. Nolte-'t Hoen, E.N.; Buermans, H.P.; Waasdorp, M.; Stoorvogel, W.; Wauben, M.H.; 't Hoen, P.A. Deep sequencing of RNA from immune cell-derived vesicles uncovers the selective incorporation of small non-coding RNA biotypes with potential regulatory functions. *Nucleic Acids Res.* **2012**, *40*, 9272–9285.
124. Mittelbrunn, M.; Gutierrez-Vazquez, C.; Villarroya-Beltri, C.; Gonzalez, S.; Sanchez-Cabo, F.; Gonzalez, M.A.; Bernad, A.; Sanchez-Madrid, F. Unidirectional transfer of microRNA-loaded exosomes from T-cells to antigen-presenting cells. *Nat. Commun.* **2011**, *2*, 282.
125. Monks, C.R.; Freiberg, B.A.; Kupfer, H.; Sciaky, N.; Kupfer, A. Three-dimensional segregation of supramolecular activation clusters in T-cells. *Nature* **1998**, *395*, 82–86.

Reprinted from *IJMS*. Cite as: Yu, S.; Zhang, R.; Zhu, C.; Cheng, J.; Wang, H.; Wu, J. MicroRNA-143 Downregulates Interleukin-13 Receptor Alpha1 in Human Mast Cells. *Int. J. Mol. Sci.* **2013**, *14*, 16958-16969.

Article

MicroRNA-143 Downregulates Interleukin-13 Receptor Alpha1 in Human Mast Cells

Shaoqing Yu [1,2], Ruxin Zhang [1,*], Chunshen Zhu [3], Jianqiu Cheng [3], Hong Wang [1] and Jing Wu [1]

[1] Department of Otolaryngology, Huadong Hospital, Fudan University, Shanghai 200040, China; E-Mails: yu_shaoqing@163.com (S.Y.); whhb1984@163.com (H.W.); dushixingniu@163.com (J.W.)

[2] Department of Otolaryngology, Tongji Hospital, Tongji University, Shanghai 200065, China

[3] Department of Otolaryngology, Jinan General Hospital of PLA, Jinan 250031, Shandong, China; E-Mails: zhucs1@sina.cn (C.Z.); showing100@163.com (J.C.)

* Author to whom correspondence should be addressed; E-Mail: nosezhang@163.com; Tel./Fax: +86-021-6611-1041.

Received: 3 June 2013; in revised form: 8 July 2013 / Accepted: 2 August 2013 / Published: 19 August 2013

Abstract: MicroRNA-143 (miR-143) was found to be downregulated in allergic rhinitis, and bioinformatics analysis predicted that IL-13Rα1 was a target gene of miR-143. To understand the molecular mechanisms of miR-143 involved in the pathogenesis of allergic inflammation, recombinant miR-143 plasmid vectors were constructed, and human mast cell-1(HMC-1) cells which play a central role in the allergic response were used for study. The plasmids were transfected into HMC-1 cells using a lentiviral vector. Expression of IL-13Rα1 mRNA was then detected by reverse transcriptase polymerase chain reaction (RT-PCR) and Western Blotting. The miR-143 lentiviral vector was successfully stably transfected in HMC-1 cells for target gene expression. Compared to the control, the target gene IL-13Rα1 was less expressed in HMC-1 transfected with miR-143 as determined by RT-PCR and Western Blotting ($p < 0.05$); this difference in expression was statistically significant and the inhibition efficiency was 71%. It indicates that miR-143 directly targets IL-13Rα1 and suppresses IL-13Rα1 expression in HMC-1 cells. Therefore, miR-143 may be associated with allergic reaction in human mast cells.

Key words: microRNA-143; IL-13Rα1; mast cell; lentiviral vector; allergy

1. Introduction

Allergy is a chronic inflammatory condition and many inflammatory factors, such as IL-4, IL-13, IL-8, *etc*. have been related to this disorder [1,2]; however, how these factors regulate mechanisms responsible for allergic reactions has not yet been thoroughly elucidated. Recently, it has been found that microRNAs (miRNAs) inhibit protein translation to regulate gene expression; miRAs are small RNAs produced by cells through a unique process, involving RNA polymerase II, microprocessor protein complex, and the RNAase III/Dicer endonuclease complex. miRNA ribonucleoprotein complex is formed at the end of this process [3]. The biological functions of miRNAs depend on their ability to silence gene expression, primarily through degradation of the target mRNA and/or translational suppression mediated by the RNA-induced silencing complex (RISC) [3]. The function of miRNAs in gene regulation has been investigated in a variety of diseases including allergic disease.

In our previous study, miRNA profiles and RT-PCR were used to reveal the difference between nasal mucosa biopsies of patients with allergic rhinitis (AR) and those of healthy volunteers. When compared with normal tissue, microRNA-143 (miR-143) was the most significantly downregulated miRNA in nasal mucosa of tissues exhibiting AR. It was found that expression of miR-143 in smooth muscle cells of airways triggers allergic conditions such as asthma [4]. However, we have not yet been able to understand how miR-143 regulates allergic inflammation in upper airways such as AR.

To study the mechanism of miR-143 regulation in allergic inflammation, we investigated the target genes of miR-143 through the TargetScan procedure (http://genes.mit.edu/tscan/targetscanS.html and http://pictar.mdc-berlin.de/) to predict miR-143 target genes, and found that the IL-13Rα1 gene 3' UTR had 15 sequential pairing bases with miR-143 (Table 1), which indicates that IL-13Rα1, known to play an important role in allergy, may be a potential target of miR-143. To investigate this, we used mast cells, which are an important AR target cell and built the miR-143 overexpression system. This system was transfected into mast cells to observe changes induced through target gene expression. The aim of this study was thus to identify the effects of miR-143 on IL-13Rα1 in mast cells to further an understanding of the molecular mechanisms of allergic inflammation.

Table 1. Target prediction of IL-13Rα1 for Has-miR-143 based on conservation.

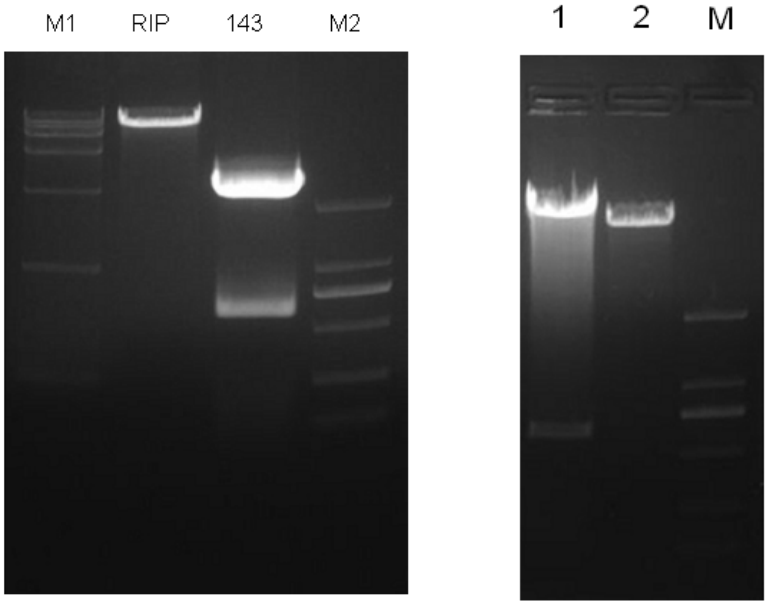
Nuclei mapped to alignments	Nuclei mapped to sequence	Structure of predicted duplex	Probabilities	Free energies kcal/mol
1397	1338	_GAG__GCAG__G_____CAUCUCA_: _CUC__UGUC__C_____GUAGAGU_	0.30	−24.1

2. Results

2.1. Construction and Identification of miR-143 pLenO-RIP Plasmid

We connected the target gene DNA fragment into pLenO-RIP empty vector (Figure 1a), and transformed the plasmid ligation product into *E. coli* DH5. Subsequently, we extracted the plasmids from amplified DH5. An approximately 600 bp DNA fragment was identified from the recombinant plasmid (Figure 1b). Finally, DNA sequencing confirmed the recombinant plasmid DNA sequence.

Figure 1. Images of gel of plasmid digest of pLenO-RIP, PUC-miR-143 and recombinant plasmid. Lanes 1–4 (from left to right) were products of DNA marker (DL15000), pLenO-RIP, PUC-miR-143 and DNA marker (DL2000), respectively (**a**). From left to right were positive clones, negative control and marker (DL2000) respectively. The approximate 600bp size of the product can be found in positive clones (**b**).



(**a**) (**b**)

2.2. Lentiviral Vector Packaging and Identification

The lentiviral packaging systems, comprising four kinds of plasmid DNA solutions, were co-transfected into 293T cells. Cells grew well and strong fluorescence intensity was observed under the fluorescent microscope, indicating that the virus packaging was successful (Figure 2). Forty-eight hours after transfection, supernatant viral material was collected and concentrated, then a 10-fold dilution was transformed into a 100–10^{-5} concentration gradient to infect 293T cells, and the supernatant viral material was collected. 48 h later, the virus titer was determined via the virus titer formula (the Tu/mL) = (GFP − positive cell count × virus supernatant dilution factor)/0.1. The production of virus droplets was 2.9×10^8 of TU/mL, and they were then prepared for transfection of HMC-1 cells.

Forty-eight hours after developing HMC-1 cell culture, the lentiviral vector of miR-143 was transfected into cell culture, and RFP blank lentivirus was used as the control. Approximately 48 h–72 h after infection, the cell culture was found in good condition under an inverted fluorescence microscope, wherein the infection efficiency was up to 80% (Figure 3). The results indicate that the miR-143 lentiviral vector was successfully transfected into HMC-1 cells, and that it expressed the relevant target genes in a stable manner.

Figure 2. Different concentrations of virus were co-transfected into 293T cells: 0.1 μL (**a**), 0.1 μL (**b**) and 0.01 μL (**c**). Cells grew well and strong fluorescence intensity was observed under the fluorescent microscope, suggesting that the virus packaging was successful.

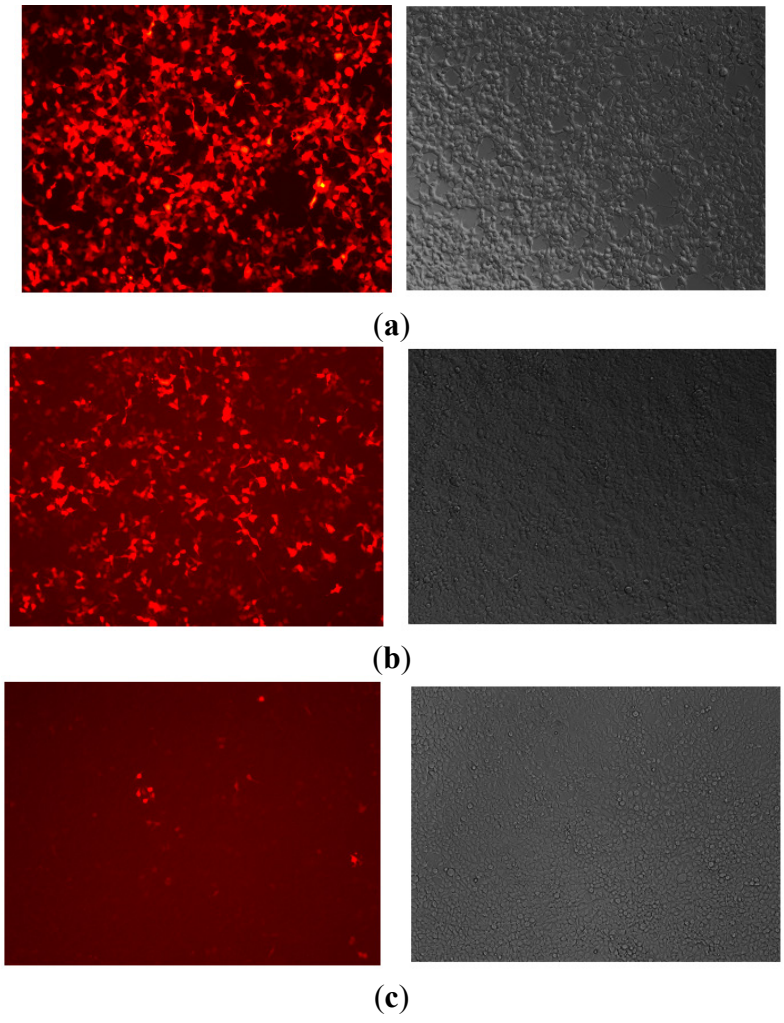
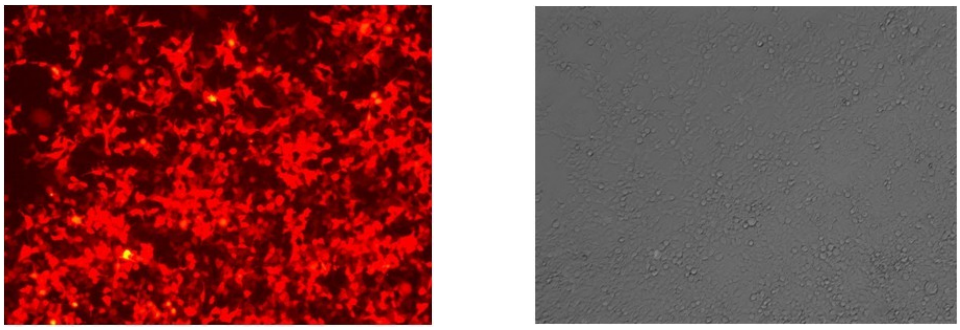


Figure 3. miR-143 transfected HMC-1 cells (**a**) fluorescent vision and (**b**) normal vision. 36 h after transfection, the RFP red fluorescent protein gene were seen expressed in up to 80% cells under fluorescent vision, it showed that the target genes were successfully expressed in the cells. (original magnification ×40)

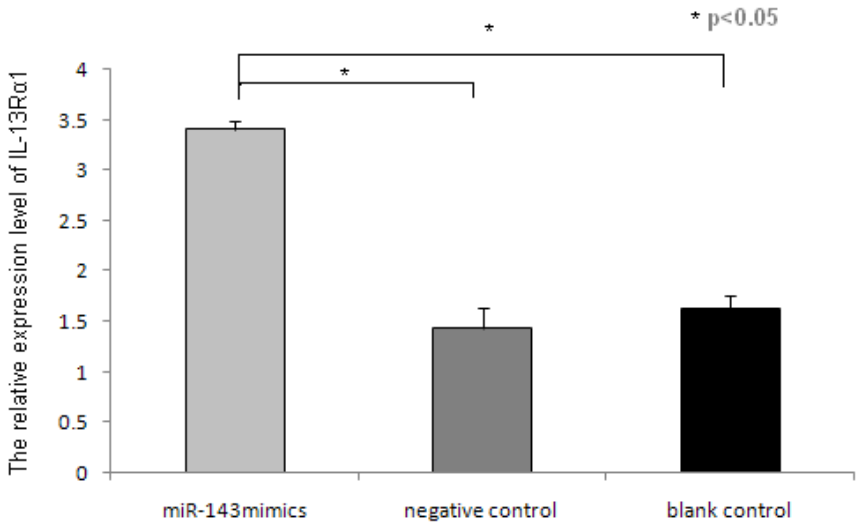


(**a**) (**b**)

2.3. miR-143 Suppresses IL-13Rα1 Expression in HMC-1 Cells

The cumulative data for mRNA expression of IL-13Rα1 are presented in Figure 4. Compared with the control, IL-13Rα1 mRNA expression was significantly downregulated in miR-143 transfected cells ($p > 0.05$). No significant changes of IL-13Rα1 mRNA levels were observed in HMC-1 cells and empty vector transfected cells ($p > 0.05$).

Figure 4. Expression of target gene IL-13Rα1 in different groups by RT-PCR. For the three groups: miR-143 mimics transfected cells, negative control group (empty vector transfected cells) and blank control group (HMC-1 cells), and every experiment was conducted in triplicate. The vertical axis represents the relative expression level of IL-13Rα1 control to 18sRNA (ΔCt). With the same cycle number, among them, miR-143 transfected cells had less amplification of IL-13Rα1 control to 18sRNA (* $p < 0.05$ for each).

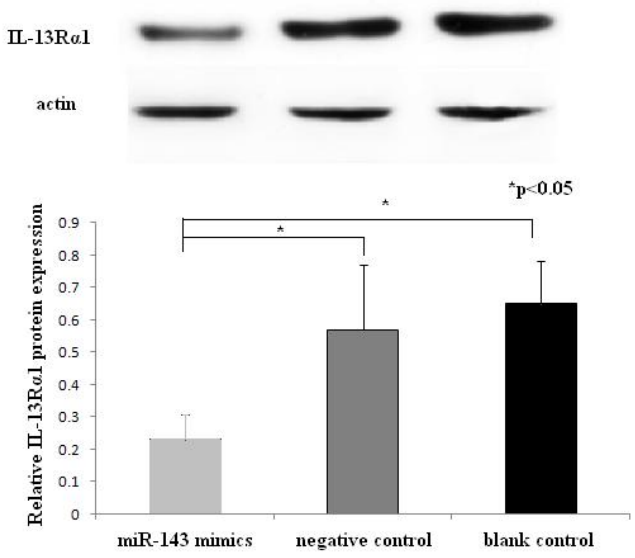


As shown in Figure 4, the expression of IL-13Rα1 in the miR-143 transfected HMC-1 cells group was reported to be 3.41; this value is relative to the internal reference 18sRNA. Furthermore, the expression of IL-13Rα1 in untreated controls HMC-1 and empty vector transfected control group were 1.63 and 1.43, respectively. The results indicated there was significantly reduced expression of IL-13Rα1 when miR-143 is overexpressed in HMC-1 cells, with an inhibition efficiency of 71%. This indicates that IL-13Rα1 is a target gene of miR-143 in HMC-1 cells, and that miR-143 can significantly inhibit the target gene IL-13Rα1 expression in HMC-1 cells.

2.4. Western Blotting Results of IL-13Rα1 Downregulation in miR-143 Transfected HMC-1 Cells

Western blot analysis showed that the expression of IL-13Rα1 protein in the miR-143 group was much lower than in the negative and blank control groups ($p < 0.05$ for each). There was no significant difference in the expression of IL-13Rα1 protein between the negative control and blank control group ($p > 0.05$). As shown in Figure 5, the over-expression of miR-143 caused reduction in IL-13Rα1 protein expression in HMC-1 cells.

Figure 5. Western blot analysis of IL-13Rα1 protein 48 h after transfection. For the three groups: miR-143 mimics, negative control group and blank control group. Experiments were done in duplicate. Western blot analysis showed that the relative expression of IL-13Rα1 protein in the miR-143 group was much lower than in negative and blank control groups (***** $p < 0.05$ for each).



3. Discussion

MiRNAs are short (20–24 nt), non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms, because they affect both the stability and translation of mRNAs [5]. In mammals, thousands of miRNAs that perform diverse functions have been identified.

Among them, miR-143 was found to be associated with allergic rhinitis in our research [6], but the mechanism of its action was unclear. In order to study this mechanism, we determined the target genes of miR-143 and investigated related mechanisms. The target genes of miR-143 were screened using miRanda (http://www.microrna.org) software. It was found that miR-143 may target multiple genes, including NOVA1, ZIC3 and MARCKS etc. Among these target genes, IL-13α1 has been the focus of allergic inflammation-associated research in recent years [7]. miR-143 caused allergic inflammation, which might be associated with the target gene IL-13α1. Therefore, we carried out this microRNA intervention project.

To determine the gene regulation of miRNA, methods of overexpression and suppression of miRNAs are often used in research studies focused on this subject [8]. In this study, the overexpression of miR-143 was chosen for exploring the regulation of miR-143 on genes of inflammatory cytokines in allergic inflammation. The precursor sequence of miR-143 was synthesized directly, and then this sequence amplified genes using the primer extension method. The product was digested and inserted into the lentivirus expression vector; thus lentiviral packaging plasmids were used to transfect 293T cells. The results indicated that the miR-143 lentiviral expression vector was successfully built which could be perfomed for the next step.

Lentiviral transduction vectora were used in this study because they have the ability to infect different types of cells. In addition, lentiviral transduction vectors can also carry exogenous genes and integrate them into the genome of the host cell for the purpose of achieving long-term stable expression [9]. It will not encode viral proteins, so it can be used for miRNA overexpression transfection experiments [10]. The viral vector carrying a red fluorescent protein gene RFP can be transfected into HMC-1 by lentiviral vectors [11]. The high expression of RFP indicates the high efficiency of miR-143 transfection. Moreover, the sustained expression of RFP in HMC-1 also showed that there was long-term overexpression of miR-143 in cells.

It has been reported that miR-143 plays an important role in regulating smooth muscle cell (SMC) fate and plasticity. Thomas Boettger's study revealed that the expression of miR-143, which formed a small cluster on mouse chromosome 18, strongly correlated with the number of SMCs. Vascular smooth muscle cells (VSMCs) from miR-143-deficient mice were locked in the synthetic state, which incapacitated their contractile abilities and favored neo-intimal lesion development, thereby revealing an unanticipated role of miR-143 in the regulation of VSMC phenotype [12]. Pleiotropic cytokines, IFN-beta and IFN-gamma, could stimulate miR-143 expression of smooth muscle cell in airways, contributing to airway allergic diseases such as asthma [13]. Some studies also demonstrated that miR-143 deficiency is associated not only with altered vasocontraction but also with impaired vasodilation [14]. Furthermore, some authors showed that miR-143 is a critical regulator of cell cycle activity in stem cells [15]. Recently it was discovered that miR-143 has tumor suppressor activity [16], but the role of miR-143 in allergic inflammation, especially in functional cells, such as mast cell, EOS, and T cell has been little investigated. This is the first study to illustrate that the allergic reaction in mast cells was attributable to inhibition of IL-13Rα1 by miR-143.

IL-13 is an immunoregulatory cytokine predominantly secreted by activated Th2 cells. Over the past several years, it has become evident that IL-13 is a key mediator in the pathogenesis of allergic inflammation. Like IL-4, IL-13 responds by signaling through the T-cell antigen receptor and mast cells. IL-13 responds through basophils when there is a cross-linkage of the high-affinity receptor for IgE. It is also produced by activated eosinophils. IL-13 plays a pivotal role in IgE-dependent inflammatory reactions, and it acts on B cells to produce IgE. IL-13 has been implicated in a variety of allergic responses, including airway hypersensitivity, mucus hypersecretion, AR, and asthma. IL-13 induces many of the same responses as IL-4 and shares a receptor subunit with IL-4 [17,18].

IL-13 mediates its effects by interacting with a complex receptor system comprised of IL-4Rα and two IL-13 binding proteins: IL-13Rα1 and IL-13Rα2. The expression of IL-13 receptors has been detected on human B cells, basophils, eosinophils, mast cells, endothelial cells, fibroblasts, monocytes, macrophages, respiratory epithelial cells, and smooth muscle cells. Human IL-13Rα1 is a single gene on chromosome Xq13 [19]. It binds IL-13 with low affinity, and also combines with IL-4Rα to form a high-affinity IL-13 binding complex. A variety of cells, including B cells, basophils, eosinophils, mast cells, endothelial cells, fibroblasts, monocytes, epithelial cells, and smooth muscle cells, respond to IL-13 owing to this binding complex.

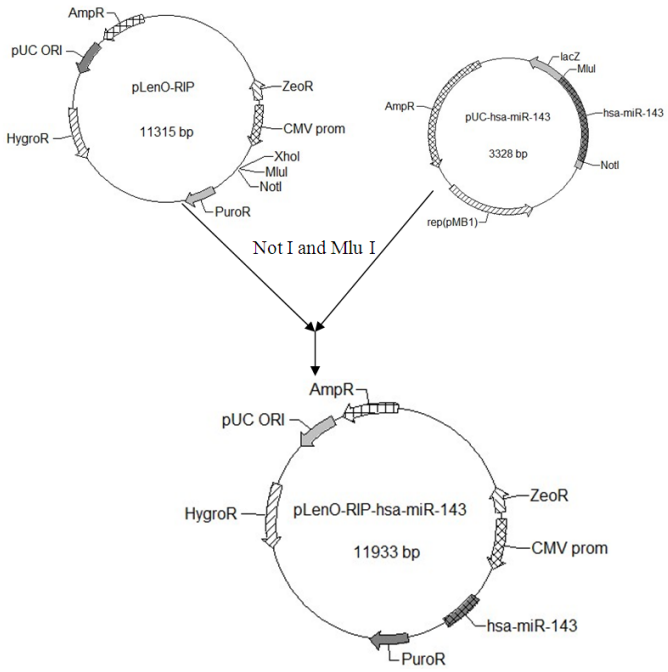
Some studies have shown that human mast cell can express IL-13Rα1 and can be activated by IL-13 [20]. In this study, the overexpression of miR-143 can suppress the expression of IL-13Rα1 in mast cells, which may be stimulated by combined IL-13/IL-4. This ultimately translates into minimizing the allergic response. Thus, the role of miR-143 in allergic response may be associated with regulation of IL-13 pathway. Nevertheless, further investigation is needed to validate these inferences.

4. Experimental Section

4.1. Construction of miR-143 Target Sequence-Luciferase Reporter Plasmid

Using the miR-143 target sequence obtained from Ensembl, we designed its pre-miRNA sequence (GCGCAGCGCCCUGUCUCCCAGCCUGAGGUGCAGUGCUGCAUCUCUGGUCAGUUGGGAG UCUGAGAUGAAGCACUGUAGCUCAGGAAGAGAGAAGUUGUUCUGCAGC). We extended both sides of the sequence of pre-miRNA about 200 bp as the pri-miRNA. The sequence with an over-hanging Mlu I site and a Not I site was synthesized by Genscript (Shanghai, China) and digested to obtain the target gene fragment. The miR 143-coding genomic DNA fragment was cloned down stream from the RFP gene of the lentiviral vector pLenO-RIP by digesting with Not I and Mlu I (Figure 6). The constructs were confirmed by DNA sequencing.

Figure 6. Construction of miR-143 target sequence-luciferase reporter plasmid.



4.2. Lentivirus Production and Transfection

Lentivirus vector was used to transfect genes. This virus packaging system was composed of pRsv-REV, pMDlg-pRRE pMD2G, and interference plasmid. With the help of Lectivirus vector, we could determine the gene expression plasmid, the (pLenO-RIP + miR-143), which also containing the red fluorescent protein (RFP). The production of lentiviral vector pLenO-RIP + miR-143 was performed by simultaneously delivering lentiviral transfer vectors and packaging plasmids (pRsv-REV, pMDlg-pRRE and pMD2.G) into 293 T cells.

Pseudo-viral particles that were generated by 293 T cells within 48 h were centrifuged at 100,000× *g* for 2 h and frozen at −70 °C for future experiments. 293 T cells were seeded into 24-well plates, and lentivirus was used to transduce for 48 h with a multiplicity of infection (MOI) of 10.

4.3. Cell Culture

The HMC-1 cells were cultured at 37 °C in an atmosphere of 5% CO_2. In this case, HMC-1 cells were cultured in DMEM media supplemented with 10% FBS, and the media were replaced routinely. The cells were subjected to transfection when the cell density was about 60%–70%.

4.4. Transfection and Luciferase Assays

To eliminate the interference of transfection reagent, cell lines were seeded into 24-well plates and set up as follows: (a) miR-143 mimics group, in which the cells were transfected with lentivirus at a multiplicity of infection (MOI) of 10 for 48 h; (b) negative control group, in which the cells were

subjected to transfection with lentiviral without miR-143; and (c) blank control group, in which the HMC-1 cells were cultured routinely without any treatment. All cells of groups were cultured in normal conditions. The cells were subjected to transfection when the cell density was about 60%–70%. All transfections were performed in quadruplicate and repeated in at least three independent experiments.

After transfection, cell count was determined at 24 h, 48 h and 72 h, respectively for the purpose of determining cell proliferation. The expression of RFP was observed as red fluorescent protein under fluorescent microscope. Moreover, RNA and protein were extracted after transfection at 48 h or 72 h, respectively.

4.5. RNA Isolation and Reverse Transcription and Polymerase Chain Reaction (RT-PCR) for miR-143-Target Gene Expression

Total cellular RNA was extracted from 1×10^6 cells using TrizolTM reagent (Invitrogen Inc., Carlsbad, California, MD, USA) according to the manufacturer's protocol. The concentration and purity of RNA were determined spectrophotometrically. Then, the synthesis of cDNA was performed according to a cDNA cycle kit K1621 (Fermentas Inc., Burlington Canada). To determine the expression of the target gene of miR-143, IL-13Rα1, we performed fluorescent quantitative real time RT-PCR assay. The sequences of the primers (TaKaRa Inc., Dalian, China) specific for IL-13Rα1 were performed with sense (ACCCGAGGGAGCCAGCTCAA) and antisense (GGTGCTACACTGGGACCCCACT) primers, wherein the expected size of the amplified sequence was 111 bp. 18sRNA was used as control. Then, the incubation of cDNA and primer was performed at 95 °C for 15 s, and the PCR reaction proceeded for 45 cycles as per the following conditions: 95 °C for 15 s and 60 °C for 30 s, in a programmable thermal cycler (Eppendorf realplex.2s, Eppendorf Co., Ltd., Hamburger, Germany) using a thermostable Taq DNA polymerase (SYBR premix ex taq, TaKaRa Inc., Dalian, China). Experiments were done in triplicate. For each sample, the amounts of the target and an endogenous control (18sRNA) were determined. To obtain a normalized target value, the amount of the target molecule was divided by the amount of the endogenous reference.

4.6. Western Blotting Analysis

Forty-eight hours after transfection, HMC-1 cells were harvested and centrifuged, and total protein was extracted. Protein concentrations were determined using the BCA protein assay. After heated for 10 min at 100 °C, 20 g of denatured protein was subjected to 10% SDS-PAGE. Then proteins were transferred electrophoretically for 1 h at 200 mA at 4 °C onto PVDF membranes. Membranes were blocked for 1 h at room temperature in TBS containing 5% non-fat dry milk. Blots were washed 3 times for 10 min each with 0.1% TBS-T and subsequently treated overnight at 4 °C with primary antibodies against IL-13Rα1 and actin (1:500). After washing 3 times for 10 min each with 0.1% TBS-T, the blots were incubated with anti-mouse antibody (1:5000) conjugated with horseradish peroxidase for 1 h at room temperature. Bands were visualized by using EZ-ECL detection reagents. The scanned

images were quantified using Quantity One software. Actin used as an endogenous protein for normalization. Experiments were done in duplicate. The ratio of IL-13Rα1/actin was used for semi-quantification and comparison between different groups.

4.7. Statistical Analysis

All data were analyzed by SPSS10.0 software (SPSS, Chicago, IL, USA, 2000). Microarray data were analyzed by Cluster Analysis. All data were represented by mean values ± standard deviation. P values less than 0.05 were considered statistically significant. Independent simple T test was used to compare the results of different groups.

5. Conclusions

We successfully constructed the lentiviral miR-143 overexpression vector and transfected it into human mast cell. In this study, we also identified IL-13Rα1 as a target gene of miR-143, and that it was inhibited by overexpression of miR-143 in mast cells. All these findings indicate that miR-143 may play important roles in triggering allergic inflammation.

Acknowledgments

This study was supported by a grant from the National Science Foundation of China (No. 81170893) and Postdoctoral Science Foundation of China (No. 20100470108).

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

1. Meltzer, E.O.; Nathan, R.; Derebery, J.; Stang, P.E.; Campbell, U.B.; Yeh, W.S.; Corrao, M.; Stanford, R. Sleep, quality of life, and productivity impact of nasal symptoms in the United States: Findings from the burden of rhinitis in America survey. *Allergy Asthma Proc*. **2009**, *30*, 244–254.
2. Blaiss, M.S. Allergic rhinitis: Direct and indirect costs. *Allergy Asthma Proc.* **2010**, *31*, 375–380.
3. Lee, Y.; Ahn, C.; Han, J.; Choi, H.; Kim, J.; Yim, J.; Lee, J.; Provost, P.; Rådmark, O.; Kim, S.; *et al*. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **2003**, *425*, 415–419.
4. Ji, X.; Li, J.; Xian, X. microRNAs: Potential regulators of airway smooth muscle cell plasticity involved in asthma-induced airway remodeling. *Asian Biomed.* **2013**, *7*, 3–14.
5. Farh, K.K.; Grimson, A.; Jan, C.; Lewis, B.P.; Johnston, W.K.; Lim, L.P.; Burge, C.B.; Bartel, D.P. The widespread impact of mammalian microRNAs on mRNA repression and evolution. *Science* **2005**, *310*, 1817–1821.

6. Yu, S.; Zhang, R.; Liu, G.; Yan, Z.; Hu, H.; Yu, S.; Zhang, J. Microarray analysis of differentially expressed microRNAs in allergic rhinitis. *Am. J. Rhinol. Allergy* **2011**, *25*, e242–e246.
7. Oh, C.K.; Geba, G.P.; Molfino, N. Investigational therapeutics targeting the IL-4/IL-13/STAT-6 pathway for the treatment of asthma. *Eur. Respir. Rev*. **2010**, *19*, 46–54.
8. Zhang, X.; Graves, P.R.; Zeng, Y. Stable Argonaute2 overexpression differentially regulates microRNA production. *Biochim. Biophys. Acta* **2009**, *1789*, 153–159.
9. Pfeifer, A.; Hofmann, A. Lentiviral transgenesis. *Methods Mol. Biol.* **2009**, *530*, 391–405.
10. Loewen, N.; Poeschla, E.M. Lentiviral vectors. *Adv. Biochem. Eng. Biotechnol.* **2005**, *99*, 169–191.
11. Scherr, M.; Battmer, K.; Ganser, A.; Eder, M. Modulation of gene expression by lentiviral-mediated delivery of small interfering RNA. *Cell Cycle* **2003**, *2*, 251–257.
12. Elia, L.; Quintavalle, M.; Zhang, J.; Contu, R.; Cossu, L.; Latronico, M.V.; Peterson, K.L.; Indolfi, C.; Catalucci, D.; Chen, J.; *et al*. The knockout of miR-143 and -145 alters smooth muscle cell maintenance and vascular homeostasis in mice: correlates with human disease. *Cell Death Differ*. **2009**,*16*, 1590–1598.
13. Boettger, T.; Beetz, N.; Kostin, S.; Schneider, J.; Krüger, M.; Hein, L.; Braun, T. Acquisition of the contractile phenotype by murine arterial smooth muscle cells depends on the miR143/145 gene cluster. *J. Clin. Invest.* **2009**, *119*, 2634–2647.
14. Norata, G.D.; Pinna, C.; Zappella, F.; Elia, L.; Sala, A.; Condorelli, G.; Catapano, A.L. microRNA 143–145 deficiency impairs vascular function. *Int. J. Immunopathol. Pharmacol*. **2012**, *25*, 467–474.
15. Lai, V.K.; Ashraf, M.; Jiang, S.; Haider, K. microRNA-143 is a critical regulator of cell cycle activity in stem cells with co-overexpression of Akt and angiopoietin-1 via transcriptional regulation of Erk5/cyclin D1 signaling. *Cell Cycle* **2012**, *11*, 767–777.
16. Ugras, S.; Brill, E.; Jacobsen, A.; Hafner, M.; Socci, N.D.; Decarolis, P.L.; Khanin, R.; O'Connor , R.; Mihailovic, A.; Taylor, B.S.; *et al*. Small RNA sequencing and functional characterization reveals microRNA-143 tumor suppressor activity in liposarcoma. *Cancer Res*. **2011**, *71*, 5659–5669.
17. Kelly-Welch, A.; Hanson, E.M.; Keegan, A.D. Interleukin-13 (IL-13) pathway. *Sci. STKE* **2005**, *2005*, cm8.
18. Munitz, A.; Brandt, E.B.; Mingler, M.; Finkelman, F.D.; Rothenberg, M.E. Distinct roles for IL-13 and IL-4 via IL-13 receptor alpha1 and the type II IL-4 receptor in asthma pathogenesis. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7240–7245.
19. Guo, J.; Apiou, F.; Mellerin, M.P.; Lebeau, B.; Jacques, Y.; Minvielle, S. Chromosome mapping and expression of the human interleukin-13 receptor. *Genomics* **1997**, *42*, 141–145.
20. Kaur, D.; Hollins, F.; Woodman, L.; Yang, W.; Monk, P.; May, R.; Bradding, P.; Brightling, C.E. Mast cells express IL-13R alpha 1: IL-13 promotes human lung mast cell proliferation and Fc epsilon RI expression. *Allergy* **2006**, *61*, 1047–1053.

Reprinted from *IJMS*. Cite as: Kozlowska, E.; Krzyzosiak, W.J.; Koscianska, E. Regulation of Huntingtin Gene Expression by miRNA-137, -214, -148a, and Their Respective isomiRs. *Int. J. Mol. Sci.* **2013**, *14*, 16999-17016.

Article

Regulation of Huntingtin Gene Expression by miRNA-137, -214, -148a, and Their Respective isomiRs

Emilia Kozlowska, Wlodzimierz J. Krzyzosiak * and Edyta Koscianska *

Department of Molecular Biomedicine, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14 Str., 61-704 Poznan, Poland; E-Mail: emiliak@ibch.poznan.pl

* Authors to whom correspondence should be addressed;
E-Mails: wlodkrzy@ibch.poznan.pl (W.J.K.); edytak@ibch.poznan.pl (E.K.);
Tel.: +48-61-852-8503, Fax: +48-61-852-0532.

Received: 30 May 2013; in revised form: 6 August 2013 / Accepted: 8 August 2013 / Published: 19 August 2013

Abstract: With the advent of deep sequencing technology, a variety of miRNA length and sequence variants, termed isomiRNAs (isomiRs), have been discovered. However, the functional roles of these commonly detected isomiRs remain unknown. In this paper, we demonstrated that miRNAs regulate the expression of the *HTT* gene, whose mutation leads to Huntington's disease (HD), a hereditary degenerative disorder. Specifically, we validated the interactions of canonical miRNAs, miR-137, miR-214, and miR-148a, with the HTT 3'UTR using a luciferase assay. Moreover, we applied synthetic miRNA mimics to examine whether a slight shifting of miRNA seed regions might alter the regulation of the *HTT* transcript. We also examined miR-137, miR-214, and miR-148a isomiRs and showed the activity of these isoforms on reporter constructs bearing appropriate sequences from the HTT 3'UTR. Hence, we demonstrated that certain 5'-end variants of miRNAs might be functional for the regulation of the same targets as canonical miRNAs.

Keywords: miRNA; isomiR; target validation; luciferase assay; huntingtin; Huntington's disease

1. Introduction

MicroRNAs (miRNAs) are 21- to 24-nucleotide noncoding RNAs that fine-tune gene expression. These molecules act at the posttranscriptional level through modulation of translational efficiency and/or destabilization of target transcripts (reviewed in [1]). miRNAs exert their functions through imperfect pairing with the 3' untranslated region (UTR) of target mRNAs. Nucleotides 2 through 8 of the miRNA, termed the "seed" sequence, are essential for target recognition and binding [2].

The canonical pathway of animal miRNA biogenesis includes two subsequent cleavages (reviewed in [3–6]). Briefly, precursor miRNAs (~60-nt pre-miRNAs) are generated from primary transcripts (pri-miRNAs) through cleavage with the ribonuclease Drosha and exported to the cytoplasm by Exportin-5. Then, ~22-nt miRNA duplexes are generated through cleavage with the ribonuclease Dicer. Only one miRNA strand (the guide strand) of the duplex induces Argonaute proteins (AGO) to form the programmed RNA-induced silencing complex (RISC); the other strand (the passenger strand, or miRNA*) is released and degraded. The thermodynamic stability of the ends of the miRNA duplexes plays a crucial role in miRNA strand selection.

Currently, more than 2000 mature human miRNAs have been deposited in the miRNA repository (miRBase, Release 19) [7]. The deep sequencing of short RNAs has not only enabled the identification of novel miRNAs but also revealed that miRNAs are heterogeneous and differ in length. Heterogeneous miRNA variants are referred to as isomiRNAs (isomiRs) [8]. The primary source of the heterogeneity of miRNA length is imprecise cleavage by the ribonucleases Drosha and Dicer [8–11], which can be further biased at the AGO2 binding step [12]. However, miRNA length variation might also reflect various downstream effects, such as limited miRNA degradation by exonucleases, the addition of extra nucleotides [13–15], and miRNA sequencing artifacts [16,17]. It has recently been shown that the human trans-activation response (TAR) RNA-binding protein (TRBP), a molecular partner of Dicer, might also contribute to miRNA length heterogeneity. Specifically, TRBP triggered the production of isomiRs that were longer at the 5' strand than the canonical miRNAs by a single nucleotide. As a result, different mRNAs were targeted due to changes in guide-strand selection [18]. It has also been reported in *Drosophila* that the Nibbler (Nbr) 3'–5' exonuclease trims the 3'ends of miR-34 generating isomiRs shorter than the canonical sequence [19]; however, there is no evidence for similar exonuclease activity in vertebrates.

miRNAs control the expression of the majority of human genes [20], and these molecules are involved in many physiological and pathological processes. The alteration of miRNA expression has been associated with numerous diseases, including neurodegenerative disorders, such as Huntington's disease (HD). HD is the most common fatal polyglutamine (polyQ) disorder and results from the expansion of a CAG repeat in exon 1 of the huntingtin (*HTT*) gene. The precise mechanism of HD pathogenesis is not fully understood, but both the mutant protein (reviewed in [21]) and mutant transcript might be toxic to cells (reviewed in [22]). Of particular interest is the potential involvement of miRNA in the regulation of the *HTT* gene. The global deregulation of miRNAs in samples obtained

from HD patients was demonstrated using Illumina massively parallel sequencing [23]. Most importantly, miRNA of varying lengths and/or sequences (isomiRs) were observed for the vast majority of miRNAs detected in two forebrain areas, the frontal cortex (FC) and striatum (ST), of both healthy individuals and HD patients [23].

In general, the miRNA heterogeneity observed in deep sequencing might have important functional implications. Most importantly, miRNAs with shifted 5'-ends have different seed sequences responsible for the recognition of a complementary sequence and the binding to mRNA. Therefore, it is assumed that heterogeneous 5' isomiRs might regulate different targets [10,15,24,25]. Moreover, both 5' and 3' isomiRs might exhibit modified turnover properties [24,26] and altered strand selection within the RISC because strand selection is influenced by the extent of the 3' overhang and the degree of pairing for any miRNA-miRNA* duplex [27,28].

An early evidence supporting the hypothesis of isomiRs functionality comes from an experiment that showed a difference in target cleavage between miR-142-5p and its variant, which contained two extra nucleotides at the 5'-end [29]. A putative functional role for isomiRs has been suggested in many reports because isomiRs actively associate with the RISC and translational machinery [24,30–32] (reviewed in [33]). This assumption was further supported by the observation that isomiRs exhibit differential expression across tissues and developmental stages [26,34,35]. Nevertheless, the real biological significance of isomiRs is not fully understood because few studies concerning isomiR regulation at the cellular level have been reported, and thus far, only variants of miR-133, miR-101, and miR-31 have been experimentally examined. Specifically, it was shown that 5'-isomiR-101, which is highly expressed in the brain, associates with AGO2 immunocomplexes and decreases the expression of five validated miR-101 targets but to a lesser degree than the canonical miR-101 [35]. Differential mRNA targeting was demonstrated in the case of two prevalent 5' isomiRs of the key cardiac regulator miR-133a [31]. Three miR-31 isoforms that differed only slightly in their 5'- and/or 3'-end sequences were compared (namely, hsa-miR-31, ptr-miR-31, and mmu-miR-31), implicating isomiR-31s in the concordant and discordant regulation of six known target genes [36].

In this paper, we validated miRNA-mRNA interactions that might be involved in the regulation of the HTT transcript. Specifically, we experimentally assessed the validity of three predicted interactions and demonstrated that the canonical miR-137, miR-214, and miR-148a bind to the 3'UTR of the *HTT* gene. These results provide the first evidence that miR-137 and miR-148a regulate the expression of huntingtin and confirm that this regulation is also mediated by miR-214, as previously reported [37]. Moreover, using luciferase reporter assays, we investigated the regulation of huntingtin using select miRNA isoforms. We focused on 5'-end isomiRs with the shifted seed sequence that is the primary determinant of mRNA target recognition. Here, we showed that certain 5'-end isomiRs of miR-214 are functional for the downregulation of huntingtin expression.

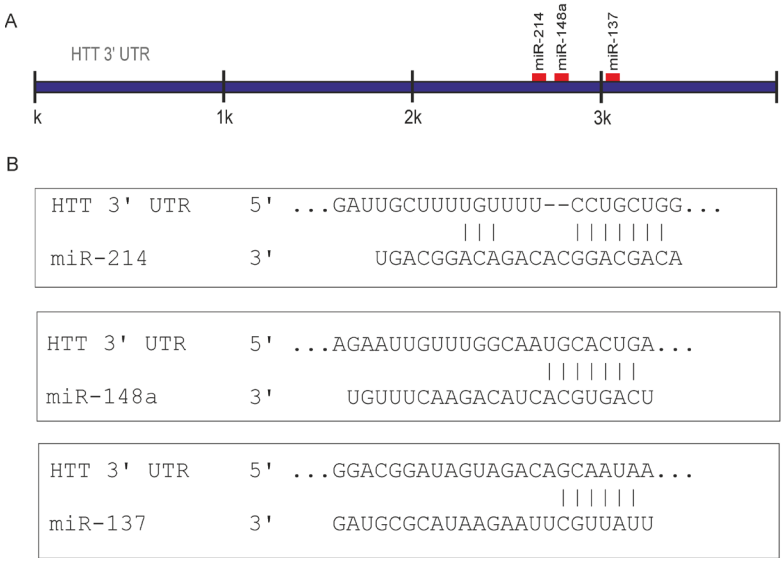
2. Results and Discussion

2.1. Prediction of miRNA Interactions with the HTT Transcript

In a previous study, we predicted potential miRNA interactions with mRNAs derived from genes triggering hereditary neurological disorders known as trinucleotide repeat expansion diseases (TREDs), including Huntington's disease (HD) [38]. The results of this in-depth *in silico* analysis prompted further research on the potential miRNA-mediated regulation of the HTT transcript in the context of the pathogenesis and therapy of HD. We compared different target prediction algorithms and verified our predictions using the available data gathered in various databases dedicated to miRNA target prediction (e.g., miRWalk database [39] and miRTarBase [40]). We selected interactions with miR-137, miR-214, and miR-148a for experimental verification. The deregulation of the expression of these miRNAs in HD patients or in cellular models of HD has been reported. Specifically, miR-137 was downregulated in the striatum of HD patients [23], while both miR-214 and miR-148a were upregulated in STHdh^{Q111}/Hdh^{Q111} cells [41]. Moreover, miR-137 is highly expressed in the nervous system, suggesting the involvement/potential role of this miRNA in the pathogenesis of HD. miR-137 has also been recently identified as a direct target of the repressor element-1 (RE-1) silencing transcription factor (REST) [42]. The second candidate, miR-214, has been positively verified in previous studies; miR-214, along with three other miRNAs (miRs 150, 146a, and 125b), downregulated the expression of huntingtin [37]. The same study also showed that these miRNAs affect the formation of mutHTT aggregates, the toxicity induced by mutHTT, and the expression of brain-derived neurotrophic factor (BDNF), thereby collectively contributing to HD pathogenesis.

The candidate miRNAs (miRs 137, 214, and 148a) ranked high in the results generated by either algorithm based on conservation criteria, *i.e.*, Diana-micro T [43], miRanda [44], or PicTar [45]. However, our prediction was primarily based on the use of the TargetScanHuman algorithm (Release 6.2) [46]. According to TargetScan, a site for miR-137 is highly conserved among vertebrates, and sites for miRs 214 and 148a are poorly conserved among mammals or vertebrates. In addition, the miR-137 and miR-148a sites were 8mers (defined as exact matches to positions 2–8 of the mature miRNA, followed by an adenine), while the selected miR-214 site was a 7mer-m8 (an exact match to positions 2–8 of the mature miRNA). The positions of the miR-137, miR-214, and miR-148a binding sites in the 3'UTR of the huntingtin transcript and the base pairing of these miRNAs with target sequences are presented in Figure 1. The binding parameters of these miRNAs met the recommended bioinformatics criteria, and their experimental validation was of particular interest in the light of current knowledge of potential involvement of miRNAs in neurodegeneration and the entire competing endogenous RNA (ceRNA) activity network [47], which recently has been shown to be implicated in neurodegenerative diseases including HD [48,49].

Figure 1. Graphical presentation of selected miRNA target site distribution in the 3' untranslated region (3'UTR) of the huntingtin transcript. To predict miRNAs that potentially target the HTT 3'UTR, the TargetScanHuman algorithm (Release 6.2) [46] was used. (**A**) Regions of interaction for the miRNAs selected for experimental validation; (**B**) miRNA base pairing with an appropriate target sequence is schematically presented.

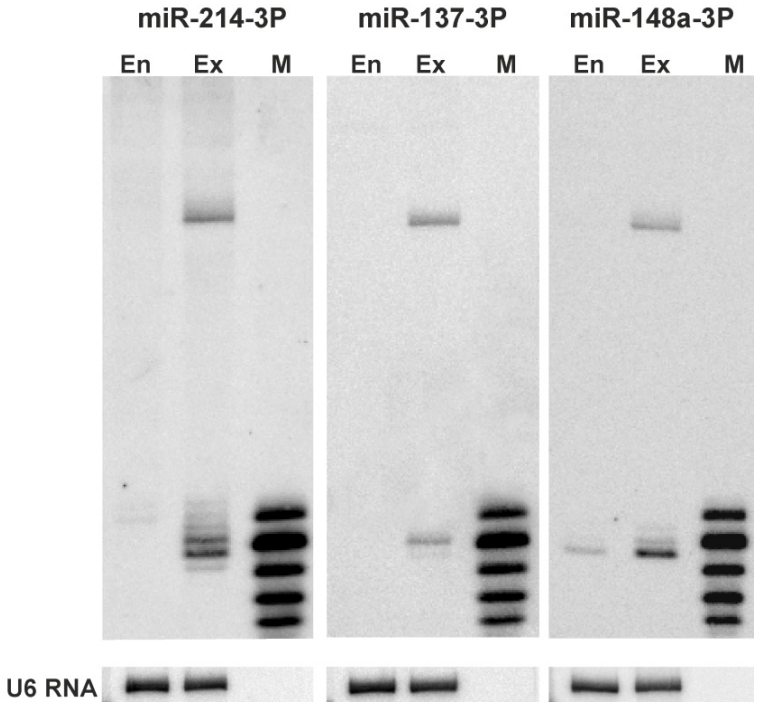


2.2. Canonical miR-137, miR-214, and miR-148a Regulate the Expression of the HTT Gene

For the experimental validation of the predicted binding of the selected canonical miRNAs (miRs 137, 214, and 148a) to their target sites in the HTT 3'UTR, experiments using reporter constructs and luciferase assays were performed as described previously [50]. However, sequences carrying binding sites for the appropriate miRNAs were cloned into pmirGLO vector (Promega), which is considered optimal for miRNA-mRNA interaction studies. Constructs bearing single miRNA binding sites were generated and defined as wild-type reporters (WT). Constructs with mutations that disrupted native pairing within the binding region (5' seed site) of the candidate miRNAs (MUT) and constructs that showed perfect complementarity (PM) to these sites were also generated to provide negative and positive controls, respectively (details in the Experimental section).

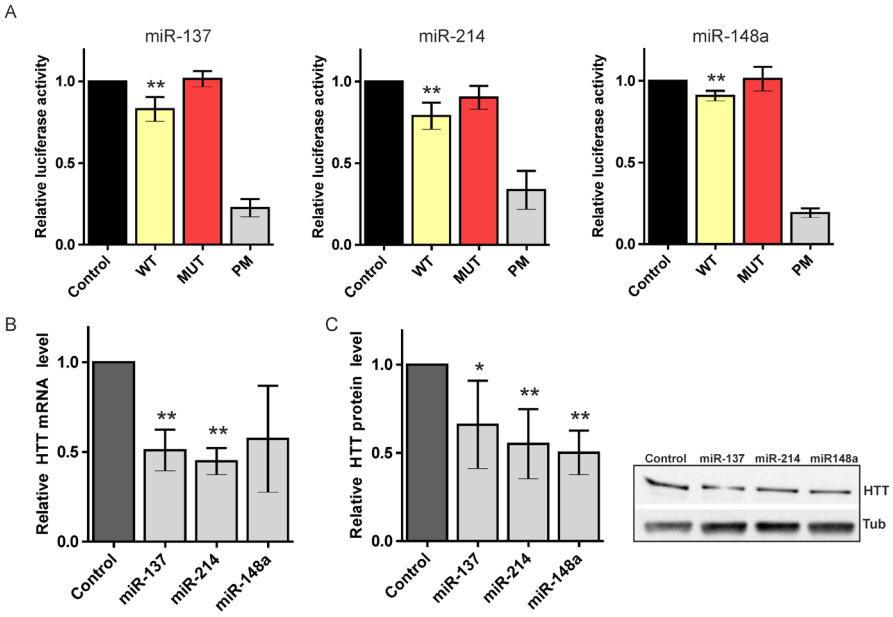
We transfected HEK293T cells with either reporter carrying potential miRNA binding sites. Four constructs were transfected into cells and tested in parallel. To determine whether the miRNAs of interest were expressed in the HEK293T cells, we performed northern blot analysis. The expression of miR-137 was not detected in the HEK293T cells. miR-214 and miR-148a were expressed at low and moderate levels (Figure 2), respectively, consistent with the available deep sequencing results. Therefore, our experimental system required miRNA overexpression, and we used miRNA-coding plasmid vectors (System Biosciences, Open Biosystems) for this purpose (compare endogenous miRNA levels and those expressed from vectors in Figure 2).

Figure 2. Endogenous expression and overexpression of miR-137, miR-214, and miR-148a in HEK293T cells. Northern blot detection of miRs 137, 214, and 148a in non-treated HEK293T cells and cells transfected with miRNA-coding plasmids (System Biosciences, Open Biosystems). M denotes the size marker, end-labeled 17, 19, 21, 23, and 25-nt oligoribonucleotides. En and Ex indicate the miRNA levels, endogenous and expressed from appropriate vectors, respectively. Hybridization to U6 RNA provides a loading control.



In the luciferase assays, we obtained considerable repression of the luciferase expression after the transfection of reporter constructs for the three miRNAs tested (Figure 3A). Specifically, we observed a significant reduction in luciferase activity when reporter constructs bearing binding sites for miRs 137 and 214 were used (reductions to 83% and 79%, respectively) and a slightly weaker but reproducible and statistically significant suppression of the luciferase activity in the case of miR-148a (suppression to 87%). The luciferase activity for all of the MUT constructs showed efficient de-repression nearly equal to that in the control experiment; the positive controls (PMs) repressed luciferase at low levels, ranging from 17% to 33% (for miR-148a PM and miR-214 PM, respectively) of the empty reporter construct. These results verify the reliability of the experimental system used.

Figure 3. Regulation of the huntingtin (HTT) expression by canonical miRNAs. (**A**) Relative repression of the luciferase expression. Reporter constructs carrying a single binding site for miR-137, miR-214, and miR-148a were tested. For each luciferase. experiment, the miRNA activity on four constructs was measured in parallel: an empty pmirGlO vector (Control), a wild-type potential binding site for the appropriate miRNA (WT), a mutated binding site (MUT), and a site with full complementarity (PM). The firefly luciferase activity was normalized against *Renilla* luciferase activity. An average result from at least three independent experiments is shown (details in the text); (**B**) Relative HTT mRNA levels. Real-time PCR performed 48 h after transfection of HEK293T cells with miR-137, miR-214, and miR-148a. The bar graphs show the quantification of the HTT mRNA levels normalized to actin mRNAs based on data collected from three independent experiments; (**C**) Relative HTT protein levels. Western blot analysis of the cellular levels of HTT protein 72 h after transfection of HEK293T cells with miR-137, miR-214, and miR-148a. The bar graphs show the quantification of the protein levels detected in three western blot experiments. A representative blot is shown. The asterisks indicate statistical significance; a single asterisk at p-value < 0.05 and a double asterisk at p-value < 0.01.



We also monitored huntingtin expression at the mRNA and protein levels following the transfection of HEK293T cells with miRNA-coding plasmids. Real-time PCR performed 48 h after transfection with miR-137, miR-214, or miR-148a showed a strong decrease in the HTT mRNA level (Figure 3B). Similarly, the HTT protein level was significantly reduced 72 h posttransfection in cells overexpressing any of the miRNAs (Figure 3C, Figure S1). This observation is consistent with the finding that miRNA binding reduces the cellular levels of targeted transcripts [51,52]. However, other

studies have reported that no or minimal changes in the respective mRNA levels were observed or that these changes were only reported for certain targets [35]. Overall, miR-137, -214 and -148a were positively verified as negative regulators of the *HTT* gene. The lack of regulation of the huntingtin expression, demonstrated in both luciferase assays and western blotting, was observed for the other miRNA (miR-107) and shown for comparison as supplementary data (Figure S2). The strongest reduction in the luciferase activity and the greatest and second-greatest repression at the mRNA and protein levels were observed with miR-214. Thus, this study provides further support for the regulatory potential of miR-214, which was previously validated in a different experimental system [37]. Moreover, this study provides the first evidence of HTT regulation by miR-137 and miR-148a.

2.3. 5'-End Variants of miRNAs Are Functional and Might Regulate the Same Targets as Canonical miRNAs

Although many reports suggest isomiR functionality [24,30–32], there is still little research that address this issue experimentally. Specifically, one variant of miR-101 [35] and two isomiRs of miR-133 [31] and miR-31 [36] have been investigated. In these reports, the isomiRs were less effective than their canonical analogs [35] or exhibited differences in effectiveness depending on the regulated target [31].

Here, we determined whether the 5'-end variants of three miRNAs (5'-end isomiRs), namely miRs 137, 214 and 148a, might function in the same experimental system (*i.e.*, whether these miRs reduce the luciferase activity when appropriate reporter constructs are used). We designed and synthesized miRNA variants with seed sequences shifted by −1, +1, or +2 nt (Integrated DNA Technologies) (Figure 4A). We selected miRNA 5' isoforms that are relatively highly represented in deep sequencing data because we considered sequence abundance a prerequisite for the functionality of these molecules. We based this selection on the sequencing data gathered in the YM 500 database [53] but we also evaluated the expression levels of isomiRs in other sources [32]. The only exception was isomiR-137+1, whose sequence is barely detectable using deep sequencing. This isomiR variant was added to the analysis to examine the same miRNA seed shifts for all isomiRs tested. Moreover, trimming variants that affect the 5' end of miRNAs were reported to be abundant species, and the vast majority of these 5' isomiRs affected a single nucleotide upstream of the reference miRNA [35]. A strong correlation between the expression of miRNAs and isomiRs was also observed [30].

According to the TargetScanHuman Custom (Release 5.2) [46] prediction, none of the selected isomiRs targeted HTT (Figure 4A); thus, we verified the targeting of these molecules experimentally. In addition, we assessed *in silico* how the overall number of genes targeted by the analyzed miRNAs and isomiRs might vary due to the change introduced into their seed regions. Potential targets for the 5'-end variants of miR-137, miR-214, and miR-148a were predicted using the TargetScan Custom 5.2 algorithm [46] and are shown in the Venn diagrams by overlaps (Figure 4B). Specifically, targets for the canonical miRNAs were compared with the targets of the miRNAs with seed regions shifted by −1, +1, and +2 nt. This analysis revealed that the number of predicted targets changed, but apart from

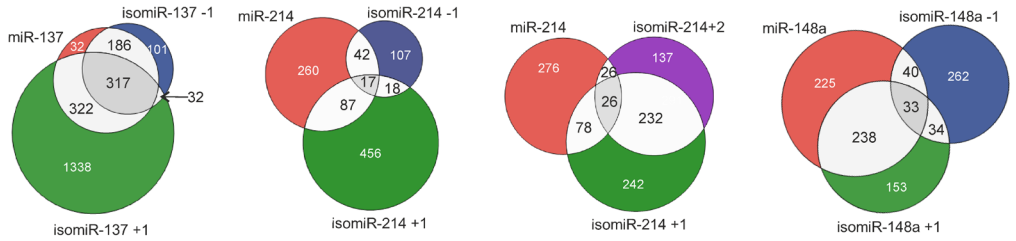
unique targets, many genes were still predicted as targets for both miRNAs and isomiRs, confirming that isomiRs might share certain common mRNA targets but not all mRNA targets [36]. These results are also consistent with the suggestion that isomiRs function cooperatively to target common biological pathways [30]. However, distinct functions for miRs and isomiRs have also been suggested [31,35].

Figure 4. Graphical presentation of selected isomiR variants and their potential to target different genes. (**A**) Nucleotide sequences of miR-137, miR-214, miR-148a, and their isoforms. miRNA sequences are marked in red, and isomiR sequences are shown in blue, green, and violet for −1-, +1-, and +2-nt seed shifting, respectively. The miRNA seed sequences are labeled with black rectangles. Information on the miRNA lengths, as well as their potential for targeting the *HTT* gene and isomiR expression levels, is also provided. (*) Ability to interact with the HTT 3'UTR, as predicted by the TargetScanHuman algorithm (Release 6.2) for miRNAs and the TargetScan Custom (Release 5.2) for isomiRs [46], (**) isomiR read number according to the YM500 database [53]; (**B**) Venn diagrams showing the predicted miRNA targets for selected isomiRs. Potential targets for the 5'-end variants of miR-137, miR-214, and miR-148a were predicted using the TargetScan Custom algorithm (Release 5.2) and are shown as overlaps in the Venn diagrams. Targets for the canonical miRNAs are compared with the targets for the miRNAs with the shifted seed regions and are depicted in the same colors as in panel A. The numbers inside the circles denote the numbers of potential targets predicted for the appropriate miRNA variants.

A

miRNA name	sequence	length nt	target HTT*	read no.**
hsa-miR-137	5'...CGUUGUUAUUGCUUAAGAAUACGCGUAGUCGAG...3'	23	Yes	
isomiR-137+1	GUUAUUGCUUAAGAAUACGCGUAG	24	No	-
isomiR-137-1	UAUUGCUUAAGAAUACGCGUAG	22	No	15275
hsa-miR-214	5'...CCUGUACAGCAGGCACAGACAGGCAGUCACA...3'	22	Yes	
isomiR-214+1	UACAGCAGGCACAGACAGGCAGU	23	No	16189
isomiR-214+2	GUACAGCAGGCACAGACAGGCAGU	24	No	686
isomiR-214-1	CAGCAGGCACAGACAGGCAGU	21	No	437
hsa-miR-148a	5'...AGAAGUCAGUGCACUACAGAACUUUGUCUC 3'	22	Yes	
isomiR-148a+1	GUCAGUGCACUACAGAACUUUGU	23	No	4980
isomiR-148a-1	CAGUGCACUACAGAACUUUGU	21	No	33873

B



To validate the regulation of the HTT transcript by canonical miRNAs in a luciferase assay, we overexpressed the desired miRNAs from plasmid vectors. To study the interactions of the HTT transcript with isomiRs, appropriate isomiR sequences had to be introduced into cells as synthetic oligonucleotides. Thus, we transfected HEK293T cells with both the miR-137 mimic and miR-137 vector (System Biosciences) to determine whether these two experimental systems generate the same results (Figure 5A). Moreover, we examined miRNA mimic activities at different final concentrations (10, 30, and 50 nM) to determine the optimal concentration for these experiments (Figure 5B). A clear correlation between the results of the luciferase experiments with the miR-137-coding plasmid and the synthetic miR-137 mimic was observed; thus, we further investigated the functionality of our 5'-end isomiRs using appropriate miRNA mimics. In the luciferase assays, we obtained considerable and significant repression of the luciferase expression after the transfection of the reporter constructs and all three miR-214 isomiR mimics, namely, isomiR-214+1, isomiR-214+2, and isomiR-214-1 (luciferase repression equal to 71%, 80%, and 79%, respectively). Moreover, this reduction in the luciferase activity was comparable to the reduction induced by the canonical miR-214 mimic (71%) (Figure 6A). In contrast, the luciferase activity was not reduced when miR-137 isomiRs were used, in the case of neither isomiR-137+1 nor isomiR-137-1, compared with the considerable repression observed using the canonical miR-137 mimic (79%) (Figure 6B). Similarly, in the case of isomiR-148a+1 and isomiR-148a-1, the activity of luciferase was slightly reduced (9% and 6%, respectively), while the reduction obtained for the canonical miR-148a mimic was much stronger (80%) (Figure 6C). The observed difference in the functionality of the analyzed isomiRs raises the question when miRNA-mRNA pairing conforms to strict rules and when some flexibility in the miRNA seed region is permitted, and which additional mechanisms other than the base paring of the seed region might affect target genes repression by isomiRs.

Several factors influence the recognition of a target site by miRNA, e.g., the sequence composition of the 3'-UTR [54], the immediate environment of the putative target site [55], and the structural accessibility of the target site [2,56]. Moreover, endogenous natural antisense transcripts transcribed from the opposite strand of a protein-coding gene or a non-protein coding gene [34] and the RNA-binding proteins [57] could directly bind to mRNA, thereby masking the miRNA binding site of a target gene and preventing the inhibitory effects of the miRNA on target gene translation. These factors, however, are of importance to canonical miRNA binding. Here, we examined several 5' isomiRs of slightly different lengths that previously demonstrated canonical miRNA targeting. Therefore, the structural features and genomic context of these molecules did not significantly differ between the canonical miRNAs and their isomiRs or between the isomiRs themselves.

A distinct feature of the functional isomiR-214 variants and the two other isomiRs examined in this study was the fact that miR-214 is a 7mer with compensatory base pairing at the 3' end (see Figure 1). Although canonical miRNA-target specificity is primarily triggered by complementarity within the seed region, non-canonical interactions depend also on 3' compensatory sites [2,58], which might be important for miR-214 and its variants. The miRNA/isomiR length was also suggested as a factor that

might affect functionality. In a study of isomiRs, the analysis of two miR-133a mimics (22/23 nt) was performed, followed by the analysis of two other variants that represented the respective other length for each miR-133a variant. However, the luciferase repression did not depend on mimic length within this range [31]. Therefore, alterations to the 3' end of the miR-133a mimic did not affect the level of mRNA repression, suggesting that the 3' end is not essential for efficient target binding in this case. Another important factor that might account for the disparate functioning of isomiRs is differential binding capacity with the Argonaute complex (affinity of a given miRNA to AGO). Previous studies have shown that some miRNA variants were differentially loaded onto AGOs, and the 5'-end nucleotide of small RNA was critical for its interaction with AGO proteins [12,59–61]. However, miR-101 was more efficiently loaded into the RISC than its isomiR [35], and the 5'-end nucleotide of isomiR-31s was not a rigorous criterion for AGO complex loading [36]. In this study, in the case of the most effective miRNA, namely miR-214, all variants were functional regardless of the different nucleotides at their 5' end (Figure 2). Small changes in the miRNA sequence profoundly affected the functional asymmetry of the miRNA duplex, altering which strand of a miRNA duplex functions in mRNA silencing [18]. Therefore, it cannot be ruled out that, in the case of the nonfunctional isomiRs of miR-137 and miR-148a, the passenger strands were incorporated into the RISC and did not target their binding sites.

Figure 5. Correlation between the results of the luciferase experiments conducted with miR-137-coding plasmid and synthetic miR-137 mimics. (**A**) Relative repression of the luciferase expression. Reporter constructs carrying a single binding site for miR-137 were tested; miRNA activity on four constructs was measured in parallel (Control, WT, MUT, and PM), as described in Figure 3. Left—miRNA expression from the synthetic oligonucleotide (miR-137 mimic), right—miRNA overexpressed from the miR-137 vector. The firefly luciferase activity was normalized against *Renilla* luciferase activity. The standard errors are calculated from three independent experiments; (**B**) The relative repression of the luciferase expression resulted from the miRNA mimic activity. Four reporter constructs were tested (Control, WT, MUT, and PM) but with the addition of miR-137 mimic at different final concentrations, specifically 10, 30, and 50 nM, as denoted in the figure. The standard errors were calculated from one experiment performed in triplicate.

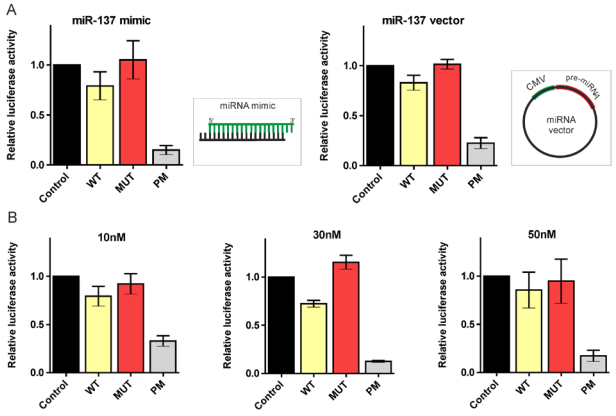
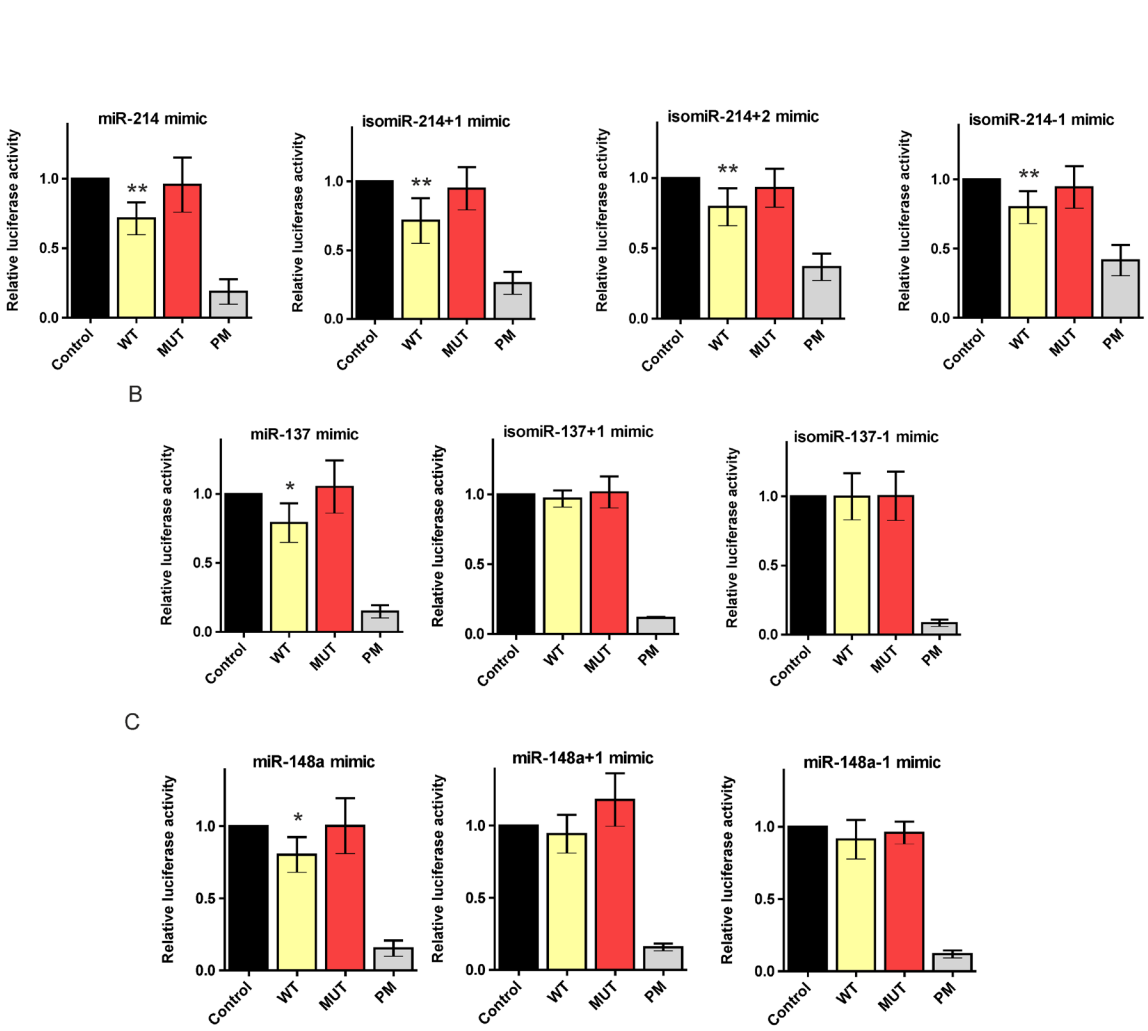


Figure 6. Regulation of the huntingtin expression by isomiRs. Relative repression of the luciferase expression for miR-214, miR-137, miR-148a, and their isomiRs (+1, +2, or −1). Reporter constructs carrying single binding sites for the appropriate miRNAs were tested, namely miR-137 (**A**), miR-214 (**B**), and miR-148a (**C**), as depicted in the figure. For each luciferase experiment, the miRNA activity on four constructs (Control, WT, MUT and PM) was measured in parallel, as described in Figures 3 and 5. The firefly luciferase activity was normalized against *Renilla* luciferase activity. The standard errors were calculated from three independent experiments. The asterisks indicate statistical significance; a single asterisk at p-value < 0.05 and a double asterisk at p-value < 0.01.



3. Experimental Section

3.1. Cell Culture

HEK293T cells were obtained from the American Type Culture Collection (ATCC) and grown in Dulbecco's Modified Eagle's Medium (DMEM, Lonza, Wakersville, MD, USA) supplemented with

8% fetal bovine serum (FBS) (Sigma-Aldrich, St. Louis, MO, USA), 2 mM L-glutamine, and an antibiotic-antimycotic solution (Sigma-Aldrich, St. Louis, MO, USA) at 37 °C in a humidified atmosphere of 5% CO_2. At 24 h prior to transfection, the HEK293T cells were plated in 12-well or 6-well dishes in DMEM growing medium and harvested 24, 48, and 72 h post-transfection for the luciferase assay, real-time PCR, and western blot analyses, respectively.

3.2. Plasmid Constructs and Synthetic miRNA Oligonucleotides

To generate reporter constructs bearing miRNA-binding sites, the pmirGLO Dual-Luciferase miRNA Target Expression Vector was used (Promega, Madison, WI, USA). This vector is based on Promega dual-luciferase technology, with firefly luciferase (*luc2*) as the primary reporter for monitoring mRNA regulation and *Renilla* luciferase (*hRluc-neo*) as a control reporter for normalization and selection. Specific oligonucleotides with *Dra*I and *Xba*I ends containing single binding sites for the analyzed miRNA (HTT b.s. for miRs 214, 137, and 148a) were synthesized (IBB Warsaw). The appropriate oligos were annealed by boiling and gradual cooling and subsequently phosphorylated and cloned into the pmirGLO vector, previously digested with *Dra*I (Fermentas, St.-Leon-Rot, Germany) and *Xba*I (Fermentas, St.-Leon-Rot, Germany) restriction enzymes, downstream of the *luc2* gene. For all miRNAs, three types of constructs were prepared, namely wild type (WT), carrying mutations (MUT) and perfect match (PM) constructs (for sequences refer to Table S1), which all have 10-nucleotide flanking sequences, as described previously [50].

For miRNA overexpression, commercial plasmid constructs expressing miRNA precursors (pri-miR-148a (Open Biosystems, Huntsville, AL, USA), pri-miR-137, or pri-miR-214 (System Biosciences, Mountain View, CA, USA)) were used. These plasmids contain pri-miRNA sequences in their natural genome context to ensure biologically relevant interactions with the endogenous processing machinery.

Synthetic miRNA mimics (miR-137, miR-214, and miR-148a mimics) and their length variants were chemically synthesized (Integrated DNA Technologies). The following modifications were introduced: (1) 2'-*O*-methyl modification on positions 1 and 2 and a two-nucleotide UU overhang on the 3' end of the miRNA mimic sense strand, (2) 5' phosphorylation and a two-nucleotide overhang based on nucleotide types found in natural pre-miRNAs on the 3' end of the miRNA mimic antisense strand. All sequences are presented as supplementary data (Table S2).

3.3. Cell Transfection

HEK293T cells were transfected using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocols. For luciferase assays, the cells were transfected in 12-well plates at ~80% confluence. For each transfection experiment, 200 ng of the appropriate reporter construct and either 250 ng of the appropriate miRNA-coding vector or 30 nM of miRNA mimic were used. The cells were harvested 24 h after transfection and assayed for luciferase activity. For miRNA overexpression required for real-time PCR and western blot analyses, the cells were grown to 80% and

60% confluence, respectively, transfected in 6-well plates with 1 μg/mL pri-miRNA plasmid vectors, and harvested at 48 and 72 h, respectively.

3.4. Luciferase Reporter Assay

After harvesting, the cells were lysed in a passive lysis buffer (Promega, Madison, WI, USA). The luciferase activity was measured using a Dual-Luciferase Reporter Assay System (Promega, Madison, WI, USA) according to the manufacturer's instructions with a Centro LB 960 luminometer (Berthold Technologies, Oak Ridge, TN, USA).

3.5. RNA Isolation and Real-Time PCR

Total RNA from HEK293T cells was isolated using TRI Reagent (MRC, Inc., BioShop, Cincinnati, OH, USA) according to the manufacturer's instructions. The RNA concentration was estimated using a NanoDrop spectrophotometer. cDNA was obtained from 500 ng of total RNA using Superscript III (Life Technologies, Carlsbad, CA, USA) and random hexamer primers (Promega, Madison, WI, USA). For subsequent quantitative real-time analyses, 50 ng of cDNA was used. Real Time PCR was performed on a LightCycler 480 II system (Roche Diagnostics, Mannheim, Germany) using TaqMan Gene Expression Assays and TaqMan Universal Master Mix II (Applied Biosystems, Foster City, CA, USA). The results obtained for the assessment of huntingtin mRNA levels were normalized to the levels of actin mRNA.

3.6. Northern Blotting

High-resolution northern blotting was performed as previously described [62,63]. Briefly, 25 μg of total RNA was extracted from HEK293T cells and resolved on a 12% denaturing polyacrylamide gel in 0.5× TBE. The RNA was transferred to a GeneScreen Plus hybridization membrane (PerkinElmer, Spokane, WA, USA) using semi-dry electroblotting (Sigma-Aldrich, St. Louis, MO, USA), immobilized by subsequent UV irradiation (120 mJ/cm^2) (UVP), and baked in an oven at 80 °C for 30 min. The membranes were probed with specific DNA oligonucleotides (Table S3) complementary to the annotated human miRNAs miR-137-3P, miR-214-3P, and miR-148a-3P (miRBase). The probes were labeled with [γ^{32}P] ATP (5000 Ci/mmol; Hartmann Analytics, Braunschweig, Germany) using USB OptiKinase (Affymetrix, Cleveland, OH, USA). The hybridizations were performed at 37 °C overnight in a PerfectHyb buffer (Sigma-Aldrich, St. Louis, MO, USA). The marker lanes contained a mixture of radiolabeled RNA oligonucleotides (17-, 19-, 21-, 23-, and 25-nt in length). Hybridizations to U6 RNA provided loading controls. Radioactive signals were quantified by phosphorimaging (Multi Gauge v3.0; Fujifilm).

3.7. Western Blotting

A total of 15 μg of protein was diluted in sample buffer containing 2-mercaptoethanol, denatured for 5 min, and separated on 3%–8% gradient Tris-Acetate gels (Invitrogen, Carlsbad, CA, USA)

in XT Tricine Buffer (BioRad, Hercules, CA, USA). After electrophoresis, the proteins were electrotransferred onto a nitrocellulose membrane (Sigma, St. Louis, MO, USA). All immunodetection steps were performed on a SNAPid (Millipore, Billerica, MA, USA) in PBS buffer containing 0.25% nonfat milk and 0.1% Tween 20, and the membranes were washed in PBS/Tween. For huntingtin and tubulin detection, the blots were probed with the primary anti-huntingtin (1:500, Millipore, Billerica, MA, USA) and anti-alpha-tubulin (1:5000, Covance, Emeryville, CA, USA) antibodies, respectively, and subsequently probed with HRP-conjugated secondary antibodies (1:500, Sigma, St. Louis, MO, USA). The immunoreaction was detected using Western Bright Quantum (Advansta, CA, USA). The protein amounts were quantified using GelPro 3.1 software (Media Cybernetics, Bethesda, MD, USA).

3.8. Statistical Analysis

All experiments were repeated at least three times. Graphs were generated using GraphPad Prism 5 (GraphPad Software). The figures for the luciferase assays were generated after averaging the results from the repeat experiments for a particular construct. The values for error bars (mean with SD) and the statistical significance were calculated using GraphPad Prism 5. The statistical significance of the luciferase reduction in the case of transfection with constructs carrying miRNA-binding sites was assessed using a one-sample *t*-test with a hypothetical value of 1 assigned to cells transfected with a control empty vector. p-values < 0.05 (two-tailed) were considered significant.

4. Conclusions

This study presents new evidence that *HTT* gene expression is regulated by miRNAs and, most importantly, demonstrates that certain isomiRs are functional and regulate the same target as canonical miRNAs.

IsomiRs are commonly reported in deep-sequencing studies and have been described in all studied organisms and tissues. The existence of miRNA variants might contribute considerably to the complexity of target regulation by miRNAs and strongly increase the regulatory potential of these molecules. The presence of isomiRs could have far-reaching implications for miRNA therapeutic applications; it must be taken into account in various diagnostic tests as well as in the design of miRNA mimics or anti-miRs as therapeutic agents. Therefore, of particular importance is to identify factors that determine the biological relevance of isomiRs.

Acknowledgments

This work was supported by funding from the Polish Ministry of Science and Higher Education (N N301 523038), the National Science Centre (2011/03/B/NZ1/03259), and the European Regional Development Fund within the Innovative Economy Programme (POIG.01.03.01-00-098/08). The real-time PCR analyses were performed on a LightCycler 480 II system (Roche) in the Laboratory of Subcellular Structures Analysis at the Institute of Bioorganic Chemistry, PAS, in Poznan.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Chekulaeva, M.; Filipowicz, W. Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. *Curr. Opin. Cell Biol.* **2009**, *21*, 452–460.
2. Bartel, D.P. MicroRNAs: Target recognition and regulatory functions. *Cell* **2009**, *136*, 215–233.
3. Kim, V.N.; Han, J.; Siomi, M.C. Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 126–139.
4. Krol, J.; Loedige, I.; Filipowicz, W. The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* **2010**, *11*, 597–610.
5. Starega-Roslan, J.; Koscianska, E.; Kozlowski, P.; Krzyzosiak, W.J. The role of the precursor structure in the biogenesis of microRNA. *Cell. Mol. Life Sci.* **2011**, *68*, 2859–2871.
6. Winter, J.; Jung, S.; Keller, S.; Gregory, R.I.; Diederichs, S. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat. Cell Biol.* **2009**, *11*, 228–234.
7. Griffiths-Jones, S.; Saini, H.K.; van Dongen, S.; Enright, A.J. miRBase: Tools for microRNA genomics. *Nucleic Acids Res.* **2008**, *36*, D154–D158.
8. Morin, R.D.; O'Connor, M.D.; Griffith, M.; Kuchenbauer, F.; Delaney, A.; Prabhu, A.L.; Zhao, Y.; McDonald, H.; Zeng, T.; Hirst, M.; *et al.* Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.* **2008**, *18*, 610–621.
9. Starega-Roslan, J.; Krol, J.; Koscianska, E.; Kozlowski, P.; Szlachcic, W.J.; Sobczak, K.; Krzyzosiak, W.J. Structural basis of microRNA length variety. *Nucleic Acids Res.* **2011**, *39*, 257–268.
10. Seitz, H.; Ghildiyal, M.; Zamore, P.D. Argonaute loading improves the 5' precision of both MicroRNAs and their miRNA* strands in flies. *Curr. Biol.* **2008**, *18*, 147–151.
11. Wu, H.; Ye, C.; Ramirez, D.; Manjunath, N. Alternative processing of primary microRNA transcripts by Drosha generates 5' end variation of mature microRNA. *PLoS One* **2009**, *4*, e7566.
12. Frank, F.; Sonenberg, N.; Nagar, B. Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature* **2010**, *465*, 818–822.
13. Landgraf, P.; Rusu, M.; Sheridan, R.; Sewer, A.; Iovino, N.; Aravin, A.; Pfeffer, S.; Rice, A.; Kamphorst, A.O.; Landthaler, M.; *et al.* A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **2007**, *129*, 1401–1414.
14. Ruby, J.G.; Jan, C.; Player, C.; Axtell, M.J.; Lee, W.; Nusbaum, C.; Ge, H.; Bartel, D.P. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* **2006**, *127*, 1193–1207.
15. Wu, H.; Neilson, J.R.; Kumar, P.; Manocha, M.; Shankar, P.; Sharp, P.A.; Manjunath, N. miRNA profiling of naive, effector and memory CD8 T cells. *PLoS One* **2007**, *2*, e1020.

16. Huse, S.M.; Huber, J.A.; Morrison, H.G.; Sogin, M.L.; Welch, D.M. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol.* **2007**, *8*, R143.
17. Tian, G.; Yin, X.; Luo, H.; Xu, X.; Bolund, L.; Zhang, X.; Gan, S.Q.; Li, N. Sequencing bias: Comparison of different protocols of microRNA library construction. *BMC Biotechnol.* **2010**, *10*, 64.
18. Lee, H.Y.; Doudna, J.A. TRBP alters human precursor microRNA processing *in vitro*. *RNA* **2012**, *18*, 2012–2019.
19. Liu, N.; Abe, M.; Sabin, L.R.; Hendriks, G.J.; Naqvi, A.S.; Yu, Z.; Cherry, S.; Bonini, N.M. The exoribonuclease Nibbler controls 3' end processing of microRNAs in *Drosophila*. *Curr. Biol.* **2011**, *21*, 1888–1893.
20. Friedman, R.C.; Farh, K.K.; Burge, C.B.; Bartel, D.P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **2009**, *19*, 92–105.
21. Ross, C.A.; Tabrizi, S.J. Huntington's disease: From molecular pathogenesis to clinical treatment. *Lancet Neurol.* **2011**, *10*, 83–98.
22. Fiszer, A.; Krzyzosiak, W.J. RNA toxicity in polyglutamine disorders: Concepts, models, and progress of research. *J. Mol. Med.* **2013**, *91*, 683–691.
23. Marti, E.; Pantano, L.; Banez-Coronel, M.; Llorens, F.; Minones-Moyano, E.; Porta, S.; Sumoy, L.; Ferrer, I.; Estivill, X. A myriad of miRNA variants in control and Huntington's disease brain regions detected by massively parallel sequencing. *Nucleic Acids Res.* **2010**, *38*, 7219–7235.
24. Chiang, H.R.; Schoenfeld, L.W.; Ruby, J.G.; Auyeung, V.C.; Spies, N.; Baek, D.; Johnston, W.K.; Russ, C.; Luo, S.; Babiarz, J.E.; *et al.* Mammalian microRNAs: Experimental evaluation of novel and previously annotated genes. *Genes Dev.* **2010**, *24*, 992–1009.
25. Kawahara, Y.; Megraw, M.; Kreider, E.; Iizasa, H.; Valente, L.; Hatzigeorgiou, A.G.; Nishikura, K. Frequency and fate of microRNA editing in human brain. *Nucleic Acids Res.* **2008**, *36*, 5270–5280.
26. Fernandez-Valverde, S.L.; Taft, R.J.; Mattick, J.S. Dynamic isomiR regulation in *Drosophila* development. *RNA* **2010**, *16*, 1881–1888.
27. Khvorova, A.; Reynolds, A.; Jayasena, S.D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **2003**, *115*, 209–216.
28. Schwarz, D.S.; Hutvagner, G.; Du, T.; Xu, Z.; Aronin, N.; Zamore, P.D. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **2003**, *115*, 199–208.
29. Azuma-Mukai, A.; Oguri, H.; Mituyama, T.; Qian, Z.R.; Asai, K.; Siomi, H.; Siomi, M.C. Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7964–7969.
30. Cloonan, N.; Wani, S.; Xu, Q.; Gu, J.; Lea, K.; Heater, S.; Barbacioru, C.; Steptoe, A.L.; Martin, H.C.; Nourbakhsh, E.; *et al.* MicroRNAs and their isomiRs function cooperatively to target common biological pathways. *Genome Biol.* **2011**, *12*, R126.
31. Humphreys, D.T.; Hynes, C.J.; Patel, H.R.; Wei, G.H.; Cannon, L.; Fatkin, D.; Suter, C.M.; Clancy, J.L.; Preiss, T. Complexity of murine cardiomyocyte miRNA biogenesis, sequence variant expression and function. *PLoS One* **2012**, *7*, e30933.

32. Lee, L.W.; Zhang, S.; Etheridge, A.; Ma, L.; Martin, D.; Galas, D.; Wang, K. Complexity of the microRNA repertoire revealed by next-generation sequencing. *RNA* **2010**, *16*, 2170–2180.
33. Neilsen, C.T.; Goodall, G.J.; Bracken, C.P. IsomiRs—The overlooked repertoire in the dynamic microRNAome. *Trends Genet.* **2012**, *28*, 544–549.
34. Faghihi, M.A.; Zhang, M.; Huang, J.; Modarresi, F.; van der Brug, M.P.; Nalls, M.A.; Cookson, M.R.; St-Laurent, G., 3rd; Wahlestedt, C. Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol.* **2010**, *11*, R56.
35. Llorens, F.; Banez-Coronel, M.; Pantano, L.; Del Rio, J.A.; Ferrer, I.; Estivill, X.; Marti, E. A highly expressed miR-101 isomiR is a functional silencing small RNA. *BMC Genomics* **2013**, *14*, 104.
36. Chan, Y.T.; Lin, Y.C.; Lin, R.J.; Kuo, H.H.; Thang, W.C.; Chiu, K.P.; Yu, A.L. Concordant and discordant regulation of target genes by miR-31 and its isoforms. *PLoS One* **2013**, *8*, e58169.
37. Sinha, M.; Ghose, J.; Bhattarcharyya, N.P. Micro RNA-214,-150,-146a and-125b target Huntingtin gene. *RNA Biol.* **2011**, *8*, 1005–1021.
38. Witkos, T.M.; Koscianska, E.; Krzyzosiak, W.J. Practical aspects of microRNA target prediction. *Curr. Mol. Med.* **2011**, *11*, 93–109.
39. Dweep, H.; Sticht, C.; Pandey, P.; Gretz, N. miRWalk—Database: Prediction of possible miRNA binding sites by "walking" the genes of three genomes. *J. Biomed. Inform.* **2011**, *44*, 839–847.
40. Hsu, S.D.; Lin, F.M.; Wu, W.Y.; Liang, C.; Huang, W.C.; Chan, W.L.; Tsai, W.T.; Chen, G.Z.; Lee, C.J.; Chiu, C.M.; *et al.* miRTarBase: A database curates experimentally validated microRNA-target interactions. *Nucleic Acids Res.* **2011**, *39*, D163–D169.
41. Sinha, M.; Ghose, J.; Das, E.; Bhattarcharyya, N.P. Altered microRNAs in STHdh(Q111)/ Hdh(Q111) cells: miR-146a targets TBP. *Biochem. Biophys. Res. Commun.* **2010**, *396*, 742–747.
42. Soldati, C.; Bithell, A.; Johnston, C.; Wong, K.Y.; Stanton, L.W.; Buckley, N.J. Dysregulation of REST-regulated coding and non-coding RNAs in a cellular model of Huntington's disease. *J. Neurochem.* **2013**, *124*, 418–430.
43. Kiriakidou, M.; Nelson, P.T.; Kouranov, A.; Fitziev, P.; Bouyioukos, C.; Mourelatos, Z.; Hatzigeorgiou, A. A combined computational-experimental approach predicts human microRNA targets. *Genes Dev.* **2004**, *18*, 1165–1178.
44. John, B.; Enright, A.J.; Aravin, A.; Tuschl, T.; Sander, C.; Marks, D.S. Human microRNA targets. *PLoS Biol.* **2004**, *2*, e363.
45. Krek, A.; Grun, D.; Poy, M.N.; Wolf, R.; Rosenberg, L.; Epstein, E.J.; MacMenamin, P.; da Piedade, I.; Gunsalus, K.C.; Stoffel, M.; *et al.* Combinatorial microRNA target predictions. *Nat. Genet.* **2005**, *37*, 495–500.
46. Lewis, B.P.; Burge, C.B.; Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **2005**, *120*, 15–20.
47. Salmena, L.; Poliseno, L.; Tay, Y.; Kats, L.; Pandolfi, P.P. A ceRNA hypothesis: The Rosetta Stone of a hidden RNA language? *Cell* **2011**, *146*, 353–358.

48. Bicchi, I.; Morena, F.; Montesano, S.; Polidoro, M.; Martino, S. MicroRNAs and molecular mechanisms of neurodegeneration. *Genes* **2013**, *4*, 244–263.
49. Costa, V.; Esposito, R.; Aprile, M.; Ciccodicola, A. Non-coding RNA and pseudogenes in neurodegenerative diseases: "The (un)Usual Suspects". *Front. Genet.* **2012**, *3*, 231.
50. Koscianska, E.; Baev, V.; Skreka, K.; Oikonomaki, K.; Rusinov, V.; Tabler, M.; Kalantidis, K. Prediction and preliminary validation of oncogene regulation by miRNAs. *BMC Mol. Biol.* **2007**, *8*, 79.
51. Baek, D.; Villen, J.; Shin, C.; Camargo, F.D.; Gygi, S.P.; Bartel, D.P. The impact of microRNAs on protein output. *Nature* **2008**, *455*, 64–71.
52. Guo, H.; Ingolia, N.T.; Weissman, J.S.; Bartel, D.P. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **2010**, *466*, 835–840.
53. Cheng, W.C.; Chung, I.F.; Huang, T.S.; Chang, S.T.; Sun, H.J.; Tsai, C.F.; Liang, M.L.; Wong, T.T.; Wang, H.W. YM500: A small RNA sequencing (smRNA-seq) database for microRNA research. *Nucleic Acids Res.* **2013**, *41*, D285–D294.
54. Robins, H.; Press, W.H. Human microRNAs target a functionally distinct population of genes with AT-rich 3' UTRs. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15557–15562.
55. Grimson, A.; Farh, K.K.; Johnston, W.K.; Garrett-Engele, P.; Lim, L.P.; Bartel, D.P. MicroRNA targeting specificity in mammals: Determinants beyond seed pairing. *Mol. Cell* **2007**, *27*, 91–105.
56. Kertesz, M.; Iovino, N.; Unnerstall, U.; Gaul, U.; Segal, E. The role of site accessibility in microRNA target recognition. *Nat. Genet.* **2007**, *39*, 1278–1284.
57. Goswami, S.; Tarapore, R.S.; Teslaa, J.J.; Grinblat, Y.; Setaluri, V.; Spiegelman, V.S. MicroRNA-340-mediated degradation of microphthalmia-associated transcription factor mRNA is inhibited by the coding region determinant-binding protein. *J. Biol. Chem.* **2010**, *285*, 20532–20540.
58. Brennecke, J.; Stark, A.; Russell, R.B.; Cohen, S.M. Principles of microRNA-target recognition. *PLoS Biol.* **2005**, *3*, e85.
59. Ebhardt, H.A.; Tsang, H.H.; Dai, D.C.; Liu, Y.; Bostan, B.; Fahlman, R.P. Meta-analysis of small RNA-sequencing errors reveals ubiquitous post-transcriptional RNA modifications. *Nucleic Acids Res.* **2009**, *37*, 2461–2470.
60. Felice, K.M.; Salzman, D.W.; Shubert-Coleman, J.; Jensen, K.P.; Furneaux, H.M. The 5' terminal uracil of let-7a is critical for the recruitment of mRNA to Argonaute2. *Biochem. J.* **2009**, *422*, 329–341.
61. Mi, S.; Cai, T.; Hu, Y.; Chen, Y.; Hodges, E.; Ni, F.; Wu, L.; Li, S.; Zhou, H.; Long, C.; *et al.* Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. *Cell* **2008**, *133*, 116–127.
62. Koscianska, E.; Starega-Roslan, J.; Czubala, K.; Krzyzosiak, W.J. High-resolution northern blot for a reliable analysis of microRNAs and their precursors. *ScientificWorldJournal* **2011**, *11*, 102–117.
63. Koscianska, E.; Starega-Roslan, J.; Sznajder, L.J.; Olejniczak, M.; Galka-Marciniak, P.; Krzyzosiak, W.J. Northern blotting analysis of microRNAs, their precursors and RNA interference triggers. *BMC Mol. Biol.* **2011**, *12*, 14.

Reprinted from *IJMS*. Cite as: Ma, J.; Yu, S.; Wang, F.; Bai, L.; Xiao, J.; Jiang, Y.; Chen, L.; Wang, J.; Jiang, A.; Li, M.; Li, X. MicroRNA Transcriptomes Relate Intermuscular Adipose Tissue to Metabolic Risk. *Int. J. Mol. Sci.* **2013**, *14*, 8611-8624.

Article

MicroRNA Transcriptomes Relate Intermuscular Adipose Tissue to Metabolic Risk

Jideng Ma [1,†], Shuzhen Yu [1,†], Fengjiao Wang [1], Lin Bai [1], Jian Xiao [1], Yanzhi Jiang [2], Lei Chen [3], Jinyong Wang [3], Anan Jiang [1], Mingzhou Li [1,*] and Xuewei Li [1,*]

[1] Institute of Animal Genetics & Breeding, College of Animal Science & Technology, Sichuan Agricultural University, Ya'an 625014, China; E-Mails: jideng_ma@sina.com (J.M.); yushuzhen1988@126.com (S.Y.); wangfengjiaosicau@gmail.com (F.W.); blin16@126.com (L.B.); jianxiao112@163.com (J.X.); lingdang317@163.com (A.J.)

[2] College of Life and Basic Sciences, Sichuan Agricultural University, Ya'an 625014, China; E-Mail: jiangyz04@163.com

[3] Chongqing Academy of Animal Science, Chongqing 402460, China; E-Mails: sicau.chen@gmail.com (L.C.); kingyou@vip.sina.com (J.W.)

† These authors contributed equally to this work.

* Authors to whom correspondence should be addressed; E-Mails: mingzhou.li@163.com (M.L.); xuewei.li@sicau.edu.cn (X.L.); Tel.: +86-835-288-5991 (M.L.); +86-835-288-6000 (X.L.); Fax: +86-835-288-6080 (M.L. & X.L.).

Received: 20 March 2013; in revised form: 15 April 2013 / Accepted: 17 April 2013 / Published: 22 April 2013

Abstract: Intermuscular adipose tissue is located between the muscle fiber bundles in skeletal muscles, and has similar metabolic features to visceral adipose tissue, which has been found to be related to a number of obesity-related diseases. Although various miRNAs are known to play crucial roles in adipose deposition and adipogenesis, the microRNA transcriptome of intermuscular adipose tissue has not, until now, been studied. Here, we sequenced the miRNA transcriptomes of porcine intermuscular adipose tissue by small RNA-sequencing and compared it to a representative subcutaneous adipose tissue. We found that the inflammation- and diabetes-related miRNAs were significantly enriched

in the intermuscular rather than in the subcutaneous adipose tissue. A functional enrichment analysis of the genes predicted to be targeted by the enriched miRNAs also indicated that intermuscular adipose tissue was associated mainly with immune and inflammation responses. Our results suggest that the intermuscular adipose tissue should be recognized as a potential metabolic risk factor of obesity.

Keywords: intermuscular adipose tissue (IMAT); metabolic risk; miRNA; pig; immune response; inflammation response; obesity; transcriptome

1. Introduction

Adipose tissues (ATs) play a vital role in energy homeostasis and process the largest energy reserve in the body of animals. The rapidly expanding adipokine family is secreted by ATs [1], and, as a result, AT has been identified as an endocrine organ that influences a variety of physiological and pathological processes (such as immunity and inflammation) [2,3] that are involved in the development of metabolic diseases such as cardiovascular disease and type 2 diabetes mellitus [4–6]. Functional and metabolic differences between the visceral and subcutaneous ATs have been well documented. Subcutaneous AT mainly affects metabolic processes, while visceral AT has been identified as a metabolic risk factor for obesity. Recent studies have revealed that the intermuscular adipose tissue (IMAT), which is located between the muscle fiber bundles in skeletal muscles, has similar functional and metabolic features as the visceral ATs [7,8]. Indeed, IMAT was found in greater amounts than visceral AT in acromegaly patients despite their increased muscle mass, suggesting that increased amounts of AT in muscles might be associated with growth hormone-induced insulin resistance [9].

MicroRNAs (miRNAs) are endogenous small non-coding RNAs that modulate gene expression at a post-transcriptional level by binding to the 3' untranslated region (3'-UTR) of the target mRNAs [10]. During the past decade, various miRNAs that play crucial roles in adipose deposition and adipogenesis have been identified. Typically, miR-143 was identified as a pro-adipogenic modulator during pre-adipocyte differentiation [11,12]. MiR-103 [13] and the miR-17-92 cluster [14] were reported to accelerate adipocyte differentiation. MiR-27a [15], miR-27b [16], miR-448 [17] and miR-15a [18] were demonstrated to suppress adipogenic differentiation. MiR-519d [19], miR-335 [20] and miR-377 [21] were associated with lipid metabolism disorders. However, features of the miRNA transcriptome of IMAT have yet to be investigated.

Sus scrofa (pig) is emerging as an ideal biomedical model for obesity and metabolic disorders in human because of the similarity in metabolic features and proportional organ sizes in these two species [22]. To decipher the unique metabolic and functional features of IMAT, we sequenced the miRNA transcriptomes of porcine IMAT by small RNA-sequencing and compared it with a

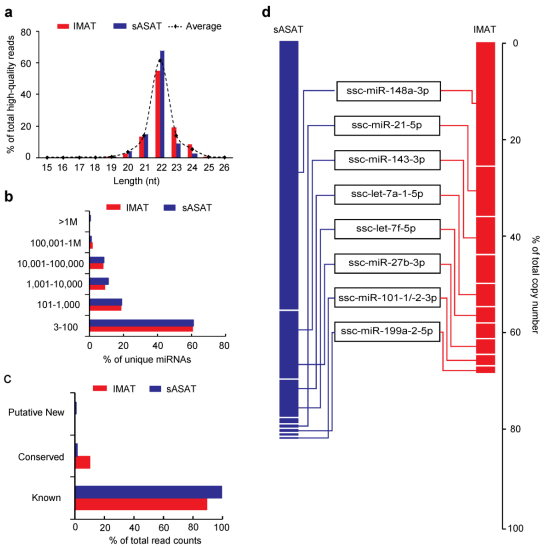
representative subcutaneous adipose tissue, superficial abdominal subcutaneous adipose tissue (sASAT). We identified various known, conserved, and putative novel porcine miRNAs in these two tissues. Notably, the sASAT-enriched miRNAs were related mainly to lipid metabolic homeostasis, while the IMAT-enriched miRNAs were related mainly to inflammation and diabetes, and the target genes of the IMAT-enriched miRNAs were primarily associated with inflammatory and diabetes processes. Together, these findings indicated the metabolic risk of IMAT. Our results will contribute to studies into the role of IMAT in obesity-related metabolic disease.

2. Results and Discussion

2.1. Transcriptome Sequencing Data

We used a small RNA-sequencing approach to sequence the miRNA transcriptomes of porcine IMAT and sASAT and obtained 17.76 million (M) and 18.50 M raw reads, respectively. More than 80% of the raw reads passed the quality filters (see Methods) and were termed the high-quality reads (IMAT: 14.93 M, 84.09%; sASAT: 14.92 M, 81.59%) (Supplementary Table S1). The high-quality reads in both transcriptomes exhibited the canonical size range distribution that is common to mammalian miRNAs (Figure 1a). The vast majority of the reads were 21–23 nucleotides (nt) in length. The 22-nt reads accounted for 61.50% of all the high-quality reads, followed by the 21-nt (14.00%) and 23-nt (13.99%) reads. This result indicated the reliability of using the small RNA-sequencing approach to generate miRNA reads as candidates for further analysis.

Figure 1. Description of miRNAs in two adipose tissues. (**a**) Length distribution of the high-quality reads; (**b**) Distribution of read counts of the identified miRNAs; (**c**) Distribution of read counts in the three defined miRNA groups; (**d**) Copy numbers of the top 10 miRNAs with highest read counts. IMAT, intermuscular adipose tissue; sASAT, superficial abdominal subcutaneous adipose tissue.



2.2. MiRNA Profiling of sASAT and IMAT

A total of 597 mature miRNAs corresponding to 453 miRNA precursors (pre-miRNAs) were identified in the two libraries by mapping them to the pig genome. In agreement with previous reports [23,24], we found that all the miRNA classes consisted of multiple mature variants (the isomiRs). The most abundant isomiR in each class was picked as the reference sequence for that class [23] based on the evidence that there was a significant positive correlation between the counts of most abundant isomiR and the total counts of all isomiRs in the same class (IMAT: Spearman's $r = 0.98$, $p < 10^{-5}$; sASAT: Spearman's $r = 0.97$, $p < 10^{-4}$).

The identified mature miRNAs and their precursors were divided into three subgroups according to alignment criteria (Supplementary Table S2) as: (1) Porcine known miRNAs: 297 miRNAs mapped to 176 known porcine pre-miRNAs; specifically, 210 were in miRBase 18.0 [25] and 87 were novel miRNA*s; (2) Porcine conserved miRNAs: 107 miRNAs mapped to 71 other known mammalian pre-miRNAs in miRBase 18.0 and these pre-miRNAs mapped to the pig genome. These miRNAs were labeled with the names of the corresponding conserved miRNAs; (3) Porcine putative new miRNAs: 230 miRNAs (longer than 18 nt and unmapped to any known mammalian pre-miRNAs in miRBase 18.0) encompassing 206 candidate pre-miRNAs that were predicted RNA hairpins derived from the pig genome, and were labeled PPN (Porcine putative new). Notably, there are the distinct pre-miRNAs coding the identical mature miRNAs, which resulting in 617 miRNAs (*i.e.*, reference sequence) corresponding to 597 unique miRNA sequences (Supplementary Table S3).

The identified miRNAs exhibited a large dynamic range of read counts ranging from 3 to millions. The vast majority of miRNAs (IMAT: 61.11%; sASAT: 61.03%) were in low abundance (3 to 100 read counts) and belonged mainly to the porcine conserved and putative new miRNA groups. Only a few miRNAs (IMAT: 2.11%; sASAT: 1.17%) were in high abundance (>100,000 read counts) and they belonged mainly to the porcine known miRNA group (Figure 1b,c). This result suggests that the low-abundance conserved and putative new miRNAs may have escaped from previous detection efforts.

We found that the top ten miRNAs with the highest abundance contributed 67.46% and 82.94% of the total counts in the IMAT and sASAT libraries, respectively, and eight miRNAs were shared by two libraries in the top 10 positions (Figure 1d). The high abundance of these miRNAs implies that they may have housekeeping cellular roles and may be the main regulatory miRNAs in adipogenesis [11,26,27] and cellular basal metabolism [28,29]. For example, let-7a-5p [12], miR-148a-3p [26], miR-21-5p [27], miR-143-3p [30] and miR-101-3p [13] have been reported to be up-regulated during 3T3-L1 pre-adipocyte differentiation, whereas miR-27b-3p was found to be down-regulated during adipogenesis of human multipotent adipose-derived stem cells [31] and miR-199a-5p was up-regulated in subcutaneous AT in obese versus non-obese individuals [13].

2.3. Inflammation- and Diabetes-Related miRNAs Enriched in IMAT

More than half of the unique miRNAs (351 of 597, 58.79%) were co-expressed in IMAT and sASAT. Only 171 (28.64%) and 75 (12.56%) of the unique miRNAs were expressed specifically in IMAT and sASAT, respectively (Figure 2a and Supplementary Table S3). It was well-known that miRNAs function in a dose-dependent manner [32], thus the less abundant miRNAs (<1000 read counts in both libraries) were considered to be less important and were filtered out. Of the 110 more abundant unique miRNAs (>1000 read counts in either library), 53 (48.18%) were determined to be differentially expressed (DE) between IMAT and sASAT using the IDEG6 program [33] (Figure 2a and Supplementary Table S4). The changes in expression patterns of the top 14 DE miRNAs with the highest read counts showed significant positive correlations between the q-PCR results and the small RNA-sequencing data (Person's $r = 0.894$, $p < 10^{-4}$), again highlighting the reliability of the small RNA-sequencing approach (Figure 2b). Moreover, in the process of q-PCR validation, we also found that all expression levels of selected miRNAs obtained by q-PCR within the biological replicates were highly correlated and with very low deviation, which not only indicated the high repeatability and reliability of the q-PCR approach but also reflected the high purity of our experimental samples (Supplementary Table S5).

Figure 2. Characteristics of the differentially expressed (DE) miRNAs between porcine sASAT and IMAT. (**a**) Distribution of 597 unique miRNAs between sASAT (blue) and IMAT (yellow). The red circle represents the 110 miRNAs with read counts >1000 in either of the two libraries. The dashed circles indicate the 45 IMAT-enriched (**left**) and eight sASAT-enriched (**right**) miRNAs ($p < 0.001$); (**b**) q-PCR validation for the top 14 DE miRNAs with highest read counts between IMAT and sASAT. Pearson's correlation was used to determine the relationship between the q-PCR and small RNA-seq results for miRNA expression levels. IMAT-NE and sASAT-NE represent normalized expression levels for the miRNAs in the IMAT and sASAT libraries, respectively; (**c**) The differential expression of 19 inflammation- and diabetes-related miRNAs between IMAT and sASAT.

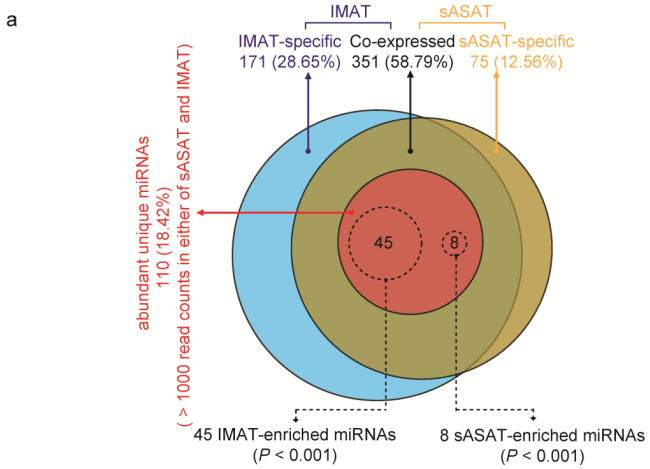
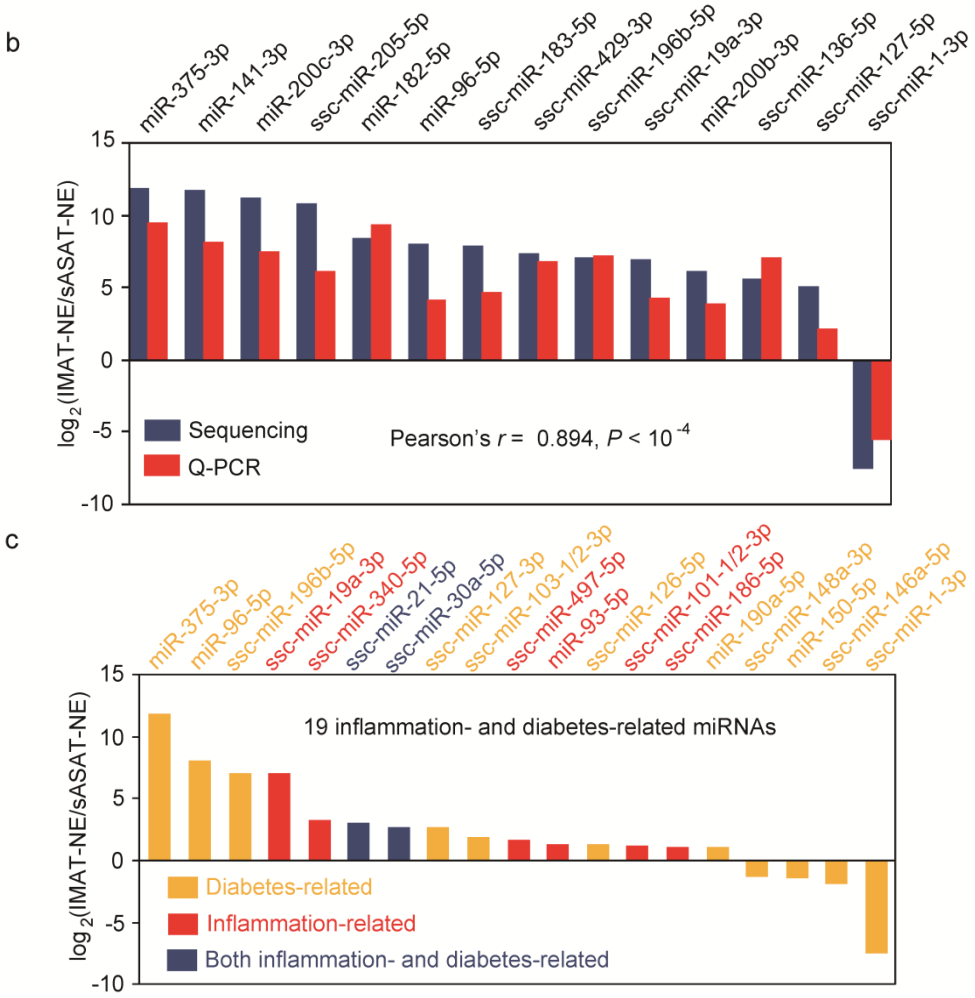


Figure 2. *Cont.*



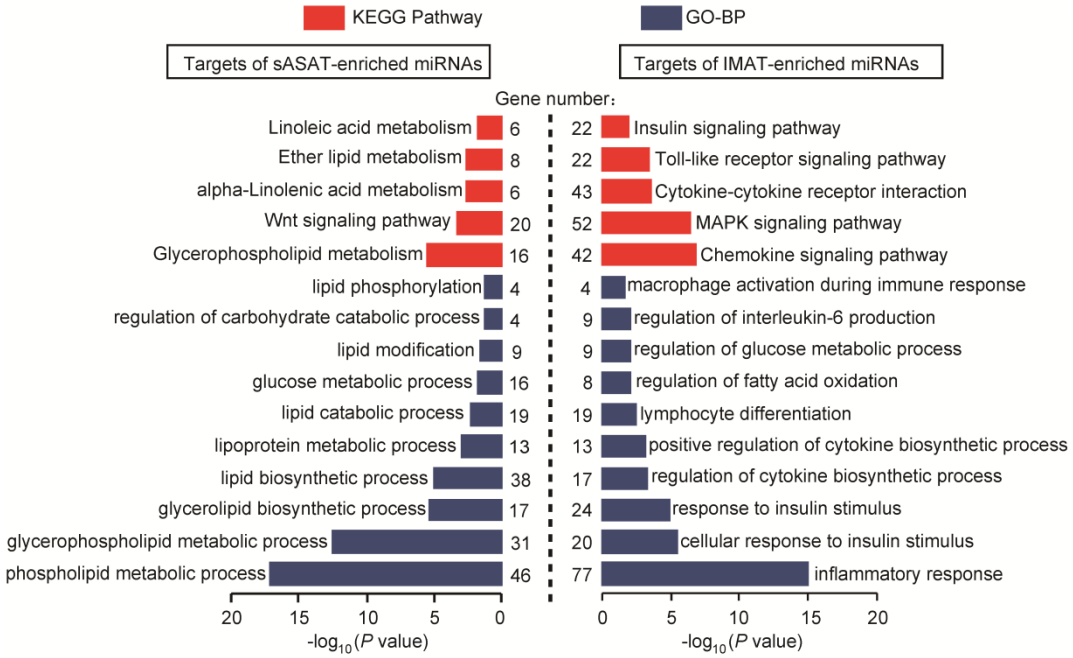
Notably, many of the DE miRNAs (19 out of 53, 35.84%) were associated with inflammation and diabetes based on the annotations assigned using the Pathway Central database (SA Biosciences, Frederick, MD, USA) (Figure 2c). Eight inflammation-related and 9 diabetes-related miRNAs were present in higher abundance in the IMAT transcriptome compared with the sASAT transcriptome (Figure 2c). MiR-21 was found to be over-expressed at the inflammation site [34] and it has been suggested that miR-21 could act as a biomarker for inflammation in the aging process and cardiovascular disease [35]. MiR-101 was reported to be related to inflammation and chondrocyte extracellular matrix degradation [36]. Circulating miR-30a was up-regulated in diabetes patients and has been associated with insulin resistance [20]. The ectopic high expression of miR-103 could induce impaired glucose homeostasis or, conversely, the silencing of miR-103 could improve glucose homeostasis and insulin sensitivity [37].

In addition, various IMAT-enriched miRNAs related to pathological responses were found (Supplementary Table S4). For example, miR-182 and miR-183 (members of the miR-183-96-182 cluster) are well-characterized oncomiRs that can promote the clonal expansion of activated helper T

lymphocytes [38,39]. MiR-200b, miR-200c and miR-141 (members of the miR-200 family) were reported to be significantly altered in bladder [40] and breast cancers [41]. MiR-191 was suggested as a biomarker for the diagnosis and prognosis of acute myeloid leukemia [42]. These results suggest that IMAT is associated mainly with inflammation- and diabetes-related responses, and should be deemed as a potential metabolic risk factor of obesity.

In contrast, the sASAT-enriched miRNAs (Supplementary Table S4) were mainly related to adipogenesis and lipid metabolism. For example, miR-378 was up-regulated in adipogenesis of human AT-derived stromal cells [43] and the over-expression of miR-378 in ST2 cell line was reported to promote lipid accumulation by enhancing *de novo* lipogenesis [44]. MiR-365 was revealed as a central regulator of brown fat differentiation and adipogenesis [45]. MiR-146a regulated mainly lipid accumulation induced by oxidized low-density lipoprotein [46].

Figure 3. KEGG pathways and gene ontology biological process (GO-BP) categories enriched in the target genes of the top eight sASAT- and IMAT-enriched miRNAs. GO-BP is the GO terms under the biological process ontology.



2.4. Functional Enrichment Analyses of miRNA Target Genes

To further highlight the distinct functional features of IMAT and sASAT, the target genes of the top eight DE miRNAs enriched in sASAT (967 mRNA genes) and IMAT (1707 mRNA genes) were predicted using PicTar [47], TargetScan human 6.2 [48] and MicroCosm Targets (version 5.0) [49] (Supplementary Table S6), and analyzed using DAVID [50] to determine whether or not they were enriched for specific functional categories and pathways. Similar to the finding for the DE miRNAs, the target genes of the IMAT-enriched miRNAs were primarily associated with inflammatory and

diabetes-related processes, such as "inflammatory response" (77 genes, $p = 1.21 \times 10^{-15}$), "cellular response to insulin stimulus" (20 genes, $p = 4.14 \times 10^{-6}$), "lymphocyte differentiation" (19 genes, $p = 3.75 \times 10^{-3}$), "regulation of interleukin-6 production" (9 genes, $p = 1.17 \times 10^{-2}$), "macrophage activation during immune response" (4 genes, $p = 2.74 \times 10^{-2}$), "chemokine and toll-like signaling pathways" (42 genes, $p = 1.61 \times 10^{-7}$) and "insulin signaling pathway" (22 genes, $p = 1.51 \times 10^{-2}$). In contrast, the target genes of the sASAT-enriched miRNAs were mainly associated with lipid and energy metabolism, such as "glycerophospholipid metabolic process" (46 genes, $p = 5.10 \times 10^{-18}$), "lipid biosynthetic process" (38 genes, $p = 7.36 \times 10^{-6}$), "glucose metabolic process" (16 genes, $p = 1.53 \times 10^{-2}$) and "Wnt signaling pathway" (20 genes, $p = 5.27 \times 10^{-4}$) (Figure 3). These results further suggested that while sASAT is mainly involved in metabolic homeostasis, IMAT is susceptive to inflammation and should be regarded as a potential metabolic risk factor.

3. Experimental Section

3.1. Animals and Sample Collection

Three 210-day-old female Landrace pigs with normal weight (111.67 ± 1.15 kg) were used. The piglets were weaned simultaneously at 28 ± 1 day of age. A starter diet provided 3.40 Mcal·kg^{-1} metabolisable energy (ME), 20.00% crude protein and 1.15% lysine from the thirtieth to sixtieth day after weaning. From the 61st to the 120th day, the diet contained 3.40 Mcal·kg^{-1} ME, 17.90% crude protein and 0.83% lysine. From the 121st to 210th day, the diet contained 3.40 Mcal kg^{-1} ME, 15.00% crude protein and 1.15% lysine. The animals were allowed access to feed and water *ad libitum* and lived under the same normal conditions.

The macroscopic IMAT were directly separated from the regions that were beneath the biceps femoris muscle fascia of porcine hind leg. Since IMAT preparation could be easily contaminated by its surrounding tissues, we paid maximum attention to eliminate the others especially, such as connective tissue and muscle tissue, and all samples were resected from central part of tissue block. The sASAT were from the subcutaneous tissue of central abdomen near the last rib. All samples were immediately frozen in liquid nitrogen and stored at −80 °C before total RNA extraction.

3.2. Small RNA Libraries Construction and High-throughput Sequencing

Total RNA was extracted using the *mir*Vana™ miRNA isolation kit (Ambion, Austin, USA) following the manufacturer's protocol. The integrity of total RNA was also tested via analysis by Bioanalyzer 2100 and RNA 6000 Nano LabChip Kit (Agilent, Palo Alto, CA, USA) with RIN number >6.0.

For a certain adipose tissue, equal amounts (5 μg) of total RNA isolated from three pigs were mixed. Approximately 15 μg of small RNA-enriched total RNA was prepared for Illumina sequencing. In general, the processing by Illumina consisted of the following successive steps: the small RNA ranged from 14 to 40 nt were purified by polyacrylamide gel electrophoresis (PAGE) and ligated

specific adapters followed by polyacrylamide gel purification. Then the modified small RNA was reverse transcripted and amplified by RT-PCR. Finally, the enriched cDNA was sequenced on Genome Analyzer Instrument (GAI, Illumina, San Diego, CA, USA). The small RNA-sequencing data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE30334.

3.3. Analysis of Small RNA-Sequencing Data

The raw reads were processed using Illumina's Genome Analyzer Pipeline software and subsequently handled as described by Li *et al*. with some improvement [23]. After trimming off the sequencing adapters, the resulting reads was successively filtered by read length (only the read with the of 14 to 27 nt were retained), sequence component (containing <80% A, C, G or T; containing no more than two N (undetermined bases)) and copy numbers (the low-abundance reads (only the read with >3 counts were retained). Then the retained reads were searched against the NCBI [51], Rfam [52] and Repbase database [53] to remove porcine known classes of RNAs (*i.e.*, mRNA, rRNA, tRNA, snRNA, snoRNA and repeats). The sequencing reads survived from above strict filter rules were deemed as "high-quality reads".

The high-quality reads were mapped to the pig genome (Sscrofa9) using NCBI Local BLAST following five steps in order: (1) map the high-quality reads to the 228 known porcine pre-miRNAs (encoding 257 miRNAs) and then to 6716 known pre-miRNAs (encoding 7952 miRNAs) from 24 other mammals in miRBase 18.0; (2) map the mapped high-quality reads to pig genome to obtain their genomic locations and annotations in Ensembl release 59 (Sscrofa 9, April 2009); (3) cluster the unmapped sequences in step 1 that mapped to the pig genome as putative novel miRNAs; and (4) predict hairpin RNA structures of the high-quality reads in step 3 from the adjacent 60 nt sequences in either direction from the pig genome using UNAFold [54]. To avoid ambiguous reads that have been assigned to multiple positions in pig genome, only reads longer than 18 nt in length were included in step 4.

3.4. miRNA Differential Expression Analysis

Program IDEG6 [33] was employed for detecting the DE miRNAs between two libraries. A unique miRNA is considered to be differentially expressed when it simultaneously obtains $p < 0.001$ under three statistical tests (a Audic-Claverie test, a Fisher exact test and a Chi-squared 2×2 test) with the Bonferroni correction.

3.5. Prediction and Functional Annotation of miRNA Target Genes

The potential targets of a certain miRNA were predicted by PicTar [47], TargetScan human 6.2 [48] and, MicroCosm Targets Version 5.0 [49], and the pairwise overlaps of results from three programs composed the final predicted targets. The predictions were according to the interactions of human

mRNA-miRNA due to the absence of porcine miRNAs in current version of above-mentioned algorithm. The gene ontology biological process (GO-BP) terms and KEGG pathway terms enriched in predicted target genes were determined using a DAVID bioinformatics resources [50].

3.6. Q-PCR Validation

The expression changes of 14 selected miRNAs were validated by an EvaGreen-based High-Specificity miRNA qRT-PCR Detection Kit (Stratagene, La Jolla, CA, USA) on the CFX96™ Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA). The q-PCR validation were carried out on three biological replicates.The primer pairs were available in Supplementary Table S7. Three endogenous control genes (U6 snRNA, 18S rRNA and Met-tRNA) [23] were used in this assay. The ΔΔCt method was used to determine the expression level differences between surveyed samples. Normalized factors (NF) of three endogenous control genes and relative quantities of objective miRNAs were analyzed using the qBase software [55].

4. Conclusions

We have generated reliable miRNA transcriptomes of porcine sASAT and IMAT, and identified many known and novel miRNAs using small RNA-sequencing approach. We found that inflammation- and diabetes-related miRNAs were enriched in IMAT compared with sASAT, which indicated the metabolic risk of the IMAT. A functional enrichment analysis of genes targeted by the enriched miRNAs also indicated that IMAT was mainly associated with the immune and inflammation response and may be a potential metabolic risk factor of obesity. The current study provides data that can be used in future studies to investigate the metabolic role of IMAT in obesity-related metabolic dysfunction. Our findings will also help promote the further development of the pig model for human metabolic research. It is also worth noting that further detailed comparision of IMAT between obese and non-obese individuals will be necessary and beneficial to decipher the role of miRNAs in adipogenesis and IMAT-related metabolic diseases.

Acknowledgments

This work was supported by grants from the National High Technology Research and Development Program of China (863 Program) (2013AA102502), the Specialized Research Fund of Ministry of Agriculture of China (NYCYTX-009), the Project of Provincial Twelfth Five Years' Animal Breeding of Sichuan Province (2011YZGG15), and the National Special Foundation for Transgenic Species of China (2011ZX08006-003) to X.L. and M.L., the Chongqing Fund for Distinguished Young Scientists (CSTC2010BA1007).

Conflict of Interest

The authors declare no conflict of interest.

References

1. Zhang, Y.; Proenca, R.; Maffei, M.; Barone, M.; Leopold, L.; Friedman, J.M. Positional cloning of the mouse obese gene and its human homologue. *Nature* **1994**, *372*, 425–432.
2. Kadowaki, T.; Yamauchi, T.; Kubota, N.; Hara, K.; Ueki, K.; Tobe, K. Adiponectin and adiponectin receptors in insulin resistance, diabetes, and the metabolic syndrome. *J. Clin. Invest.* **2006**, *116*, 1784–1792.
3. Lago, F.; Dieguez, C.; Gómez-Reino, J.; Gualillo, O. The emerging role of adipokines as mediators of inflammation and immune responses. *Cytokine Growth Factor Rev.* **2007**, *18*, 313–325.
4. Arner, P. Insulin resistance in type 2 diabetes-role of the adipokines. *Curr. Mol. Med.* **2005**, *5*, 333–339.
5. Dogru, T.; Sonmez, A.; Tasci, I.; Bozoglu, E.; Yilmaz, M.I.; Genc, H.; Erdem, G.; Gok, M.; Bingol, N.; Kilic, S. Plasma visfatin levels in patients with newly diagnosed and untreated type 2 diabetes mellitus and impaired glucose tolerance. *Diabetes Res. Clin. Pract. Suppl.* **2007**, *76*, 24–29.
6. Berg, A.H.; Scherer, P.E. Adipose tissue, inflammation, and cardiovascular disease. *Circ. Res.* **2005**, *96*, 939–949.
7. Gallagher, D.; Kuznia, P.; Heshka, S.; Albu, J.; Heymsfield, S.B.; Goodpaster, B.; Visser, M.; Harris, T.B. Adipose tissue in muscle: A novel depot similar in size to visceral adipose tissue. *Am. J. Clin. Nutr.* **2005**, *81*, 903–910.
8. Boettcher, M.; Machann, J.; Stefan, N.; Thamer, C.; Häring, H.U.; Claussen, C.D.; Fritsche, A.; Schick, F. Intermuscular adipose tissue (IMAT): Association with other adipose tissue compartments and insulin sensitivity. *J. Magn. Reson. Imaging* **2009**, *29*, 1340–1345.
9. Freda, P.U.; Shen, W.; Heymsfield, S.B.; Reyes-Vidal, C.M.; Geer, E.B.; Bruce, J.N.; Gallagher, D. Lower visceral and subcutaneous but higher intermuscular adipose tissue depots in patients with growth hormone and insulin-like growth factor I excess due to acromegaly. *J. Clin. Endocrinol. Metab.* **2008**, *93*, 2334–2343.
10. Nelson, P.; Kiriakidou, M.; Sharma, A.; Maniataki, E.; Mourelatos, Z. The microRNA world: Small is mighty. *Trends Biochem. Sci.* **2003**, *28*, 534–540.
11. Xie, H.; Lim, B.; Lodish, H.F. MicroRNAs induced during adipogenesis that accelerate fat cell development are downregulated in obesity. *Diabetes* **2009**, *58*, 1050–1057.
12. Kajimoto, K.; Naraba, H.; Iwai, N. MicroRNA and 3T3-L1 pre-adipocyte differentiation. *RNA* **2006**, *12*, 1626–1632.
13. Ortega, F.J.; Moreno-Navarrete, J.M.; Pardo, G.; Sabater, M.; Hummel, M.; Ferrer, A.; Rodriguez-Hermosa, J.I.; Ruiz, B.; Ricart, W.; Peral, B. MiRNA expression profile of human subcutaneous adipose and during adipocyte differentiation. *PLoS One* **2010**, *5*, e9022.
14. Wang, Q.; Li, Y.C.; Wang, J.; Kong, J.; Qi, Y.; Quigg, R.J.; Li, X. miR-17-92 cluster accelerates adipocyte differentiation by negatively regulating tumor-suppressor Rb2/p130. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 2889–2894.

15. Lin, Q.; Gao, Z.; Alarcon, R.M.; Ye, J.; Yun, Z. A role of miR-27 in the regulation of adipogenesis. *FEBS J.* **2009**, *276*, 2348–2358.
16. Kim, S.Y.; Kim, A.Y.; Lee, H.W.; Son, Y.H.; Lee, G.Y.; Lee, J.-W.; Lee, Y.S.; Kim, J.B. miR-27a is a negative regulator of adipocyte differentiation via suppressing *PPARγ* expression. *Biochem. Biophys. Res. Commun.* **2010**, *392*, 323–328.
17. Kinoshita, M.; Ono, K.; Horie, T.; Nagao, K.; Nishi, H.; Kuwabara, Y.; Takanabe-Mori, R.; Hasegawa, K.; Kita, T.; Kimura, T. Regulation of adipocyte differentiation by activation of serotonin (5-HT) receptors *5-HT2AR* and *5-HT2CR* and involvement of microRNA-448-mediated repression of *KLF5*. *Mol. Endocrinol.* **2010**, *24*, 1978–1987.
18. Andersen, D.C.; Jensen, C.H.; Schneider, M.; Nossent, A.Y.; Eskildsen, T.; Hansen, J.L.; Teisner, B.; Sheikh, S.P. MicroRNA-15a fine-tunes the level of Delta-like 1 homolog (*DLK1*) in proliferating 3T3-L1 preadipocytes. *Exp. Cell Res.* **2010**, *316*, 1681–1691.
19. Martinelli, R.; Nardelli, C.; Pilone, V.; Buonomo, T.; Liguori, R.; Castanò, I.; Buono, P.; Masone, S.; Persico, G.; Forestieri, P. miR-519d overexpression is associated with human obesity. *Obesity* **2012**, *18*, 2170–2176.
20. Ferland-McCollough, D.; Ozanne, S.; Siddle, K.; Willis, A.; Bushell, M. The involvement of microRNAs in Type 2 diabetes. *Biochem. Soc. Trans.* **2010**, *38*, 1565.
21. Wang, Q.; Wang, Y.; Minto, A.W.; Wang, J.; Shi, Q.; Li, X.; Quigg, R.J. MicroRNA-377 is up-regulated and can lead to increased fibronectin production in diabetic nephropathy. *FASEB J.* **2008**, *22*, 4126–4135.
22. Spurlock, M.E.; Gabler, N.K. The development of porcine models of obesity and the metabolic syndrome. *J. Nutr.* **2008**, *138*, 397–402.
23. Li, M.; Xia, Y.; Gu, Y.; Zhang, K.; Lang, Q.; Chen, L.; Guan, J.; Luo, Z.; Chen, H.; Li, Y. MicroRNAome of porcine pre-and postnatal development. *PLoS One* **2010**, *5*, e11541.
24. Xie, S.-S.; Li, X.-Y.; Liu, T.; Cao, J.-H.; Zhong, Q.; Zhao, S.-H. Discovery of porcine microRNAs in multiple tissues by a Solexa deep sequencing approach. *PLoS One* **2011**, *6*, e16235.
25. Kozomara, A.; Griffiths-Jones, S. miRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **2011**, *39*, D152–D157.
26. Qin, L.; Chen, Y.; Niu, Y.; Chen, W.; Wang, Q.; Xiao, S.; Li, A.; Xie, Y.; Li, J.; Zhao, X. A deep investigation into the adipogenesis mechanism: Profile of microRNAs regulating adipogenesis by modulating the canonical Wnt/β-catenin signaling pathway. *BMC Genomics* **2010**, *11*, 320.
27. Kim, Y.J.; Hwang, S.J.; Bae, Y.C.; Jung, J.S. MiR-21 regulates adipogenic differentiation through the modulation of *TGF-β* signaling in mesenchymal stem cells derived from human adipose tissue. *Stem Cells* **2009**, *27*, 3093–3102.
28. Hoekstra, M.; van der Lans, C.A.; Halvorsen, B.; Gullestad, L.; Kuiper, J.; Aukrust, P.; van Berkel, T.J.; Biessen, E.A. The peripheral blood mononuclear cell microRNA signature of coronary artery disease. *Biochem. Biophys. Res. Commun.* **2010**, *394*, 792–797.

29. Ro, S.; Park, C.; Young, D.; Sanders, K.M.; Yan, W. Tissue-dependent paired expression of miRNAs. *Nucleic Acids Res.* **2007**, *35*, 5944–5953.
30. Esau, C.; Kang, X.; Peralta, E.; Hanson, E.; Marcusson, E.G.; Ravichandran, L.V.; Sun, Y.; Koo, S.; Perera, R.J.; Jain, R. MicroRNA-143 regulates adipocyte differentiation. *J. Biol. Chem.* **2004**, *279*, 52361–52365.
31. Karbiener, M.; Fischer, C.; Nowitsch, S.; Opriessnig, P.; Papak, C.; Ailhaud, G.; Dani, C.; Amri, E.-Z.; Scheideler, M. microRNA miR-27b impairs human adipocyte differentiation and targets *PPARγ*. *Biochem. Biophys. Res. Commun.* **2009**, *390*, 247–251.
32. Carlsbecker, A.; Lee, J.-Y.; Roberts, C.J.; Dettmer, J.; Lehesranta, S.; Zhou, J.; Lindgren, O.; Moreno-Risueno, M.A.; Vatén, A.; Thitamadee, S. Cell signalling by microRNA165/6 directs gene dose-dependent root cell fate. *Nature* **2010**, *465*, 316–321.
33. Romualdi, C.; Bortoluzzi, S.; d'Alessi, F.; Danieli, G.A. IDEG6: A web tool for detection of differentially expressed genes in multiple tag sampling experiments. *Physiol. Genomics* **2003**, *12*, 159–162.
34. Okayama, H.; Schetter, A.; Harris, C. MicroRNAs and inflammation in the pathogenesis and progression of colon cancer. *Dig. Dis.* **2012**, *30*, 9–15.
35. Olivieri, F.; Spazzafumo, L.; Santini, G.; Lazzarini, R.; Albertini, M.C.; Rippo, M.R.; Galeazzi, R.; Abbatecola, A.M.; Marcheselli, F.; Monti, D. Age-related differences in the expression of circulating microRNAs: miR-21 as a new circulating marker of inflammaging. *Mech. Ageing Dev.* **2012**, *133*, 675–685.
36. Dai, L.; Zhang, X.; Hu, X.; Zhou, C.; Ao, Y. Silencing of microRNA-101 prevents IL-1b-induced extracellular matrix degradation in chondrocytes. *Arthritis Res. Ther.* **2012**, *14*, R268.
37. Trajkovski, M.; Hausser, J.; Soutschek, J.; Bhat, B.; Akin, A.; Zavolan, M.; Heim, M.H.; Stoffel, M. MicroRNAs 103 and 107 regulate insulin sensitivity. *Nature* **2011**, *474*, 649–653.
38. Stittrich, A.-B.; Haftmann, C.; Sgouroudis, E.; Kühl, A.A.; Hegazy, A.N.; Panse, I.; Riedel, R.; Flossdorf, M.; Dong, J.; Fuhrmann, F. The microRNA miR-182 is induced by IL-2 and promotes clonal expansion of activated helper T lymphocytes. *Nat. Immunol.* **2010**, *11*, 1057–1062.
39. Sarver, A.L.; Li, L.; Subramanian, S. MicroRNA miR-183 functions as an oncogene by targeting the transcription factor *EGR1* and promoting tumor cell migration. *Cancer Res.* **2010**, *70*, 9570–9580.
40. Wiklund, E.D.; Bramsen, J.B.; Hulf, T.; Dyrskjøt, L.; Ramanathan, R.; Hansen, T.B.; Villadsen, S.B.; Gao, S.; Ostenfeld, M.S.; Borre, M. Coordinated epigenetic repression of the miR-200 family and miR-205 in invasive bladder cancer. *Int. J. Cancer* **2011**, *128*, 1327–1334.
41. Isacke, C. MicroRNA-200 family modulation in distinct breast cancer phenotypes. *PLoS One* **2012**, *7*, e47709.
42. Garzon, R.; Volinia, S.; Liu, C.-G.; Fernandez-Cymering, C.; Palumbo, T.; Pichiorri, F.; Fabbri, M.; Coombes, K.; Alder, H.; Nakamura, T. MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia. *Blood* **2008**, *111*, 3183–3189.

43. Zaragosi, L.-E.; Wdziekonski, B.; Brigand, K.L.; Villageois, P.; Mari, B.; Waldmann, R.; Dani, C.; Barbry, P. Small RNA sequencing reveals miR-642a-3p as a novel adipocyte-specific microRNA and miR-30 as a key regulator of human adipogenesis. *Genome Biol.* **2011**, *12*, R64.
44. Gerin, I.; Bommer, G.T.; McCoin, C.S.; Sousa, K.M.; Krishnan, V.; MacDougald, O.A. Roles for miRNA-378/378* in adipocyte gene expression and lipogenesis. *Am. J. Physiol. Endocrinol. Metab.* **2010**, *299*, E198–E206.
45. Sun, L.; Xie, H.; Mori, M.A.; Alexander, R.; Yuan, B.; Hattangadi, S.M.; Liu, Q.; Kahn, C.R.; Lodish, H.F. Mir193b-365 is essential for brown fat differentiation. *Nat. Cell Biol.* **2011**, *13*, 958–965.
46. Yang, K.; He, Y.S.; Wang, X.Q.; Lu, L.; Chen, Q.J.; Liu, J.; Sun, Z.; Shen, W.F. MiR-146a inhibits oxidized low-density lipoprotein-induced lipid accumulation and inflammatory response via targeting toll-like receptor 4. *FEBS Lett.* **2011**, *585*, 854–860.
47. Krek, A.; Grün, D.; Poy, M.N.; Wolf, R.; Rosenberg, L.; Epstein, E.J.; MacMenamin, P.; da Piedade, I.; Gunsalus, K.C.; Stoffel, M. Combinatorial microRNA target predictions. *Nat. Genet.* **2005**, *37*, 495–500.
48. Lewis, B.P.; Burge, C.B.; Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **2005**, *120*, 15–20.
49. Griffiths-Jones, S.; Saini, H.K.; van Dongen, S.; Enright, A.J. miRBase: Tools for microRNA genomics. *Nucleic Acids Res.* **2008**, *36*, D154–D158.
50. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2008**, *4*, 44–57.
51. Pruitt, K.D.; Tatusova, T.; Klimke, W.; Maglott, D.R. NCBI Reference Sequences: Current status, policy and new initiatives. *Nucleic Acids Res.* **2009**, *37*, D32–D36.
52. Gardner, P.P.; Daub, J.; Tate, J.G.; Nawrocki, E.P.; Kolbe, D.L.; Lindgreen, S.; Wilkinson, A.C.; Finn, R.D.; Griffiths-Jones, S.; Eddy, S.R.; *et al*. Rfam: Updates to the RNA families database. *Nucleic Acids Res.* **2009**, *37*, D136–D140.
53. Kohany, O.; Gentles, A.J.; Hankus, L.; Jurka, J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinforma.* **2006**, *7*, 474.
54. Markham, N.R.; Zuker, M. UNAFold: Software for nucleic acid folding and hybridization. *Methods Mol. Biol.* **2008**, *453*, 3–31.
55. Hellemans, J.; Mortier, G.; de Paepe, A.; Speleman, F.; Vandesompele, J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* **2007**, *8*, R19.

Reprinted from *IJMS*. Cite as: Maegdefessel, L.; Spin, J.M.; Adam, M.; Raaz, U.; Toh, R.; Nakagami, F.; Tsao, P.S. Micromanaging Abdominal Aortic Aneurysms. *Int. J. Mol. Sci.* **2013**, *14*, 14374-14394.

Review

Micromanaging Abdominal Aortic Aneurysms

Lars Maegdefessel [1,†], Joshua M. Spin [2,†], Matti Adam [2], Uwe Raaz [2], Ryuji Toh [2], Futoshi Nakagami [2] and Philip S. Tsao [2,*]

[1] Department of Medicine, Karolinska Institute, Stockholm SE-17176, Sweden; E-Mail: lars.maegdefessel@ki.se

[2] Division of Cardiovascular Medicine, Stanford University, Stanford, CA 94305-5406, USA; E-Mails: josh.spin@gmail.com (J.M.S.); matti.adam@stanford.edu (M.A.); uraaz@stanford.edu (U.R.); rtoh1214@hotmail.com (R.T.); saruwaku@gmail.com (F.N.)

[†] These authors contributed equally to this work.

[*] Author to whom correspondence should be addressed; E-Mail: ptsao@stanford.edu; Tel.: +1-650-498-6317; Fax: +1-650-725-2178.

Received: 26 April 2013; in revised form: 25 June 2013 / Accepted: 26 June 2013 / Published: 11 July 2013

Abstract: The contribution of abdominal aortic aneurysm (AAA) disease to human morbidity and mortality has increased in the aging, industrialized world. In response, extraordinary efforts have been launched to determine the molecular and pathophysiological characteristics of the diseased aorta. This work aims to develop novel diagnostic and therapeutic strategies to limit AAA expansion and, ultimately, rupture. Contributions from multiple research groups have uncovered a complex transcriptional and post-transcriptional regulatory milieu, which is believed to be essential for maintaining aortic vascular homeostasis. Recently, novel small noncoding RNAs, called microRNAs, have been identified as important transcriptional and post-transcriptional inhibitors of gene expression. MicroRNAs are thought to "fine tune" the translational output of their target messenger RNAs (mRNAs) by promoting mRNA degradation or inhibiting translation. With the discovery that microRNAs act as powerful regulators in the context of a wide variety of diseases, it is only logical that microRNAs be thoroughly explored as potential therapeutic

entities. This current review summarizes interesting findings regarding the intriguing roles and benefits of microRNA expression modulation during AAA initiation and propagation. These studies utilize disease-relevant murine models, as well as human tissue from patients undergoing surgical aortic aneurysm repair. Furthermore, we critically examine future therapeutic strategies with regard to their clinical and translational feasibility.

Keywords: microRNA; aortic aneurysm; fibrosis; vascular smooth muscle cells; inflammation; biomarker

1. Abdominal Aortic Aneurysm Disease

Abdominal aortic aneurysms (AAAs) are defined as permanent dilations of the abdominal aorta. The diagnosis of AAA is commonly an accidental finding, although an increasing number of screening programs target particularly high-risk populations [1]. Screening demonstrates that disease prevalence ranges from 1.3% (45–54 years of age) to 12.5% in men (75–84 years of age), and in women from 0% in the youngest to 5.2% in the oldest age groups [2]. Some recently performed analyses, however, suggest lower prevalence in certain subpopulations [3]. The most feared clinical consequence of AAA progression is acute rupture, which carries a mortality of ~80% [4]. The number of deaths attributed to AAA rupture is around 15,000 annually in the United States [5]. However, this incidence is likely underestimated, since AAA rupture is often not recognized as the cause of death. As many as 60% of patients with AAAs die of other cardiovascular causes, such as myocardial infarction or stroke, thereby suggesting a relationship between AAAs and atherosclerosis [6].

Known predictors of AAA growth include diameter of the aorta at diagnosis and active smoking [7]. Some studies have demonstrated that the incidence and progression of AAA are also related to hypertension and age [8]. However, smoking is considered to be the major modifiable risk factor for development of AAA. Indeed, AAA is more closely associated with cigarette smoking than any other tobacco-related disease, excepting lung cancer. The vast majority of AAA patients (>90%) have a history of smoking [9]. As mentioned above, the prevalence of AAAs is greater in men than in women. However, there is emerging evidence that women present with an increased risk of AAA rupture at smaller aortic diameters than men [10,11].

To date, the only available treatment option for AAA has been surgical repair [1]. The classic approach includes the insertion of an intraluminal graft via open access to the aneurysmal aorta. This has now largely been replaced by endovascular stenting approaches. Besides their lack of indication in early stages of the disease, the current interventional methods carry significant operative risk, and thus appear effective only in preventing aortic rupture [4]. Until now, no pharmacological approach has been identified which effectively limits AAA progression or the risk of rupture in humans. What has been lacking is a detailed understanding of the mechanisms of AAA initiation and expansion.

2. Pathology and Cellular Mechanisms

Others have previously discussed the multiple potential cellular and molecular mechanisms associated with AAA development [12,13]. In this article, we will primarily focus on recognized crucial molecular and cellular patho-mechanisms in aneurysm development that are subject to microRNA (miR) regulatory control. Modulation of these miRs could evolve into new therapeutic strategies on the molecular level to combat the burden of aortic aneurysms.

The complex pathology of AAAs is characterized by progressive aortic dilation, promoted by an imbalance between vascular smooth muscle cell (VSMC) proliferation and apoptosis, as well as impairment of extracellular matrix (ECM) synthesis and degradation. These effects are due (at least in part) to transmural aortic inflammation and its disruptive effects on vessel wall homeostasis [8,14,15].

2.1. Impaired Homeostasis of Vascular Smooth Muscle Cells and Extracellular Matrix

Inherited syndromes associated with aneurysm formation suggest the importance of disruption of VSMC and ECM homeostasis in aortic dilation [16], although these familial conditions are more typically associated with ascending thoracic aortic aneurysms (TAAs). The aortic pathology of TAA is characterized by elastic fiber fragmentation and loss, proteoglycan accumulation, as well as focal or diffuse regional VSMC degradation and loss [17]. The role of TGF-β signaling dysregulation in this process is complex. Marfan Syndrome (MFS) and Loeys-Dietz Syndrome (LDS), caused respectively by mutations in fibrillin-1 and TGF-β receptors I and II, predispose to ascending thoracic aortic aneurysms (TAAs), but are much less often associated with AAA [18]. The same is true of familial SMAD3 mutations [19,20].

While considerable evidence points to excessive TGF-β signaling in the various familial TAA-associated conditions, animal models have connected AAA to decreased TGF-β activity [21]. While TGF-β receptor 2 is down-regulated in human AAA tissues [22], no association has been found between genetic polymorphisms in TGF-β receptors and serum TGF-β1 concentration in humans with AAA [23]. Systemic blockade of TGF-β activity augments AngII-induced AAAs in C57BL/6 mice as well as hypercholesterolemic mice, and appears associated with VSMC apoptosis, elastin degradation, and increased inflammatory activity in the aortic wall. In a rat model with chimeric aneurysms located in the infrarenal aorta, TGF-β1 overexpression via endovascular delivery of an adenoviral construct stabilizes pre-existing aortic aneurysms [21].

As a side note, TAAs also occur in families (without syndromic features) due to mutations in SMC contractile protein genes, including SMC-specific isoforms of α-actin (ACTA2) and myosin heavy chain (MYH11), along with the kinase that controls SMC contraction (MYLK) [17,24,25].

Multiple matrix metalloproteases (MMPs), which degrade ECM and are important regulators of aortic vessel wall integrity and morphology, have been extensively studied in human AAA, MMP-9 in particular. A recent meta-analysis included eight case-control studies comparing blood MMP-9 concentration between patients with AAAs and control subjects. Despite wide heterogeneity in

circulating levels (30–750 ng/L), significantly higher MMP-9 concentrations were found in AAA patients [26].

2.2. Inflammation

Various inflammatory cell types are enriched in AAA tissues, especially macrophages. In a rabbit model of AAA induced by periaortic application of calcium chloride, there is striking macrophage accumulation within the adventitia [27]. This feature is also observed in porcine pancreatic elastase (PPE)-infusion induced AAAs in rats [28]. In $ApoE^{-/-}$ mice infused with angiotensin II (AngII), macrophage infiltration within the medial layers of the aorta is accompanied by medial rupture as an early characteristic [29], while profound accumulation of macrophages in the adventitia is seen throughout AAA progression [30].

Furthermore, macrophages appear to actively contribute to AAA development. CCR2 and monocyte chemoattractant protein-1 (MCP-1) interactions are important for macrophage-mediated inflammatory responses, including monocyte chemotaxis. Deficiency of CCR2 in mice limits the formation of AngII- and calcium chloride-induced AAAs [31,32]. Myeloid differentiation factor 88 (MyD88), an adaptor protein central to toll-like receptor signaling, also seems to play a pivotal role in macrophage-mediated vascular inflammation as deficiency of this molecule in macrophages diminishes murine AngII-induced AAAs [33].

T- and B-lymphocytes are frequently observed in AAAs [29,34]. A functional deficiency of CD4+ CD25+ T-regulatory cells was reported in patients with AAAs, and disruption of the balance of T-helper type 1 and type 2 cell function induces AAA in mice with allografted aortas [35].

Neutrophils are also present in human, and animal model AAAs [36]. The adhesion molecule L-selectin was found to be an important mediator for neutrophil recruitment in PPE-induced AAA formation in mice [37]. Neutrophil depletion in mice with aortic perfusion of PPE leads to attenuation of AAAs [36].

Further, many cytokines and chemokines play roles in AAA development [38]. Tumor necrosis factor (TNF)-α, a landmark cytokine in many inflammatory responses, is increased in plasma from patients with AAA and in human AAA tissues [39–41]. TNF-α-converting enzyme (TACE/ADAM17), and osteoprotegerin (a secreted glycoprotein member of the TNF receptor superfamily) are enhanced in human AAAs [39,40,42]. Genetic deficiency or pharmacological inhibition of TNF-α by administration of infliximab attenuates calcium chloride-induced AAAs in mice [43].

3. MicroRNA Biogenesis and Function

MicroRNA (miRNAs) are a class of well-conserved, short, non-coding RNAs that have emerged as key post-transcriptional regulators of gene expression in animals and plants. miRNAs have been described to play major roles in most, if not all, biological processes by influencing stability and translation of messenger RNAs [44]. miRNA genes are transcribed by RNA polymerase II as capped

and polyadenylated primary miRNA transcripts (pri-miRNA) [45]. Pri-miRNA processing occurs in two steps, catalyzed by the enzymes Drosha and Dicer in cooperation with a dsRNA binding protein, "DiGeorge syndrome critical region gene 8" (DGCR8) [46]. In the first step, the Drosha-DGCR8 complex processes pri-miRNA into a ~70-nucleotide precursor hairpin (pre-miRNA), which is then exported to the cytoplasm. Some pre-miRNAs are produced from very short introns (mirtrons) as a result of splicing and debranching, bypassing the Drosha-DGCR8 step [47]. Nuclear export of pre-miRNAs is mediated by the transport receptor exportin 5 (XPO5) [48].

In the cytoplasm Dicer matures pre-miRNA into an imperfect RNA duplex. The strand with the weakest base pairing at the 5' terminus is loaded into the miRNA-induced silencing complex (miRISC), and is therefore considered to be biologically active [49]. While both strands of the duplex are produced in equal amounts by transcription, their accumulation into the miRISC is asymmetric [50]. Initially, the non-miRISC strand was assumed to be an inactive passenger designated the *(star)-strand. However, systemic computational analysis has demonstrated that star-strands may contain well-conserved target recognition sites, indicating functional relevance [51]. Indeed, several recent publications have reported star-strands to be biologically active, widening the potential regulatory potency of miRNA-duplexes [52–54].

After the selected strand is loaded into the miRISC, the miRNA guides the miRISC to bind to the 3'UTR of its target sequence. The seed sequence (the first two to eight nucleotides) is considered the most important for target recognition and silencing of the mRNA [55,56]. Translation of the mRNA is inhibited after association of the miRISC with its target sequence. Efficient mRNA targeting requires continuous base pairing of the seed region to the target mRNA. Furthermore, Ago(argonaute)-proteins and the glycine-tryptophan protein of 182 kDa (GW182), core components of the miRISC, are directly associated with miRNAs, and are needed for effective translational repression, mRNA destabilization, and degradation. The exact mechanisms of translational arrest by the miRNA:mRNA complex are still a matter of debate, although both initiation and elongation steps of translation are thought to be affected [57,58].

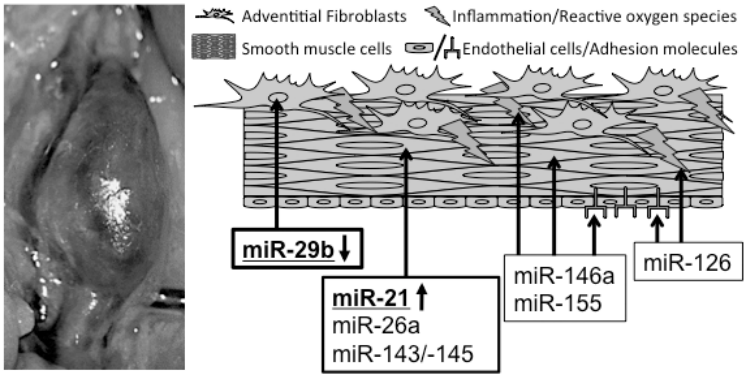
4. miRs in AAA Disease

In recent years, several miRs have been found to regulate vascular pathologies, in general, and aortic aneurysm (thoracic and abdominal) disease, in particular (Table 1). We performed a systematic published literature search on articles investigating miRNA expression and function in aortic aneurysm disease. This current review focuses mainly on miRNAs that have not only been detected as being potentially dys-regulated in human aneurysmal tissue, but have also been thoroughly studied in functional experiments, thus accessing a therapeutic strategy of beneficially altering miRNA expression to limit AAA progression (Figure 1).

Table 1. Regulatory role of microRNAs in murine abdominal aortic aneurysm (AAA) disease models (AFB = adventitial fibroblasts; AngII = angiotensin II; ASMC = aortic smooth muscle cells; PPE = porcine pancreatic elastase).

microRNA	Model of AAA induction	Effect on AAA progression
miR-21	PPE-infusion in C57BL/6 mice and AngII-infusion in *ApoE*$^{-/-}$ mice [59]	Regulates proliferation and apoptosis in ASMCs via PTEN/PI3K/AKT; induction of miR-21 through NFκB
miR-26a	PPE-infusion in C57BL/6 mice and AngII-infusion in *ApoE*$^{-/-}$ mice [60]	Inhibition of ASMC-differentiation via SMAD-1 and SMAD-4 depression
miR-29b	AngII in 1.5-year-old C57BL/6 [61]; PPE-infusion in C57BL/6 mice and AngII in *ApoE*$^{-/-}$ mice [62]	Modulating the fibrotic response in aortic wall through several collagen isoforms; repression of miR-29b in AFBs through TGF-β
miR-143/145	miR-143/145 knockout and *ApoE*$^{-/-}$ mice [63]	Regulation of ASMC homeostasis and differentiation

Figure 1. Association between microRNAs and murine abdominal aortic aneurysm formation. microRNAs (miRs) in bold and underlined have been established as regulators of aneurysm disease, utilizing gain- and loss-of function studies. All other miRs are suspected and potential disease-related modulators.



4.1. miR-21

miR-21 is considered an onco-miRNA, with increased expression in many solid tumors, where it promotes cell proliferation, migration and anti-apoptosis [64]. Data indicate that miR-21 is also highly expressed in VSMCs, and implicate it in the regulation of SMC phenotype in vascular disorders, such as post-injury neointimal lesions [65,66].

Interestingly, miR-21 stimulation induces up-regulation of smooth-muscle restricted contractile proteins through silencing of "programmed cell death protein" (PDCD)-4 expression, a known tumor suppressor protein. These findings suggest that miR-21 could regulate both VSMC contractile function [67] and proliferation [68]. miR-21 also targets multiple members of the dedicator of cytokinesis (DOCK) superfamily and modulates the activity of ras-related C3 botulinum toxin substrate 1 (Rac1) small GTPase to regulate VSMC phenotype [69].

miR-21 regulates growth and survival of VSMCs by decreasing the expression of "phosphatase and tensin homolog" (PTEN) and inducing expression of Bcl-2, resulting in pro-proliferative and anti-apoptotic effects in a carotid injury model in rats [68]. Regarding homeostasis, miR-21 promotes VSMC differentiation in response to TGF-β1 and BMP-4 [67]. These factors were shown to stimulate the processing of miR-21 in human pulmonary artery smooth muscle cells from the pri-miR to the mature miR via SMAD proteins. Additionally, miR-21 has been shown to regulate hypoxia-induced pulmonary VSMC proliferation and migration by regulating PDCD4, Sprouty 2 (SPRY2), and peroxisome proliferator-activated receptor-α (PPARα), known for their anti-proliferative and anti-migratory effects on VSMCs [70].

Interestingly, a recent report indicates that miR-21 is induced in tissue of arteriosclerosis obliterans of the lower extremities, even with <10% stenosis, and also is induced in VSMCs in response to platelet-derived growth factor (PDGF)-BB and/or hypoxia. In this report, tropomyosin 1 (TPM1) was identified as a target gene for miR-21. TPM1 reduction leads to a reduction in cytoskeletal stability, promoting VSMC proliferation and migration [71].

Furthermore, cyclic stretch has been shown to modulate miR-21 expression at the transcriptional level via FBJ murine osteosarcoma viral oncogene homolog (c-fos/AP-1) in cultured human aortic SMCs [72]. While moderate stretch is essential for maintaining vessel wall structure and vascular homeostasis [73], exacerbated stretch, as in hypertension, could promote pathological vascular remodeling by stimulating SMC proliferation, apoptosis, and migration and abnormal extracellular matrix deposition [74,75].

In endothelial cells (ECs), prolonged shear stress up-regulates the expression of miR-21 through modulation of the phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K)/v-akt murine thymoma viral oncogene (Akt) pathway, which leads to an increase of nitric oxide (NO) production while reducing apoptosis [76]. miR-21 is also expressed in endothelial progenitor cells (EPCs), where it suppresses high mobility group AT-hook 2 (Hmga2) expression, a chromatin-associated protein that modulates transcription through altering chromatin structure. Thus, inducing overexpression of miR-21 decreases proliferation and limits EPC angiogenesis *in vitro* and *in vivo* [77].

In regards to aortic dilatation, we discovered that miR-21 was significantly up-regulated in two established murine models of AAA disease, the PPE-infusion model in C57B/L6 mice and the AngII-infusion in $ApoE^{-/-}$ mice [59]. Out of the aforementioned VSMC-specific miR-21 target genes that alter proliferation and apoptosis, PTEN was the only target gene to be significantly down-regulated at three different time points during aneurysm development and progression. PTEN, a lipid and protein phosphatase and important tumor suppressor gene, acts as a key negative regulator of the PI3K pathway. Systemic injection of a locked-nucleic-acid (LNA) modified antagomiR against miR-21 diminished the pro-proliferative impact of down-regulated PTEN, leading to a significant increase in expansion of AAAs. Further down-regulation of aortic PTEN with a pre-miR-21-loaded lentivirus had significant protective effects on aneurysm expansion by inducing massive proliferation in the aortic wall in both murine models [59].

As mentioned above, smoking is considered to be the major modifiable risk factor for AAA disease. In our study, nicotine (a major constituent of tobacco smoke) accelerated AAA growth in both murine aneurysm models, and caused an augmented increase in miR-21 levels, which appeared to be a protective response to limit further aneurysm expansion and rupture. *In vitro* studies utilizing human aortic SMCs and ECs, as well as adventitial fibroblasts showed aortic SMCs to be the most responsive to miR-21 modulation. Our group also showed that miR-21 induction in nicotine, as well as AngII and interleukin-6 (IL-6) pre-treated SMCs, is dependent on NF-κB signaling. In support of these findings, we found increased expression of miR-21 and down-regulated PTEN in samples obtained from human AAA patients undergoing surgical repair of their enlarged infrarenal aorta compared to control abdominal aorta from organ donors. Notably, miR-21 was even further up-regulated (with PTEN being further decreased) in smokers with AAA disease compared with non-smokers [59].

4.2. miR-26a

Employing *in vitro* experiments with human aortic SMCs, Leeper and colleagues [60] found that miR-26a promotes the synthetic phenotype through regulation of SMAD1 and SMAD4, contributing to the regulatory circuit of TGF-β signaling-associated pathways. Overexpression of SMAD-1 and SMAD-4 was inducible with anti-miR-26a treatment. In two mouse models of aneurysm formation (PPE- and AngII-infusion), miR-26 levels were decreased, which might contribute to AAA formation through enhanced SMC apoptosis. Thus, miR-26 regulation in aneurysmal tissue with AAA development may in fact be causal, and not compensatory.

4.3. miR-29b

The miR-29 family of miRs contains three members (miR-29a, miR-29b, and miR-29c) that are encoded by two separate loci, giving rise to bi-cistronic precursor miRs (miR-29a/b1 and miR-29b2/c). This family targets numerous gene transcripts that encode ECM proteins involved in fibrotic responses, including several collagen isoforms (e.g., COL1A1, COL1A2, COL3A1), fibrillin-1, and elastin (ELN) [78], and is known to modulate gene expression during development and aging of the aorta [61] and during the progression of aortic aneurysms [61,62].

Other fibrosis-related responses and diseases, such as liver [79] and kidney fibrosis [80], systemic sclerosis [81], as well as cardiac fibrosis in response to myocardial ischemia [78], have all been linked to repressed levels of miR-29. TGF-β-associated pathways are important regulators of miR-29 expression, leading to triggering of the fibrotic response by decreasing miR-29 levels in cardiac fibroblasts, hepatic stellate cells, and dermal fibroblasts, and leading to a substantial increase in the aforementioned ECM target genes [78,81,82].

Based on these observations, miR-29 seems to be a crucial regulator of aortic aneurysm disease through modulating genes and pathways which are responsible for ECM composition and dynamics. We found that miR-29b was the only member of the miR-29 family to be significantly

down-regulated at three different time points during murine AAA development and progression [82]. Further decreasing of miR-29b expression with a LNA-anti-miR-29b led to an acceleration of collagen encoding gene expression (COL1A1, COL2A1, COL3A1, COL5A1), as well as elastin (ELN). Furthermore, matrix-metalloproteinases-2 and -9 (MMP2 and MMP9) were down-regulated in LNA-anti-miR-29b-transduced mice. These results were reproducible in two independent mouse AAA models, (PPE-and AngII-infusion), and led to a significant decrease in aneurysm expansion compared to a scrambled-control-miR injected group.

Human AAA tissue samples displayed a similar pattern of reduced miR-29b expression with increased collagen gene expression in comparison to non-aneurysmal organ donor controls. These results suggest that the aortic wall, which weakens due to steadily increasing diameter, acts to induce expression of collagens by repressing miR-29b levels, providing additional support to the aortic wall in an attempt to limit the risk for rupture.

Aging is a well-established risk factor for aneurysm development. Boon *et al.* were the first to publish a study connecting miR regulation to aortic dilatation and aging. They discovered that expression of the miR-29 family was increased in the aging mouse aorta [61]. Rather than utilizing the more commonly employed ApoE$^{-/-}$ or LDL receptor$^{-/-}$ mice, Boon and colleagues studied AngII infusion in 18-month-old C57BL/6 (wild type) mice. In these mice, AngII infusion increased miR-29b expression in samples derived from the entire aorta, which would seem to suggest that with aging the protective role of miR-29b during AAA development may be diminished. In accordance with our aforementioned results, Boon *et al.* found that systemic treatment with an LNA-modified anti-miR-29b significantly increased the expression of collagen isoforms (COL1A1, COL3A1), as well as ELN, and decreased suprarenal aortic dilatation in aged AngII-treated mice.

5. miR-143/145

Probably the most extensively studied miR in VSMC pathology is the miR-143/145 cluster, which is transcribed as a bi-cistronic transcript from a common promoter, which in turn is regulated by serum response factor (SRF), myocardin, and myocardin-related transcription factor-A [83]. MiR-143/145 is dramatically reduced in several vascular disease models, e.g., carotid balloon/wire injury, carotid ligation in rats, and in ApoE$^{-/-}$ mice [83–85].

miR-143/145 alters SMC phenotypic switching in response to vascular injury, influencing both the synthetic/proliferative and the contractile/differentiated states [63,83–86]. Studies from several different groups have shown that these effects are partly mediated by targeting of multiple transcription factors, including KLF4, KLF5, and ELK-1 [84–86]. Further, down-regulation of miR-143/145 is sufficient to up-regulate PDGF receptor (PDGF-R), protein kinase C (PKC) epsilon, and fascin, an actin bundling protein of podosomes. These last are thought to be necessary for vascular wall matrix remodeling, potentially affecting the progression of aortic dilatation [87]. Interestingly, one of the first reports regarding the role of miRs in aneurysm disease showed that miR-143/145 expression is

reduced in aortas from patients with thoracic aortic aneurysm, permitting dedifferentiation of aortic VSMC with a resultant decrease in contractile function [63].

Finally, miR-143/145 may be secreted in microvesicles derived from ECs (which otherwise do not usually express these miRs) [84]. It has been proposed that shear stress-induced KLF-2 may stimulate expression of miR-143/145 in ECs [88], leading to miR secretion in microvesicles and transfer into VSMCs [84]. EC-derived microvesicles containing miR-143/145 can reduce atherosclerotic lesions when injected into $ApoE^{-/-}$ mice [88].

6. Other miRs

A growing body of literature highlights the role of miRs in the regulation of angiogenesis and inflammation [12–14,89]. Smooth muscle degradation, along with decreased VSMC proliferation, decreased ECM synthesis and impaired ECM remodeling, have all previously been linked to AAA development. Clearly, these contributing mechanisms of aortic dilation may be regulated through miRs. However, the miRs described below have not yet been directly tied to aortic aneurysm initiation, propagation, or rupture.

6.1. miR-126

One of the most intriguing miRs as regards vascular inflammation is miR-126, an EC-enriched miR, which negatively regulates VCAM-1 expression [90,91]. Apoptotic bodies are released from ECs during atherosclerotic progression, and have been shown to contain miR-126. miR-126 decreases the expression of G-protein signaling 16 (RGS16) in ECs, thereby up-regulating the chemokine (C-X-C motif) ligand 12 (CXCL12) receptor. CXCL12 activation then decreases EC apoptosis and recruits progenitor cells at the lesion site, reducing the atherosclerotic burden *in vivo*, and contributing to plaque stabilization [92].

6.2. miR-146a

Alterations associated with aging in blood vessels include a decrease in compliance and an increase in vascular inflammatory response, which could promote AAA propagation. Several reports show dysregulation of miRs in the vasculature during aging. In particular, miR-146a expression is decreased in senescent ECs. It targets NADPH oxidase 4 (NOX4), decreasing reactive oxygen species (ROS) production. These data suggest that the reduction in miR-146 expression potentially enhances aging effects through NOX4-derived ROS [93]. In another study, miR-146a and KLF4 were found to form a feedback loop, regulating each other's expression and VSMC proliferation. The authors propose that miR-146a regulates KLF4, which competes with KLF5 binding to the miR-146a promoter to inhibit transcription [94].

6.3. miR-155

miR-155 is another miR of potential interest in AAA disease progression due to its effects on the renin-angiotensin-system (RAS). miR-155 is induced by TNF (which independently has been shown to contribute to AAA development) [95], and then negatively regulates the expression of the transcription factor "v-ets erythroblastosis virus E26 oncogene homolog 1" (Ets-1) [96]. AngII-induced overexpression of miR-155 results in a decrease in Ets-1, affecting expression of downstream targets such as VCAM-1, fms-related tyrosine kinase 1 (FLT1) and MCP1, and impairing lymphocyte adhesion to ECs [96]. miR-155 also has been shown to target the angiotensin II type 1 receptor (AT1R), resulting in decreased AngII-induced migration of ECs [96].

While the above-described effects of miR-155 might suggest an anti-inflammatory role, Nazari-Jahantigh *et al.* validated miR-155 in macrophages as a crucial component of atherosclerosis development. In these cells, miR-155 promoted the expression of MCP-1/CCL2, and directly suppressed Bcl-6, a transcription factor that inhibits NF-κB [97]. It has also been described that hematopoietic deficiency of miR-155 increases atherosclerotic plaque size and instability [98], possibly by inhibition of lipid uptake and inflammatory responses in monocytes. Clearly, findings thus far regarding the role of miR-155 have been somewhat ambiguous.

In addition to these miRs, Pahl *et al.* examined miR-regulation in human abdominal aortic tissue of patients undergoing elective open repair with samples collected at autopsy or obtained from a pre-existing tissue biobank [99], utilizing microRNA-array. Out of a total of 847 miRs, 3 miRs presented as significantly up- (miR-181a*, miR-146a, miR-21) and 5 miRs as down-regulated (miR-133b, miR133a, miR331-3p, miR30c-2*, miR-204) in patients with AAAs compared to controls. However, using an additional tissue set, qRT-PCR was only able to confirm the down-regulated miRs from the array.

7. Therapeutic Approaches Using miR Modulators

The identification of both the underlying causes of vascular disease, as well as appropriate interventions, remain great challenges to both basic vascular biology and everyday clinical practice. The traditional methods of drug design, involving enzymes, cell surface receptors, and other proteins, appear sometimes less effective in the treatment of cardiovascular diseases, due to the highly sensitive nature of the targeted systems.

In this dismaying scenario, the discovery of an entirely new method of gene regulation by miRs, and their recent validation as markers and modulators of vascular functionality during pathological conditions, provide new hope for innovative therapies. Research in recent years has recognized the crucial regulatory roles that miRs play in vascular diseases such as myocardial infarction, stroke, and aortic aneurysm [100].

Intriguingly, miRs also appear to represent valid therapeutic targets, because modulation of their expression *in vivo* with either antisense RNA molecules or miR-mimics/pre-miRs has been

shown to effectively modulate cardiovascular disease in various animal models [101]. Inhibition or overexpression of a single miR can induce or attenuate pathological responses in the cardiovascular system, as a result of the regulated coordination of numerous target genes involved in complex physiological and disease phenotypes. The most important difference between modulating miRs, and the traditional therapeutic approach is that standard drugs typically interact with specific cellular targets, whereas miRs have the capability of modulating entire functional networks [102].

miR modulation is performed by supplying antagomiRs (or anti-miRs; synthetic reverse compliments of oligonucleotides) that bind to a target miR and silence it, or by using pre-miRs/miR-mimics that act similarly to the original miR [101]. Recent animal and even human efficacy data indicate that antagomiRs have the potential to become a whole new class of drugs. These inhibitors of miR expression have several significant advantages, which make them very attractive from a drug development standpoint, including small size, as well as frequent conservation of their target miRs across species. Using lessons learned from antisense technologies (e.g., siRNA), potent oligonucleotide chemistries to inhibit miRs are currently being investigated [103]. These efforts have given rise to candidates that bind to their putative miR targets with remarkable affinity and specificity, and which have desirable drug-like qualities, including increased stability and favorable pharmacokinetics.

The most common type of modification being utilized to protect antagomiRs from immediate degradation *in vivo* is the addition of a locked nucleic acid (LNA). LNA contains a class of bicyclic RNA analogs in which the furanose ring in the sugar-phosphate backbone is chemically locked in a RNA-mimicking N-type (C3'-endo) conformation by the introduction of a 2'-*O*,4'-*C*-methylene bridge. This modification leads to nuclease resistance, as well as an increase in binding affinity to the targeted miR, which is accomplished by Watson-Crick complementary base pairing [104]. Regarding the use of antagomiRs in humans, there have been no immunogenic or toxicological safety issues reported to date. However, the major drawback of these substances at this point seems to be the necessity of repeated delivery of doses for long-term therapeutic effects. This becomes a critical issue when the route of delivery is an invasive procedure, such as systemic injection [105]. The antagomiR that has advanced the farthest in clinical trials to date is Miravirsen (anti-122) for patients with chronic hepatitis-C (HCV) infections. Recently published data from a Phase 2a trial demonstrated that the drug was not only safe, but also well tolerated, providing prolonged antiviral activity well after the last dose of monotherapy [106].

Unlike antagomiRs, the prospect of delivery of injectable, naked miR-mimics and/or pre-miRs has remained problematic. For now, lenti- as well as adeno-associated viruses (AAV) represent efficacious delivery platforms for miRs, but these carry the risks common to most gene therapies. Lentiviral vectors, for example, are derived from HIV type 1 (HIV-1), and thus the production of wild-type HIV through homologous recombination of the virus remains a major safety concern. However, recent lentiviral vector developments permitting deletion of the U3 promoter region of the long terminal repeats from the virus, leading to self-inactivation, may resolve this issue, making them a promising vector for future applications [107].

miR-mimic and pre-miR development also present difficulties related to the need to deliver synthetic RNA duplexes in which one strand (the "guide" strand) is identical to the miR of interest, while the complementary strand ("passenger" strand) is modified to increase stability as well as cellular uptake. Apart from the problems involved in permitting cellular uptake of double-stranded miR-mimics, the passenger strand has the potential to counter-productively act as an antagomiR [105].

Given the above limitations, the development of miR mimics, which do not require a viral vector represents an important therapeutic goal. Some preclinical studies have achieved this in murine models by packaging synthetic miR duplexes within lipid nanoparticles [108,109].

In summary, the ability to modulate miR activity through systemic delivery of miR inhibitors or mimics without toxicity provides unprecedented opportunities for intervening in disease processes. While challenges such as potential off-target effects and the urgent need for local and/or cell-type specific delivery mechanisms remain, the pace of discovery in this field portends new, feasible clinical therapeutic approaches in patients.

8. miRs as Biomarkers in AAA Disease

At the outset, it is necessary to point out that, to date, no easily accessible and reproducibly measurable biomarker has been identified with prognostic value for AAA growth, or even for the potential to rupture [1,110].

Recently, miRs have received much attention regarding their suitability as biomarkers for vascular disease. Following pioneering work from the cancer field, several cardiovascular studies have found substantial variations in miR expression in numerous clinical specimen subtypes (e.g., blood, urine, saliva, *etc*.) [111–114]. Measuring levels of circulating miRs has several advantages and offers novel opportunities. For example, as with nucleic acids, miRs can be both amplified and detected with high sensitivity and specificity. Also, miR-microarrays and quantitative PCR (qPCR) methodology allows the quantification of many miRs in a single experiment. There is evidence that the combined analysis of many miRs and their co-expression patterns (miR networks) enhances their predictive power as biomarkers. Furthermore, miRs are relatively stable over time in human blood and appear to be protected from degradation through various mechanisms [115].

Despite this, the quantitative analysis of miRs in material such as blood and urine comes with certain disadvantages. Firstly, the concentrations of most circulating miRs are typically very low (with the exception of whole blood samples), making reliable quantitation and normalization a challenge with existing technology. Also, there exists no consensus for miR normalization controls. Beyond this, current qPCR and microarray technologies are still quite time-consuming (several hours) compared with some protein-based biomarker tests such as troponin or C-reactive protein, which can offer results within minutes [116]. For now, the added value of miR-based biomarkers remains to be established by more rigorous testing and optimization.

Despite these hurdles, several laboratories have already obtained profiles of circulating miRs in cardiovascular disease and explored their biomarker potential. Immediately apparent are certain

inconsistencies between studies, where the same or highly similar settings have been studied. This is partially attributed to the current immaturity of the field, which still includes technical issues such as variability of RNA extraction protocols, different means of nucleic acid detection, and the aforementioned normalization procedures. However, many studies are also simply clinically underpowered, and/or do not use appropriate controls matched for potentially confounding factors such as age, sex, medication, comorbidities, and tissue source. Also, there has been minimal comparison of miRs to traditional reference biomarkers.

The first study to look at expression levels of circulating miRs in AAA disease was performed by Kin *et al*. The authors investigated a subset of miRs, which they identified to be significantly altered in abdominal aortic tissue samples from patients with AAA undergoing surgical repair when compared with non-aneurysmal thoracic aortic specimen from patients undergoing aortic valve replacement [117]. Interestingly, miRs that were up-regulated in AAA tissue samples appeared significantly down-regulated in plasma from patients with AAA compared to a small group of healthy volunteers. These included miRs-15a/b, -29b, -124a, -126, -146a, -155, and -223. Clearly, further studies in larger cohorts are necessary to explore the diagnostic, and, even more important, the predictive capabilities of miRs as biomarkers in AAA disease.

9. Summary and Perspectives

The demonstration that miRs play crucial roles in cardiovascular disease and can be easily regulated *in vitro* and *in vivo* by antagomiRs and pre-miRs/miR-mimics has tremendously accelerated miR research and nourished hopes that the agents used and verified in animal models could some day be employed in humans with AAA disease. miRs represent a relatively young, but rapidly advancing, field of basic biological and translational research with potentially new and innovative therapeutic applications. For vascular diseases in particular, the availability of local (coated stents and/or balloons) or cell type-specific delivery mechanisms would significantly increase the value of miR therapeutics in everyday clinical practice.

Acknowledgments

We would like to thank all past and current lab members of our laboratories at Stanford and Stockholm for their determination to generate the data for parts of the research presented in this present review. Our own research projects are supported by grants from the National Institutes of Health (1P50HL083800-01 to PST; 5K08 HL080567 to JMS), the Stanford Cardiovascular Institute (to JMS), the American Heart Association (0840172N to PST, 09POST2260118 to LM), the Karolinska Institute Cardiovascular Program Career Development Grant, and the Swedish Heart-Lung-Foundation (20120615 both to LM).

Conflict of Interest

The authors declare no conflict of interest.

References

1. Golledge, J.; Muller, J.; Daugherty, A.; Norman, P. Abdominal aortic aneurysm: Pathogenesis and implications for management. *Arterioscler. Thromb. Vasc. Biol*. **2006**, *26*, 2605–2613.
2. Go, A.S.; Mozaffarian, D.; Roger, V.L.; Benjamin, E.J.; Berry, J.D.; Borden, W.B.; Bravata, D.M.; Dai, S.; Ford, E.S.; Fox, C.S.; Franco, S.; *et al*. Executive summary: Heart disease and stroke statistics—2013 update: A report from the American Heart Association. *Circulation* **2013**, *127*, 143–152.
3. Svensjo, S.; Martin Björck, M.; Gürtelschmid, M.; Gidlund, K.D.; Hellberg, A.; Wanhainen, A. Low prevalence of abdominal aortic aneurysm among 65-year-old Swedish men indicates a change in the epidemiology of the disease. *Circulation* **2011**, *124*, 1118–1123.
4. Golledge, J.; Norman, P.E. Current status of medical management for abdominal aortic aneurysm. *Atherosclerosis* **2011**, *217*, 57–63.
5. Thom, T.; Haase, N.; Rosamond, W.; Howard, V.J.; Rumsfeld, J.; Manolio, T.; Zheng, Z.J.; Flegal, K.; O'Donnell, C.; Kittner, S.; *et al*. Heart disease and stroke statistics—2006 update: A report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* **2006**, *113*, e85–e151.
6. Golledge, J.; Tsao, P.S.; Dalman, R.L.; Norman, P.E. Circulating markers of abdominal aortic aneurysm presence and progression. *Circulation* **2008**, *118*, 2382–2392.
7. Jones, D.W.; Easton, J.D.; Halperin, J.L.; Hirsch, A.T.; Matsumoto, A.H.; O'Gara, P.T.; Safian, R.D.; Schwartz, G.L.; Spittell, J.A.; American Heart Association. Atherosclerotic Vascular Disease Conference: Writing Group V: Medical decision making and therapy. *Circulation* **2004**, *109*, 2634–2642.
8. Weintraub, N.L. Understanding abdominal aortic aneurysm. *N. Engl. J. Med*. **2009**, *361*, 1114–1116.
9. Franks, P.J.; Edwards, R.J.; Greenhalgh, R.M.; Powell, J.T. Smoking as a risk factor for abdominal aortic aneurysm. *Ann. N. Y. Acad. Sci.* **1996**, *800*, 246–248.
10. Norman, P.E.; Powell, J.T. Abdominal aortic aneurysm: The prognosis in women is worse than in men. *Circulation* **2007**, *115*, 2865–2869.
11. Powell, J.T.; Greenhalgh, R.M. Clinical practice. Small abdominal aortic aneurysms. *N. Engl. J. Med*. **2003**, *348*, 1895–1901.
12. Lu, H.; Rateri, D.L.; Bruemmer, D.; Cassis, L.A.; Daugherty, A. Novel mechanisms of abdominal aortic aneurysms. *Curr. Atheroscler. Rep*. **2012**, *14*, 402–412.
13. Lu, H.; Rateri, D.L.; Bruemmer, D.; Cassis, L.A.; Daugherty, A. Involvement of the renin-angiotensin system in abdominal and thoracic aortic aneurysms. *Clin. Sci*. **2012**, *123*, 531–543.

14. Daugherty, A.; Cassis, L.A. Mouse models of abdominal aortic aneurysms. *Arterioscler. Thromb. Vasc. Biol*. **2004**, *24*, 429–434.
15. Milewicz, D.M. MicroRNAs, fibrotic remodeling, and aortic aneurysms. *J. Clin. Invest.* **2012**, *122*, 490–493.
16. Lindsay, M.E.; Dietz, H.C. Lessons on the pathogenesis of aneurysm from heritable conditions. *Nature* **2011**, *473*, 308–316.
17. Guo, D.; Pannu, H.; Tran-Fadulu, V.; Papke, C.L.; Yu, R.K.; Avidan, N.; Bourgeois, S.; Estrera, A.L.; Safi, H.J.; Sparks, E. Mutations in smooth muscle alpha-actin (ACTA2) lead to thoracic aortic aneurysms and dissections. *Nat. Genet*. **2007**, *39*, 1488–1493.
18. Loeys, B.L.; Schwarze, U.; Holm, T.; Callewaert, B.L.; Thomas, G.H.; Pannu, H.; de Backer, J.F.; Oswald, G.L.; Symoens, S.; Manouvrier, S. Aneurysm syndromes caused by mutations in the TGF-beta receptor. *N. Engl. J. Med*. **2006**, *355*, 788–798.
19. Regalado, E.S.; Guo, D.; Villamizar, C.; Avidan, N.; Gilchrist, D.; McGillivray, B.; Clarke, L.; Bernier, F.; Santos-Cortez, R.L.; Leal, S.M. Exome sequencing identifies SMAD3 mutations as a cause of familial thoracic aortic aneurysm and dissection with intracranial and other arterial aneurysms. *Circ. Res*. **2011**, *109*, 680–686.
20. Van de Laar, I.M.B.H.; Oldenburg, R.A.; Pals, G.; Roos-Hesselink, J.W.; de Graaf, B.M.; Verhagen, J.M.A.; Hoedemaekers, Y.M.; Willemsen, R.; Severijnen, L.; Venselaar, H. Mutations in SMAD3 cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis. *Nat. Genet*. **2011**, *43*, 121–126.
21. Dai, J.; Losy, F.; Guinault, A.M.; Pages, C.; Anegon, I.; Desgranges, P.; Becquemin, J.P.; Allaire, E. Overexpression of transforming growth factor-beta1 stabilizes already-formed aortic aneurysms: A first approach to induction of functional healing by endovascular gene therapy. *Circulation* **2005**, *112*, 1008–1015.
22. Biros, E.; Walker, P.J.; Nataatmadja, M.; West, M.; Golledge, J. Downregulation of transforming growth factor, beta receptor 2 and Notch signaling pathway in human abdominal aortic aneurysm. *Atherosclerosis* **2012**, *221*, 383–386.
23. Golledge, J.; Clancy, P.; Jones, G.T.; Cooper, M.; Palmer, L.J.; van Rij, A.M.; Norman, P.E. Possible association between genetic polymorphisms in transforming growth factor beta receptors, serum transforming growth factor beta1 concentration and abdominal aortic aneurysm. *Br. J. Surg*. **2009**, *96*, 628–632.
24. Wang, L.; Guo, D.; Cao, J.; Gong, L.; Kamm, K.E.; Regalado, E.; Li, L.; Shete, S.; He, W.; Zhu, M.; *et al*. Mutations in myosin light chain kinase cause familial aortic dissections. *Am. J. Hum. Genet.* **2010**, *87*, 701–707.
25. Zhu, L.; Vranckx, R.; van Kien, P.K.; Lalande, A.; Boisset, N.; Mathieu, F.; Wegman, M.; Glancy, L.; Gasc, J.; Brunotte, F.; *et al.* Mutations in myosin heavy chain 11 cause a syndrome associating thoracic aortic aneurysm/aortic dissection and patent ductus arteriosus. *Nat. Genet*. **2006**, *38*, 343–349.

26. Takagi, H.; Manabe, H.; Kawai, N.; Goto, S.; Umemoto, T. Circulating matrix metalloproteinase-9 concentrations and abdominal aortic aneurysm presence: A meta-analysis. *Interact. Cardiovasc. Thorac. Surg*. **2009**, *9*, 437–440.
27. Freestone, T.; Turner, R.J.; Higman, D.J.; Lever, M.J.; Powell, J.T. Influence of hypercholesterolemia and adventitial inflammation on the development of aortic aneurysm in rabbits. *Arterioscler. Thromb. Vasc. Biol*. **1997**, *17*, 10–17.
28. Anidjar, S.; Salzmann, J.L.; Gentric, D.; Lagneau, P.; Camilleri, J.P.; Michel, J.B. Elastase-induced experimental aneurysms in rats. *Circulation* **1990**, *82*, 973–981.
29. Saraff, K.; Babamusta, F.; Cassis, L.A.; Daugherty, A. Aortic dissection precedes formation of aneurysms and atherosclerosis in angiotensin II-infused, apolipoprotein E-deficient mice. *Arterioscler. Thromb. Vasc. Biol*. **2003**, *23*, 1621–1626.
30. Rateri, D.L.; Howatt, D.A.; Moorleghen, J.J.; Charnigo, R.; Cassis, L.A.; Daugherty, A. Prolonged infusion of angiotensin II in apoE(/ −) mice promotes macrophage recruitment with continued expansion of abdominal aortic aneurysm. *Am. J. Pathol.* **2011**, *179*, 1542–1548.
31. Daugherty, A.; Rateri, D.L.; Charo, I.F.; Phillip Owens, A., III; Howatt, D.A.; Cassis, L.A. Angiotensin II infusion promotes ascending aortic aneurysms: Attenuation by CCR2 deficiency in apoE−/− mice. *Clin. Sci*. **2010**, *118*, 681–689.
32. MacTaggart, J.N.; Xiong, W.; Knispel, R.; Baxter, B.T. Deletion of CCR2 but not CCR5 or CXCR3 inhibits aortic aneurysm formation. *Surgery* **2007**, *142*, 284–288.
33. Phillip Owens, A., III; Rateri, D.L.; Howatt, D.A.; Moore, K.J.; Tobias, P.S.; Curtiss, L.K.; Lu, H.; Cassis, L.A.; Daugherty, A. MyD88 deficiency attenuates angiotensin II-induced abdominal aortic aneurysm formation independent of signaling through Toll-like receptors 2 and 4. *Arterioscler. Thromb. Vasc. Biol.* **2011**, *31*, 2813–2819.
34. Daugherty, A.; Manning, M.W.; Cassis, L.A. Angiotensin II promotes atherosclerotic lesions and aneurysms in apolipoprotein E-deficient mice. *J. Clin. Invest*. **2000**, *105*, 1605–1612.
35. Yin, M.; Zhang, J.; Wang, Y.; Wang, S.; Böckler, D.; Duan, Z.; Xin, S. Deficient CD4+CD25+ T regulatory cell function in patients with abdominal aortic aneurysms. *Arterioscler. Thromb. Vasc. Biol.* **2010**, *30*, 1825–1831.
36. Eliason, J.L.; Hannawa, K.K.; Ailawadi, G.; Sinha, I.; Ford, J.W.; Deogracias, M.P.; Roelofs, K.J.; Woodrum, D.T.; Ennis, T.L.; Henke, P.K.; *et al.* Neutrophil depletion inhibits experimental abdominal aortic aneurysm formation. *Circulation* **2005**, *112*, 232–240.
37. Hannawa, K.K.; Eliason, J.L.; Woodrum, D.T.; Pearce, C.G.; Roelofs, K.J.; Grigoryants, V.; Eagleton, M.J.; Henke, P.K.; Wakefield, T.W.; Myers, D.D. *et al.* L-selectin-mediated neutrophil recruitment in experimental rodent aneurysm formation. *Circulation* **2005**, *112*, 241–247.
38. Middleton, R.K.; Lloyd, G.M.; Bown, M.J.; Cooper, N.J.; London, N.J.; Sayers, R.D. The pro-inflammatory and chemotactic cytokine microenvironment of the abdominal aortic aneurysm wall: A protein array study. *J. Vasc. Surg*. **2007**, *45*, 574–580.

39. Kaneko, H.; Anzai, T.; Horiuchi, K.; Kohno, T.; Nagai, T.; Anzai, A.; Takahashi, T.; Sasaki, A.; Shimoda, M.; Maekawa, Y.; *et al.* Tumor necrosis factor-alpha converting enzyme is a key mediator of abdominal aortic aneurysm development. *Atherosclerosis* **2011**, *218*, 470–478.
40. Satoh, H.; Nakamura, M.; Satoh, M.; Nakajima, T.; Izumoto, H.; Maesawa, C.; Kawazoe, K.; Masuda, T.; Hiramori, K. Expression and localization of tumour necrosis factor-alpha and its converting enzyme in human abdominal aortic aneurysm. *Clin. Sci*. **2004**, *106*, 301–306.
41. Juvonen, J.; Surcel, H.; Satta, J.; Teppo, A.; Bloigu, A.; Syrjälä, H.; Airaksinen, J.; Leinonen, M.; Saikku, P.; Juvonen, T. Elevated circulating levels of inflammatory cytokines in patients with abdominal aortic aneurysm. *Arterioscler. Thromb. Vasc. Biol.* **1997**, *17*, 2843–2847.
42. Koole, D.; Hurks, R.; Schoneveld, A.; Vink, A.; Golledge, J.; Moran, C.S.; de Kleijn, D.P.; van Herwaarden, J.A.; de Vries, J.; Laman, J.D.; *et al.* Osteoprotegerin is associated with aneurysm diameter and proteolysis in abdominal aortic aneurysm disease. *Arterioscler. Thromb. Vasc. Biol*. **2012**, *32*, 1497–1504.
43. Xiong, W.; MacTaggart, J.; Knispel, R.; Worth, J.; Persidsky, Y.; Baxter, B.T. Blocking TNF-α attenuates aneurysm formation in a murine model. *J. Immunol.* **2009**, *183*, 2741–2746.
44. Kloosterman, W.P.; Plasterk, R.H. The diverse functions of microRNAs in animal development and disease. *Dev. Cell* **2006**, *11*, 441–450.
45. Cai, X.; Hagedorn, C.H.; Cullen, B.R. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* **2004**, *10*, 1957–1966.
46. Yeom, K.; Lee, Y.; Han, J.; Suh, M.R.; Kim, V.N. Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. *Nucleic. Acids Res*. **2006**, *34*, 4622–4629.
47. Krol, J.; Loedige, I.; Filipowicz, W. The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* **2010**, *11*, 597–610.
48. Kim, V.N. MicroRNA precursors in motion: Exportin-5 mediates their nuclear export. *Trends Cell Biol.* **2004**, *14*, 156–159.
49. Hutvagner, G. Small RNA asymmetry in RNAi: Function in RISC assembly and gene Regulation. *FEBS Lett*. **2005**, *579*, 5850–5857.
50. Okamura, K.; Phillips, M.D.; Tyler, D.M.; Duan, H.; Chou, Y.T.; Lai, E.C. The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution. *Nat. Struct. Mol. Biol.* **2008**, *15*, 354–363.
51. Guo, L.; Lu, Z. The fate of miRNA* strand through evolutionary analysis: Implication for degradation as merely carrier strand or potential regulatory molecule? *PLoS One* **2010**, *5*, e11387.
52. Chang, K.W.; Kao, S.Y.; Wu, Y.H.; Tsai, M.M.; Tu, H.F.; Liu, C.J.; Lui, M.T.; Lin, S.C. Passenger strand miRNA miR-31* regulates the phenotypes of oral cancer cells by targeting RhoA. *Oral. Oncol*. **2013**, *49*, 27–33.

53. Yang, J.; Phillips, M.D.; Betel, D.; Mu, P.; Ventura, A.; Siepel, A.C.; Chen, K.C.; Lai, E.C. Widespread regulatory activity of vertebrate microRNA* species. *RNA* **2011**, *17*, 312–326.
54. Zhou, H.; Huang, X.; Cui, H.; Luo, X.; Tang, Y.; Chen, S.; Wu, L.; Shen, N. miR-155 and its star-form partner miR-155* cooperatively regulate type I interferon production by human plasmacytoid dendritic cells. *Blood* **2010**, *116*, 5885–5894.
55. Doench, J.G.; Sharp, P.A. Specificity of microRNA target selection in translational repression. *Genes Dev*. **2004**, *18*, 504–511.
56. Lewis, B.P.; Burge, C.B.; Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **2005**, *120*, 15–20.
57. Huntzinger, E.; Izaurralde, E. Gene silencing by microRNAs: Contributions of translational repression and mRNA decay. *Nat. Rev. Genet*. **2011**, *12*, 99–110.
58. Pillai, R.S.; Bhattacharyya, S.N.; Artus, C.G.; Zoller, T.; Cougot, N.; Basyuk, E.; Bertrand, E.; Filipowicz, W. Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* **2005**, *309*, 1573–1576.
59. Maegdefessel, L.; Azuma, J.; Toh, R.; Deng, A.; Merk, D.R.; Raiesdana, A.; Leeper, N.J.; Raaz, U.; Schoelmerich, A.M.; McConnell, M.V.; *et al.* MicroRNA-21 blocks abdominal aortic aneurysm development and nicotine-augmented expansion. *Sci. Transl. Med*. **2012**, *4*, 122ra22.
60. Leeper, N.J.; Raiesdana, A.; Kojima, Y.; Chun, H.J.; Azuma, J.; Maegdefessel, L.; Kundu, R.K.; Quertermous, T.; Tsao, P.S.; Spin, J.M. MicroRNA-26a is a novel regulator of vascular smooth muscle cell function. *J. Cell Physiol*. **2011**, *226*, 1035–1043.
61. Boon, R.A.; Seeger, T.; Heydt, S.; Fischer, A.; Hergenreider, E.; Horrevoets, A.J.G.; Vinciguerra, M.; Rosenthal, N.; Sciacca, S.; Pilato, M.; *et al*. MicroRNA-29 in aortic dilation: Implications for aneurysm formation. *Circ. Res.* **2011**, *109*, 1115–1119.
62. Maegdefessel, L.; Azuma, J.; Toh, R.; Merk, D.R.; Deng, A.; Chin, J.T.; Raaz, U.; Schoelmerich, A.M.; Raiesdana, A.; Leeper, N.J.; *et al*. Inhibition of microRNA-29b reduces murine abdominal aortic aneurysm development. *J. Clin. Invest*. **2012**, *122*, 497–506.
63. Elia, L.; Quintavalle, M.; Zhang, J.; Contu, R.; Cossu, L.; Latronico, M.V.G.; Peterson, K.L.; Indolfi, C.; Catalucci, D.; Chen, J.; *et al*. The knockout of miR-143 and -145 alters smooth muscle cell maintenance and vascular homeostasis in mice: Correlates with human disease. *Cell Death Differ*. **2009**, *16*, 1590–1598.
64. Lee, Y.S.; Dutta, A. MicroRNAs in cancer. *Annu. Rev. Pathol.* **2009**, *4*, 199–227.
65. Jazbutyte, V.; Thum, T. MicroRNA-21: From cancer to cardiovascular disease. *Curr. Drug Targets* **2010**, *11*, 926–35.
66. Cheng, Y.; Zhang, C. MicroRNA-21 in cardiovascular disease. *J. Cardiovasc. Transl. Res.* **2010**, *3*, 251–255.
67. Davis, B.N.; Hilyard, A.C.; Lagna, G.; Hata, A. SMAD proteins control DROSHA-mediated microRNA maturation. *Nature* **2008**, *454*, 56–61.

68. Ji, R.; Cheng, Y.; Yue, J.; Yang, J.; Liu, X.; Chen, H.; Dean, D.B.; Zhang, C. MicroRNA expression signature and antisense-mediated depletion reveal an essential role of MicroRNA in vascular neointimal lesion formation. *Circ. Res*. **2007**, *100*, 1579–1588.
69. Kang, H.; Hata, A. MicroRNA regulation of smooth muscle gene expression and phenotype. *Curr. Opin. Hematol*. **2012**, *19*, 224–231.
70. Sarkar, J.; Gou, D.; Turaka, P.; Viktorova, E.; Ramchandran, R.; Usha Raj, J. MicroRNA-21 plays a role in hypoxia-mediated pulmonary artery smooth muscle cell proliferation and migration. *Am. J. Physiol. Lung Cell Mol. Physiol.* **2010**, *299*, L861–L871.
71. Wang, M.; Li, W.; Chang, G.; Ye, C.; Ou, J.; Li, X.; Liu, Y.; Cheang, T.; Huang, X.; Wang, S. MicroRNA-21 regulates vascular smooth muscle cell function via targeting tropomyosin 1 in arteriosclerosis obliterans of lower extremities. *Arterioscler. Thromb. Vasc. Biol.* **2011**, *31*, 2044–2053.
72. Song, J.T.; Hu, B.; Qu, H.Y.; Bi, C.L.; Huang, X.Z.; Zhang, M. Mechanical stretch modulates MicroRNA 21 expression, participating in proliferation and apoptosis in cultured human aortic smooth muscle cells. *PLoS One* **2012**, *7*, e47657.
73. Chapman, G.B.; Durante, W.; Hellums, J.D.; Schafer, A.I. Physiological cyclic stretch causes cell cycle arrest in cultured vascular smooth muscle cells. *Am. J. Physiol. Heart Circ. Physiol*. **2000**, *278*, H748–H754.
74. Cheng, W.; Wang, B.; Chen, S.; Chang, H.; Shyu, K. Mechanical stretch induces the apoptosis regulator PUMA in vascular smooth muscle cells. *Cardiovasc. Res*. **2012**, *93*, 181–189.
75. Li, C.; Wernig, F.; Leitges, M.; Hu, Y.; Xu, Q. Mechanical stress-activated PKCdelta regulates smooth muscle cell migration. *FASEB J.* **2003**, *17*, 2106–2108.
76. Weber, M.; Baker, M.B.; Moore, J.P.; Searles, C.D. MiR-21 is induced in endothelial cells by shear stress and modulates apoptosis and eNOS activity. *Biochem. Biophys. Res. Commun*. **2010**, *393*, 643–648.
77. Zhu, S.; Deng, S.; Ma, Q.; Zhang, T.; Jia, C.; Zhuo, D.; Yang, F.; Wei, J.; Wang, L.; Dykxhoorn, D.M.; *et al*. microRNA-10A* and microRNA-21 modulate endothelial progenitor cell senescence via suppressing Hmga2. *Circ. Res*. **2013**, *112*, 152–164.
78. van Rooij, E.; Sutherland, L.B.; Liu, N.; Williams, A.H.; McAnally, J.; Gerard, R.D.; Richardson, J.A.; Olson, E.N. A signature pattern of stress-responsive microRNAs that can evoke cardiac hypertrophy and heart failure. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 18255–18260.
79. Kwiecinski, M.; Noetel, A.; Elfimova, N.; Trebicka, J.; Schievenbusch, S.; Strack, I.; Molnar, L.; von Brandenstein, M.; Töx, U.; Nischt, R. *et al*. Hepatocyte growth factor (HGF) inhibits collagen I and IV synthesis in hepatic stellate cells by miRNA-29 induction. *PLoS One* **2011**, *6*, e24568.

80. Wang, B.; Komers, R.; Carew, R.; Winbanks, C.E.; Xu, B.; Herman-Edelstein, M.; Koh, P.; Thomas, M.; Jandeleit-Dahm, K.; Gregorevic, P.; *et al.* Suppression of microRNA-29 expression by TGF-β1 promotes collagen expression and renal fibrosis. *J. Am. Soc. Nephrol.* **2012**, *23*, 252–265.
81. Maurer, B.; Stanczyk, J.; Jüngel, A.; Akhmetshina, A.; Trenkmann, M.; Brock, M.; Kowal-Bielecka, O.; Gay, R.E.; Michel, B.A.; Distler, J.H.; *et al*. MicroRNA-29, a key regulator of collagen expression in systemic sclerosis. *Arthritis Rheum* **2010**, *62*, 1733–1743.
82. Ogawa, T.; Iizuka, M.; Sekiya, Y.; Yoshizato, K.; Ikeda, K.; Kawada, N. Suppression of type I collagen production by microRNA-29b in cultured human stellate cells. *Biochem. Biophys. Res. Commun*. **2010**, *391*, 316–321.
83. Boettger, T.; Beetz, N.; Kostin, S.; Schneider, J.; Krüger, M.; Hein, L.; Braun, T. Acquisition of the contractile phenotype by murine arterial smooth muscle cells depends on the Mir143/145 gene cluster. *J. Clin. Invest*. **2009**, *119*, 2634–2647.
84. Cheng, Y.; Liu, X.; Yang, J.; Lin, Y.; Xu, D.Z.; Lu, Q.; Deitch, E.A.; Huo, Y.; Delphin, E.S.; Zhang, C. MicroRNA-145, a novel smooth muscle cell phenotypic marker and modulator, controls vascular neointimal lesion formation. *Circ. Res*. **2009**, *105*, 158–166.
85. Cordes, K.R.; Sheehy, N.T.; White, M.P.; Berry, E.C.; Morton, S.U.; Muth, A.N.; Lee, T.H.; Miano, J.M.; Ivey, K.N.; Srivastava, D. miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature* **2009**, *460*, 705–710.
86. Xin, M.; Small, E.M.; Sutherland, L.B.; Qi, X.; McAnally, J.; Plato, C.F.; Richardson, J.A.; Bassel-Duby, R.; Olson, E.N. MicroRNAs miR-143 and miR-145 modulate cytoskeletal dynamics and responsiveness of smooth muscle cells to injury. *Genes Dev*. **2009**, *23*, 2166–2178.
87. Quintavalle, M.; Elia, L.; Condorelli, G.; Courtneidge, S.A. MicroRNA control of podosome formation in vascular smooth muscle cells *in vivo* and *in vitro*. *J. Cell Biol*. **2010**, *189*, 13–22.
88. Hergenreider, E.; Heydt, S.; Tréguer, K.; Boettger, T.; Horrevoets, A.J.; Zeiher, A.M.; Scheffer, M.P.; Frangakis, A.S.; Yin, X.; Mayr, M.; *et al.* Atheroprotective communication between endothelial cells and smooth muscle cells through miRNAs. *Nat. Cell Biol*. **2012**, *14*, 249–256.
89. Golledge, A.L.; Walker, P.; Norman, P.E.; Golledge, J. A systematic review of studies examining inflammation associated cytokines in human abdominal aortic aneurysm samples. *Dis. Markers* **2009**, *26*, 181–188.
90. Harris, T.A.; Yamakuchi, M.; Ferlito, M.; Mendell, J.T.; Lowenstein, C.J. MicroRNA-126 regulates endothelial expression of vascular cell adhesion molecule 1. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1516–1521.
91. Asgeirsdóttir, S.A.; van Solingen, C.; Kurniati, N.F.; Zwiers, P.J.; Heeringa, P.; van Meurs, M.; Satchell, S.C.; Saleem, M.A.; Mathieson, P.W.; Banas, B.; *et al.* MicroRNA-126 contributes to renal microvascular heterogeneity of VCAM-1 protein expression in acute inflammation. *Am. J. Physiol. Renal Physiol.* **2012**, *302*, F1630–F1639.

92. Zernecke, A.; Bidzhekov, K.; Noels, H.; Shagdarsuren, E.; Gan, L.; Denecke, B.; Hristov, M.; Köppel, T.; Jahantigh, M.N.; Lutgens, E.; *et al.* Delivery of microRNA-126 by apoptotic bodies induces CXCL12-dependent vascular protection. *Sci. Signal*. **2009**, *2*, ra81.
93. Vasa-Nicotera, M.; Chen, H.; Tucci, P.; Yang, A.L.; Saintigny, G.; Menghini, R.; Mahè, C.; Agostini, M.; Knight, R.A.; Melino, G.; *et al.* miR-146a is modulated in human endothelial cell with aging. *Atherosclerosis* **2011**, *217*, 326–330.
94. Sun, S.G.; Zheng, B.; Han, M.; Fang, X.M.; Li, H.X.; Miao, S.B.; Su, M.; Han, Y.; Shi, H.J.; Wen, J.K. miR-146a and Kruppel-like factor 4 form a feedback loop to participate in vascular smooth muscle cell proliferation. *EMBO Rep*. **2011**, *12*, 56–62.
95. Suárez, Y.; Wang, C.; Manes, T.D.; Pober, J.S. Cutting edge: TNF-induced microRNAs regulate TNF-induced expression of E-selectin and intercellular adhesion molecule-1 on human endothelial cells: Feedback control of inflammation. *J. Immunol*. **2010**, *184*, 21–25.
96. Zhu, N.; Zhang, D.; Chen, S.; Liu, X.; Lin, L.; Huang, X.; Guo, Z.; Liu, J.; Wang, Y.; Yuan, W.; Qin, Y. Endothelial enriched microRNAs regulate angiotensin II-induced endothelial inflammation and migration. *Atherosclerosis* **2011**, *215*, 286–293.
97. Nazari-Jahantigh, M.; Wei, Y.; Noels, H.; Akhtar, S.; Zhou, Z.; Koenen, R.R.; Heyll, K.; Gremse, F.; Kiessling, F.; Grommes, J.; *et al*. MicroRNA-155 promotes atherosclerosis by repressing Bcl6 in macrophages. *J. Clin. Invest.* **2012**, *122*, 4190–4202.
98. Donners, M.M.; Wolfs, I.M.; Stöger, L.J.; van der Vorst, E.P.; Pöttgens, C.C.; Heymans, S.; Schroen, B.; Gijbels, M.J.; de Winther, M.P. Hematopoietic miR155 deficiency enhances atherosclerosis and decreases plaque stability in hyperlipidemic mice. *PLoS One* **2012**, *7*, e35877.
99. Pahl, M.C.; Derr, K.; Gäbel, G.; Hinterseher, I.; Elmore, J.R.; Schworer, C.M.; Peeler, T.C.; Franklin, D.P.; Gray, J.L.; Carey, D.J. *et al*. MicroRNA expression signature in human abdominal aortic aneurysms. *BMC Med. Genomics* **2012**, *5*, 25.
100. Small, E.M.; Olson, E.N. Pervasive roles of microRNAs in cardiovascular biology. *Nature* **2011**, *469*, 336–342.
101. Van Rooij, E.; Olson, E.N. MicroRNA therapeutics for cardiovascular disease: Opportunities and obstacles. *Nat. Rev. Drug Discov.* **2012**, *11*, 860–872.
102. Small, E.M.; Frost, R.J.; Olson, E.N. MicroRNAs add a new dimension to cardiovascular disease. *Circulation* **2010**, *121*, 1022–1032.
103. Van Rooij, E.; Purcell, A.L.; Levin, A.A. Developing microRNA therapeutics. *Circ. Res*. **2012**, *110*, 496–507.
104. Stenvang, J.; Petri, A.; Lindow, M.; Obad, S.; Kauppinen, S. Inhibition of microRNA function by antimiR oligonucleotides. *Silence* **2012**, *3*, 1.
105. Mendell, J.T.; Olson, E.N. MicroRNAs in stress signaling and human disease. *Cell* **2012**, *148*, 1172–1187.

106. Janssen, H.L.A.; Reesink, H.W.; Lawitz, E.J.; Zeuzem, S.; Rodriguez-Torres, M.; Patel, K.; van der Meer, A.J.; Patick, A.K.; Chen, A.; Zhou, Y.; *et al.* Treatment of HCV infection by targeting microRNA. *N. Engl. J. Med.* **2013**, *368*, 1685–1694.
107. Mishra, P.K.; Tyagi, N.; Kumar, M.; Tyagi, S.C. MicroRNAs as a therapeutic target for cardiovascular diseases. *J. Cell Mol. Med.* **2009**, *13*, 778–789.
108. Pramanik, D.; Campbell, N.R.; Karikari, C.; Chivukula, R.; Kent, O.A.; Mendell, J.T.; Maitra, A. Restitution of tumor suppressor microRNAs using a systemic nanovector inhibits pancreatic cancer growth in mice. *Mol. Cancer Ther.* **2011**, *10*, 1470–1480.
109. Trang, P.; Wiggins, J.F.; Daige, C.L.; Cho, C.; Omotola, M.; Brown, D.; Weidhaas, J.B.; Bader, A.G.; Slack, F.J. Systemic delivery of tumor suppressor microRNA mimics using a neutral lipid emulsion inhibits lung tumors in mice. *Mol. Ther*. **2011**, *19*, 1116–1122.
110. Moxon, J.V.; Parr, A.; Emeto, T.I.; Walker, P.; Norman, P.E.; Golledge, J. Diagnosis and monitoring of abdominal aortic aneurysm: Current status and future prospects. *Curr. Probl. Cardiol*. **2010**, *35*, 512–548.
111. D'Alessandra, Y.; Devanna, P.; Limana, F.; Straino, S.; di Carlo, A.; Brambilla, P.G.; Rubino, M.; Carena, M.C.; Spazzafumo, L.; de Simone, M.; *et al.* Circulating microRNAs are new and sensitive biomarkers of myocardial infarction. *Eur. Heart J.* **2010**, *31*, 2765–2773.
112. Fichtlscherer, S.; de Rosa, S.; Fox, H.; Schwietz, T.; Fischer, A.; Liebetrau, C.; Weber, M.; Hamm, C.W.; Röxe, T.; Müller-Ardogan, M.; *et al.* Circulating microRNAs in patients with coronary artery disease. *Circ. Res*. **2010**, *107*, 677–684.
113. Tijsen, A.J.; Creemers, E.E.; Moerland, P.D.; de Windt, L.J.; van der Wal, A.C.; Kok, W.E.; Pinto, Y.M. MiR423-5p as a circulating biomarker for heart failure. *Circ. Res.* **2010**, *106*, 1035–1039.
114. Zampetaki, A.; Willeit, P.; Tilling, L.; Drozdov, I.; Prokopi, M.; Renard, J.M.; Mayr, A.; Weger, S.; Schett, G.; Shah, A.; *et al*. Prospective study on circulating MicroRNAs and risk of myocardial infarction. *J. Am. Coll Cardiol*. **2012**, *60*, 290–299.
115. Engelhardt, S. Small RNA biomarkers come of age. *J. Am. Coll Cardiol*. **2012**, *60*, 300–303.
116. Zampetaki, A.; Mayr, M. Analytical challenges and technical limitations in assessing circulating miRNAs. *Thromb. Haemost*. **2012**, *108*, 592–598.
117. Kin, K.; Miyagawa, S.; Fukushima, S.; Shirakawa, Y.; Torikai, K.; Shimamura, K.; Daimon, T.; Kawahara, Y.; Kuratani, T.; Sawa, Y. Tissue- and plasma-specific microRNA signatures for atherosclerotic abdominal aortic aneurysm. *J. Am. Heart Assoc*. **2012**, *1*, e000745.

MDPI AG
Klybeckstrasse 64
4057 Basel, Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
http://www.mdpi.com/

IJMS Editorial Office
E-mail: ijms@mdpi.com
http://www.mdpi.com/journal/ijms

MDPI • Basel • Beijing
ISBN 978-3-03842-010-1
www.mdpi.com