



Regulation by non-coding RNAs

Volume 2

Edited by

Nicholas Delihias

Printed Edition of the Special Issue Published in *IJMS*



Nicholas Delihias (Ed.)

Regulation by non-coding RNAs

Volume 2



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 Delihias

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)

© 2014 by the authors; licensee MDPI, Basel, Switzerland. All articles in this volume are Open Access distributed under the Creative Commons Attribution 3.0 license (<http://creativecommons.org/licenses/by/3.0/>), which allows users to download, copy and build upon published articles even for commercial purposes, as long as the author and publisher are properly credited, which ensures maximum dissemination and a wider impact of our publications. However, the dissemination and distribution of copies of this book as a whole is restricted to MDPI, Basel, Switzerland.

Table of Contents

Volume 2

| | |
|---|-----|
| Nicholas Delihias Preface <i>Guest Editor</i> | 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 <i>Int. J. Mol. Sci.</i> 2013, <i>14</i> (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 <i>Int. J. Mol. Sci.</i> 2013, <i>14</i> (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 <i>Dies1</i> in a TGF-Independent Manner Reprinted from <i>Int. J. Mol. Sci.</i> 2013, <i>14</i> (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 <i>Int. J. Mol. Sci.</i> 2013, <i>14</i> (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 <i>Int. J. Mol. Sci.</i> 2013, <i>14</i> (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 <i>Int. J. Mol. Sci.</i> 2013, <i>14</i> (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 G ebelin, 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. Nicol as and Rosa M. Ruiz-V azquez

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 Delihias

Guest Editor

Reprinted from *IJMS*. Cite as: Morceau, F.; Chateauvieux, S.; Gaigneaux, A.; Dicato, M.; Diederich, M. Long and Short Non-Coding RNAs as Regulators of Hematopoietic Differentiation. *Int. J. Mol. Sci.* **2013**, *14*, 14744-14770.

Review

Long and Short Non-Coding RNAs as Regulators of Hematopoietic Differentiation

Franck Morceau^{1,†}, Sébastien Chateauvieux^{1,†}, Anthoula Gaigneaux¹, Mario Dicato¹ and Marc Diederich^{2,*}

¹ Laboratoire de Biologie Moléculaire et Cellulaire du Cancer, Kirchberg Hospital, 9 Rue Edward Steichen, 2540, Luxembourg; E-Mails: franck.morceau@lbmcc.lu (F.M.); sebastien.chateauvieux@lbmcc.lu (S.C.); anthoula.gaigneaux@lbmcc.lu (A.G.); mdicato@gmail.com (M.D.)

² Department of Pharmacy, College of Pharmacy, Seoul National University, Seoul 151-742, Korea

† *These authors contributed equally to this work.*

* Author to whom correspondence should be addressed; E-Mail: marc.diederich@lbmcc.lu; Tel.: +82-2-880-8919.

Received: 3 June 2013; in revised form: 5 July 2013 / Accepted: 9 July 2013 /

Published: 15 July 2013

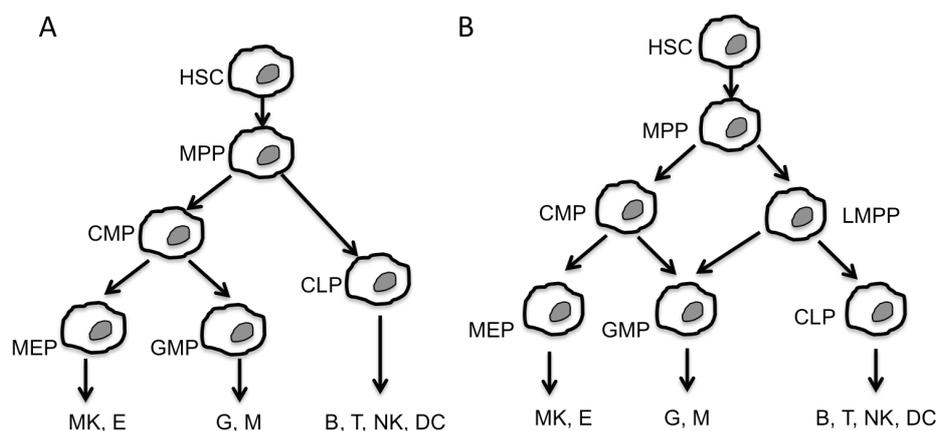
Abstract: Genomic analyses estimated that the proportion of the genome encoding proteins corresponds to approximately 1.5%, while at least 66% are transcribed, suggesting that many non-coding DNA-regions generate non-coding RNAs (ncRNAs). The relevance of these ncRNAs in biological, physiological as well as in pathological processes increased over the last two decades with the understanding of their implication in complex regulatory networks. This review particularly focuses on the involvement of two large families of ncRNAs, namely microRNAs (miRNAs) and long non-coding RNAs (lncRNAs) in the regulation of hematopoiesis. To date, miRNAs have been widely studied, leading to a wealth of data about processing, regulation and mechanisms of action and more specifically, their involvement in hematopoietic differentiation. Notably, the interaction of miRNAs with the regulatory network of transcription factors is well documented whereas roles, regulation and mechanisms of lncRNAs remain largely unexplored in hematopoiesis; this review gathers current data about lncRNAs as well as both potential and confirmed roles in normal and pathological hematopoiesis.

Keywords: ncRNA; miRNA; lncRNA; transcription factors; regulatory network; erythropoiesis; leukemia; lymphoma; differentiation

1. Introduction

Hematopoiesis is the physiological process leading to the production of all circulating blood cells. Hematopoietic stem cells (HSCs) are pluripotent cells with high self-renewal capacity. Asymmetric division is the main approved model for self-renewal and commitment towards a specific differentiation pathway [1], which allows maintaining a steady state HSCs population by preserving hematopoietic homeostasis. As shown in Figure 1, at least two models have been proposed for the hematopoietic hierarchy.

Figure 1. Two models of the hierarchical hematopoiesis process. Hematopoietic cell differentiation proceeds by successive hierarchical maturation steps. (A) Pluripotent hematopoietic stem cells (HSC) give rise to multipotential progenitors (MPP) leading to common lymphocyte progenitors (CLP) and common myeloid progenitors (CMP). CLPs directly generate cells of the immune system. CMPs give rise to megakaryocyte-erythroid progenitors (MEP) and granulocyte-macrophage progenitors (GMP); (B) The alternative model differs by the involvement of an intermediate lymphoid-primed multipotential progenitor (LMPP) to generate GMP and CMP. Both models lead to the production of differentiated hematopoietic cells (M, monocyte; G, granulocytes; E, Erythrocyte; MK, megakaryocyte; T and B, lymphocytes; NK, natural killers DC, dendritic cells).



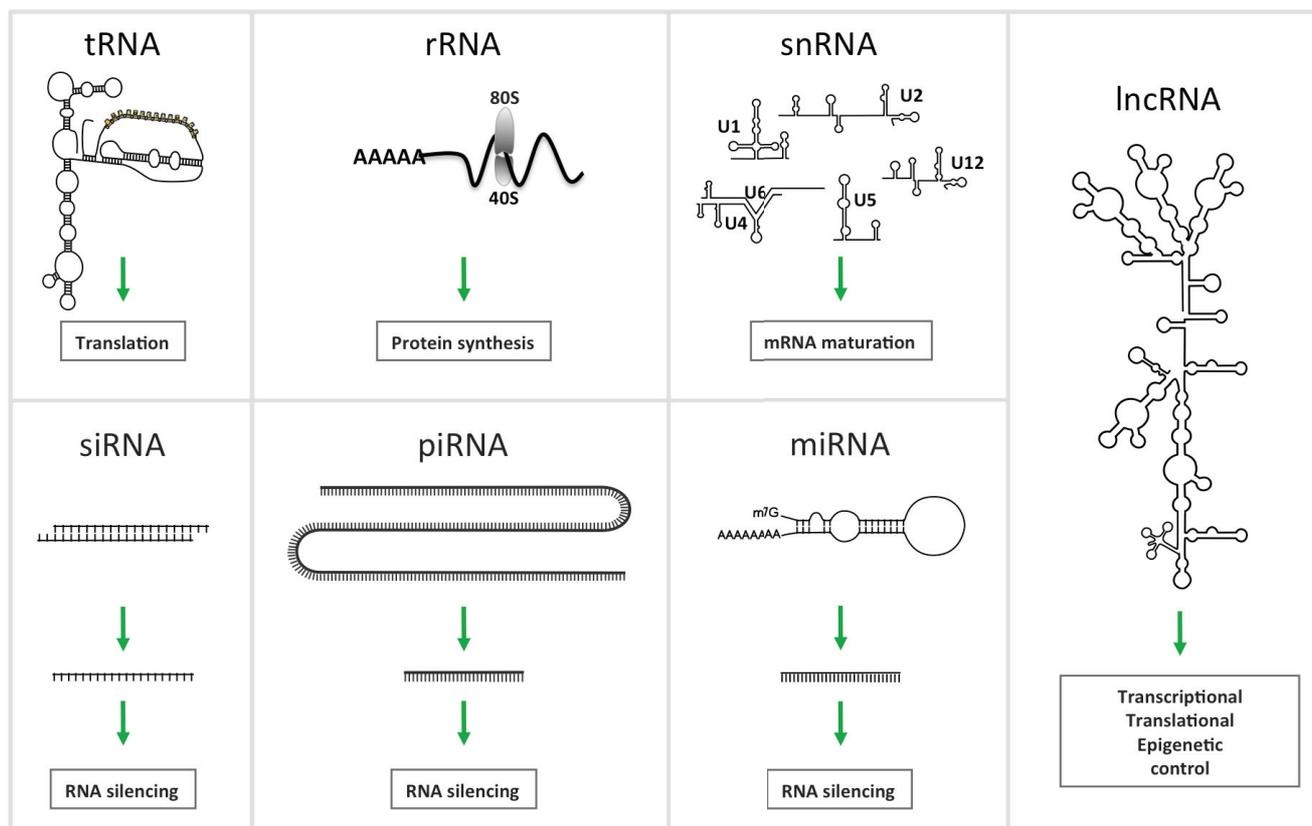
Cell fate decision is regulated by a complex network of extra- and intracellular regulatory factors to ensure HSCs commitment, survival, differentiation as well as maturation depending on physiological requirements. Many cytokines and growth factors activate cell-signaling pathways controlling posttranslational modifications of transcription factors (TF), protein-protein interactions, enzyme activation, binding of proteins to both DNA and RNA, stability of proteins and mRNAs and activation of epigenetic regulators. The ultimate step of this intricate combination of activated and inactivated proteins is the transcription or silencing of genes. To reach this status of differentiated cells, a continuous fine-tuning of genetic programs is essential for the control of the rhythm of cellular

divisions, which decreases with cell differentiation, to prevent early cell death of progenitors and precursors as well as to express specific phenotypes. Besides key factors required for self-renewal and commitment of HSCs, complex regulatory networks of TFs mark the cell paths and participate up until their destiny. During hematopoiesis, TF activities depend on interactions between themselves and with cofactors. In the myeloid and lymphoid branches of hematopoiesis, several TFs are differentially involved in divergent pathways. Further TF families are crucial in hematopoietic cell fate decisions. Lichtinger *et al.*, showed that the combination and interdependent regulation of T-cell acute lymphocytic leukemia protein (TAL)1/ stem cell leukemia (SCL), Friend leukemia integration (FLI)1 and CCAAT-enhancer-binding protein (C/EBP) is required for the correct temporal expression of lineage specific genes [2]. Malinge *et al.*, reported that the TF Ikaros, which is expressed in erythro-megakaryocyte progenitors [3], blocks terminal megakaryocytic maturation through the inhibition of GATA-1 expression in correlation with target gene inhibition including LIM domain only (Lmo)2. This possibly also involves the inhibition of GATA-1 cofactor Friend of GATA (FOG)-1 by Ikaros. As an alternative mechanism, Notch signaling pathway is also affected by Ikaros-mediated inhibition of megakaryocytic differentiation. Indeed, Ikaros was shown to inhibit the Notch-induced megakaryocytic pattern from hematopoietic progenitors [4]. Lmo2 serves as a bridge between GATA-1 and the SCL complex. Tripic *et al.*, showed that SCL increases GATA-1 transcriptional activity in the murine erythroid cell line G1E-ER4, which displays an inducible GATA-1 construct. Conversely, in the absence of Lmo2, GATA-1 plays a repressive role on target genes in correlation with the absence of SCL complex association.

Besides, transcription factors play a central role in hematopoietic development, from HSC commitment to terminal differentiation and death. They are differentially and temporally expressed along the differentiation process. Furthermore, an additional level of regulation has joined the network of regulatory factors with the involvement of non-coding RNAs (ncRNAs). This term defines RNA transcripts without protein-coding capacity including constitutively expressed housekeeping small RNAs, ribosomal (rRNAs), transfer (tRNAs), small nuclear (snRNAs), small nucleolar (snoRNAs), transfer-messenger (tmRNAs) and telomerase RNAs. Furthermore, regulatory ncRNAs have been also described including the family of microRNAs (miRNAs or miRs) and the family of larger regulatory ncRNAs, long non-coding (lncRNAs) (Figure 2). This review updates knowledge about regulatory ncRNAs in hematopoiesis by especially focusing on miRNAs as well as the lncRNAs. Particular attention is given to the miRNAs/transcription factors forming regulatory network.

2. Non-Coding RNA

Genomic analyses determined that the proportion of the genome coding for proteins corresponds approximately to 1.5%, while at least two thirds are transcribed, suggesting that many non-coding sequences are transcribed into ncRNA. In addition, while protein-coding sequences yet represent a minority of the genome of multicellular organisms, their proportion further declines with increasing complexity of the organism, with a concomitant increase in the amount of non-coding regions in intergenic or intronic sequences [5–7]. A non-coding RNA (ncRNA) or non-protein-coding RNA (npcRNA) is a functional RNA molecule not translated into a protein.

Figure 2. Schematic representation of different non coding RNAs.

The discovery of RNA interference (RNAi) [8] in *C. elegans* and earlier, the identification of a new class of small RNAs known as miRNAs [9] led to greater attention to ncRNA and their involvement in the regulation of biological processes. The total number of ncRNAs is still unknown, but through transcriptomic and bioinformatic studies, one suspects the existence of several thousands of ncRNAs. A large number of them remain to be fully identified, and the functions of most of them have not yet been validated. NcRNAs are grouped into several RNA families, subclassified according to their function, size, structure and conservation (Figure 2). We thus find the ubiquitous and well-known transfer RNA (tRNA) required as the physical link between the nucleotide sequence of nucleic acids (mRNA) and the amino acid sequence of proteins; The ribosomal RNA (rRNA) as RNA component of the ribosome, and essential for protein synthesis in all living organisms. Small nuclear RNAs (snRNAs), form a class of RNA molecules with an average length of 150 nucleotides localized within the nucleus of eukaryotic cells. Their primary function is pre-mRNA (hnRNA) processing, for which they are always associated with a set of specific proteins and the complexes are referred to as small nuclear ribonucleoproteins (snRNP). A subclass of snRNA is called small nucleolar RNAs (snoRNAs) localized in the nucleolus and implicated in the maturation of RNA molecules by guidance of chemical modifications targeting mainly rRNAs, tRNAs and snRNAs. Piwi-interacting RNAs (piRNAs) form the largest class of small non-coding RNA molecules expressed in animal cells and form the third class of small RNA silencers. They form RNA-protein complexes by interacting with Piwi proteins and are required for both epigenetic and post-transcriptional gene silencing of retrotransposons and other genetic elements in germ line cells, particularly during spermatogenesis. Small Interfering RNAs or silencing RNAs (siRNAs) are a class of short double stranded RNA molecules (20–25 bp) playing a

role in the RNA interference (RNAi) pathway. They modulate expression of specific genes by interfering with RNA translation by complementary nucleotide sequences. Post-transcriptional activity of siRNAs was first discovered in plants [10], and their possible use in mammalian cells was rapidly shown [11], evidencing the potential as a biomedical research tool. Interference induced by siRNAs share a large part of the signaling pathways used by miRNA naturally present in mammalian cells. The presence of siRNA encoded by the human genome has never been shown. Extracellular or exosomal RNA (exRNA) designate RNA species present in body fluids (venous blood, saliva, breast milk, urine, semen, menstrual blood or vaginal fluid) composed by mRNA, tRNA, miRNA, siRNA and lncRNA, generally enclosed within vesicular bodies preventing their digestion. Biochemical evidence supports the idea that exRNA uptake is a common process, suggesting new pathways for intercellular communication. The relative abundance of certain exRNAs can be correlated to cellular signaling or specific disease states. Recently a study characterized the population of RNA in human plasma by deep sequencing, and demonstrated that most abundant RNAs are miRNAs (42.32%), followed by ribosomal RNAs (9.16% of all mapable counts), long non-coding RNAs (3.36%), piwi-interacting RNAs (1.31%), transfer RNAs (1.24%), small nuclear RNAs (0.18%), and small nucleolar RNAs (0.01%) [12].

Finally, micro RNAs (miRNAs) and long noncoding RNAs (lncRNAs) form the two families of ncRNA more widely described in our review.

2.1. MiR and Hematopoiesis

2.1.1. MiRNAs, a Family of Regulatory ncRNAs

While the first small endogenous ncRNAs (*lin-4/let-7*) were described in *C. elegans* in 1993 [13], these RNAs became by now a large family of regulatory ncRNAs referred as microRNAs (miRNAs) [14–16]. These single stranded RNAs are characterized by their size of about 17–25 nucleotides and are highly conserved during evolution [14]. Mature and functional miRNAs result from a multistage process and their expression leads to post-transcriptional silencing of target genes through epigenetic regulatory functions by repressing mRNA translation [17]. Therefore miRNAs are also involved in a complex cellular network of gene regulation. On one hand, most miRNAs are able to target several mRNAs while a specific mRNA can be targeted by several miRNAs [18,19]. On the other hand, miRNAs target TF mRNAs and in turn TFs regulate miRNA gene expression. Expression of miRNAs is then tightly regulated at both transcriptional and post-transcriptional levels. Primary miRNAs (pri-miRNA) result from the transcription of genes requiring RNA polymerase II or III, and critical epigenetic regulation steps [20]. As for TFs, miRNAs display tissue and developmental specificities. In correlation with their spatial and temporal expression, miRNAs contribute to the regulation of embryogenesis, cell proliferation, differentiation and death [21]. Therefore, deregulation of miRNA expression, caused by genetic alterations, transcriptional or processing failures is incriminated in the development of many human diseases, including cancer.

Transcriptional regulation of miRNAs is not yet fully understood while processing of the transcript in mature miRNA has been described in detail [20,22]. Briefly, in the canonical miRNA biosynthetic pathway, the transcription of miRNA genes by RNA polymerase II (or III for miRNAs encoded within

Alu repeat sequences) [23] generates a pri-miRNA displaying hairpin structures. In fact, pri-miRNA transcripts can be organized in clusters encoding multiple miRNA sequences [24]. In the nucleus, pri-miRNAs are then cleaved between the hairpin structures to generate pre-miRNAs of 60 to 110 nt. This cleavage is performed by the nuclear endoribonuclease Drosha (RNaseIII) and DiGeorge critical region 8 (DGCR8) cofactor (Pasha in *Drosophila*)-forming complex. Pre-miRNAs are then exported to the cytoplasm by Exportin-5 in a Ran-GTP dependent manner, where they undergo further processing. Cytoplasmic endonuclease RNase III Dicer associated to the transactivating response RNA-binding protein (TRBP) and the protein activator of PKR (PACT), cleave the stem loop of the hairpin structure to form an asymmetric RNA duplex of about 22 nt. Double strand RNA interacts with Argonaute within the RNA-induced silencing complex (RISC) in order to select the guide strand matching with the 3'-untranslated region (UTR) of the target mRNA, thus forming the miRISC. The complementary "passenger" strand of the duplex is released and degraded. The guide strand allows the RISC-mediated inhibition of translation of the target mRNA as the main mechanism of action. Nevertheless, destabilization of target mRNAs through the accelerated deadenylation was reported as an alternative mechanism for miRNAs action [25,26]. Recently, an additional mode of action was proposed in which endogenous miRNAs act at the transcriptional level to silence genes, involving chromatin remodeling complexes [27].

2.1.2. MiRNAs Regulate the Regulatory Network during Hematopoiesis

The crucial role of miRNAs in the regulation of specific steps of hematopoiesis is now well-documented based on the modulation of their expression during HSC differentiation processes. Furthermore, in the hematopoietic system, numerous miRNAs were described to affect translation or stability of mRNAs including mRNAs encoding TFs. Thus, interplay of TFs and miRNAs are considered essential regulators of gene expression in hematopoiesis (Figure 3).

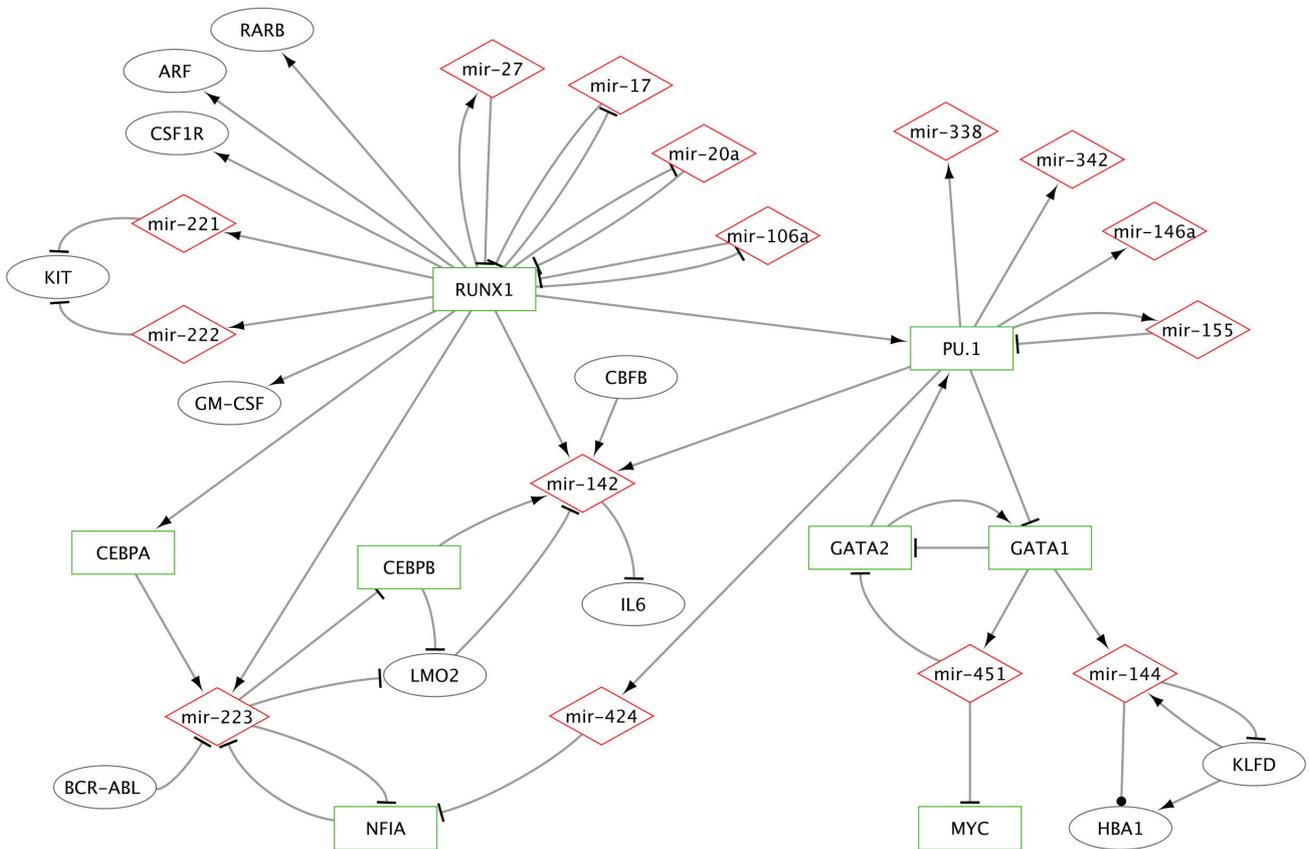
2.1.2.1. MiR-223 Is a Key Hematopoietic miRNA

One of the most extensively studied miRNA in hematopoiesis is miR-223 also described as the "fine-tuner" of granulocytic differentiation, maturation and function while its expression decreases during monocytic, erythroid [28] and mast-cell differentiation [29–32]. Nevertheless, up-regulation of miR-223 was observed by Lu *et al.* [33] and Choong *et al.*, during EPO-mediated erythroid differentiation of human CD34⁺ HSPCs [34], whereas its expression was down-regulated in K562 cells. Finally, a functional study revealed a role for miR-223 in erythroid differentiation [35]. According to the various expression levels in different types of hematopoietic cells, miR-223 is linked to several major specific transcription factors.

Expression of miR-223 is associated with granulocyte differentiation and was shown to increase in retinoic acid (RA)-induced granulocytic differentiation of leukemic cells. C/EBP α is a key factor for the commitment of hematopoietic cells towards the granulocytic lineage and its RA-mediated upregulation was anticipating miR-223 activation in NB4 and HL60 promyelocytic cells, suggesting a role for this TF in the transcriptional regulation of miR-223. In correlation, two putative C/EBP α binding sites were identified upstream of the region encoding the pre-miR-223 sequence. Results were confirmed by transfection of expression constructs, which demonstrated the region containing the

C/EBP α binding sites to be required for miR-223 responsiveness to RA. Moreover, chromatin immunoprecipitation (ChiP) assays showed that C/EBP α binds *in vivo* to miR-223 promoter. Besides C/EBP α , nuclear factor I-A (NFI-A) binding site was described in the miR-223 promoter [29], overlapping one of the two C/EBP α sites. This TF is also involved in the regulation of miR-223 expression and maintains its basal level in undifferentiated cells while C/EBP α over-expression substitutes NFI-A in RA-induced granulocytic differentiation eventually leading to miR-223 up-regulation. Furthermore, NFI-A is targeted by miR-223. Altogether, granulocytic differentiation correlates well with miR-223-mediated inhibition of NFI-A mRNA translation, therefore preventing competition with C/EBP α [29].

Figure 3. Partial representation of the network involving microRNAs (mir) and regulatory proteins in hematopoiesis. Network was built using Cytoscape 2.8.2 software [36]. Transcription factors are represented as green rectangles, miRNAs as red diamonds, and other proteins as grey ellipses. Arrows depict relationships: arrows (activation), T (inhibition), circle (undetermined/binding). RARB, Retinoic acid receptor B, ARF, Alternate Reading Frame, CSF1R, colony-stimulating factor 1 receptor; BCR-ABL, breakpoint cluster region-Abelson; CBF β , Core-binding factor subunit beta; IL6, interleukine 6; LMO2, LIM domain only 2; HBA1, hemoglobin A1; KLFD, Krüppel like factor D; RUNX1, Runt-related transcription factor 1; CEBP α , CCAAT/enhancer-binding protein alpha; CEBP β , CCAAT/enhancer-binding protein beta; GM-CSF, granulocyte macrophage-colony stimulating factor; NFIA, nuclear factor I-A.



Interestingly, a recent study reported an alternative mechanism for NFI-A regulation by endogenous miR-223. Authors demonstrated the ability of miR-223 to affect transcription of NFI-A gene in addition to a post-transcriptional effect [27]. Authors first observed that nuclear import of miR-223 increased during granulocytic differentiation of myeloid precursors. Using confocal microscopy, results showed that the nuclear compartmentalization of miR-223 was increased in RA treated HL60 cells and primary acute promyelocytic leukemia (APL) blasts cells. ChIP assays performed in HL60 cells revealed *in vivo* presence of miR-223 at complementary sequences of the NFI-A promoter. Moreover, mutation of miR-223 binding sites abolished miR-223-mediated repression of NFI-A promoter reporter constructs [27]. Beyond the increasing variety of miRNA function in hematopoiesis, these findings reveal a novel mechanism of transcriptional control.

According to the pivotal role of miR-223 in hematopoiesis, its deregulation contributes to leukemia development. Notably, it was recently established that the fusion oncokine breakpoint cluster region (Bcr)-Abelson (Abl) repressed miR-223 expression in chronic myeloid leukemia (CML). miR-223 expression was also significantly decreased in B cell lymphoproliferative disorders [37] and could serve as a prognostic factor for CLL and CML patients.

Transcription factors Satb1 and Runt-related transcription factor 1 (RUNX1) or acute myeloid leukemia 1 protein (AML1) were recently described as central regulators of the early stage of HSCs commitment and self-renewal. Many publications report RUNX1 as an essential TF for hematopoiesis initiation at both adult and embryonic levels [38–40]. In mice embryos, expression of this TF has been localized in all sites giving rise to hematopoietic cells. Hemogenic sites such as intra-aortic hematopoietic clusters are absent in RUNX1^{-/-} embryos [41] and the hematopoietic defect in RUNX1^{-/-} embryonic stem cells was rescued by ectopic expression of RUNX1 in mice [39,42]. Recently, it was also shown that the isoform RUNX1a enhances hematopoietic lineage commitment from human embryonic stem cells [43]. RUNX1 is required for both embryonic and post-natal hematopoiesis. It is also involved in the control of genes essential for myeloid differentiation. A connection between this TF and several miRNAs related to specific myeloid genes has been reported. Notably, RUNX1 also regulates miR-223 gene transcription in myeloid precursors, maintaining chromatin in a transcriptionally active state. Consequently, the fusion protein RUNX1/ETO (AML1/ETO), in which RUNX1 is inactivated, specifically triggers transcriptional silencing of miR-223 [44]. RUNX1/ETO, which causes acute myeloid leukemia (AML), acts as a dominant-negative repressor of RUNX1 target genes [45], including colony-stimulating factor 1 receptor (*CSF1R/c-fms*), GM-CSF [46,47], tumor suppressor p14(*ARF*) and the retinoic acid receptor β (*RARB*) [44].

Although up-regulation of miR-223 was correlated to decreased NFI-A expression and granulocytic differentiation, regulation of specific myeloid genes has not been yet established. Nevertheless, miR-223 together with constitutively expressed miR-142 was shown to decrease proliferation required for myeloid differentiation [48]. It is well established that miR-142 is implicated in the regulation of T cell function and development and B cell lymphoma.

2.1.2.2. MiRs-223/142/155 and Specific TFs as a Miniature Regulatory Network

Besides targeting NFI-A, miR-223 also regulates Lmo2-L/-S isoforms and C/EBP β expression in myeloid cells. A miR-223-mediated decrease of cell proliferation occurred through a miR223-C/

EBP β -Lmo2-miR142 pathway. This model is based on the findings that miR-223 targets C/EBP β as well as Lmo2 mRNAs and that C/EBP β regulates transcription of Lmo2, which is then down-regulated at both transcriptional and posttranscriptional levels. This leads to the expression of miR-142 and cell proliferation attenuation since Lmo2 negatively regulates miR-142 gene transcription [48]. Lmo2 is also involved in erythroid-specific gene expression [49] through the formation of a DNA binding complex involving also TAL1, E2A, Ldb/LNI-1 and GATA1 [50]. In agreement with the regulation of Lmo2 expression by miR-223, it has been reported that forced expression of miR-223 inhibited erythroid development of CD34⁺ HSCs [35]. Altogether these results hint for a negative role of this miRNA in erythropoiesis whereas is positive in myeloid differentiation. Similarly, NFI-A up-regulation promotes erythropoiesis while its silencing promotes granulopoiesis in correlation with miR-223 modulation [29,51,52].

Besides Lmo2, Sun *et al.*, recently demonstrated that PU.1, C/EBP β RUNX1 and the co-factor CBF β also regulated miR-142 gene transcription. Specific binding sites for these three TFs exist in the miR-142 gene promoter with PU.1 acting predominantly. Indeed, C/EBP β and RUNX1 alone led to miR-142 deficiency. However, miR-142 expression levels within hematopoietic cells depend on different combinations of PU.1 together with C/EBP β RUNX1 [53]. By regulating PU.1 expression, miR-155, another relevant miRNA in hematopoiesis, is involved in miR-142 gene transcription. miR-155 targets PU.1 mRNA triggering its repression and consequently the reduction of miR-142 production. A good illustration of the interrelationship between miR-155, PU.1 and miR-142 is the pathway leading to IL-6 expression. Lipopolysaccharide (LPS) stimulates Toll-like receptor 4 (TLR4), which induces miR-155 expression, and down-regulation of PU.1, subsequently leading to decreased miR-142 expression and thus inducing expression of its target, IL-6. This regulatory mechanism is relevant for IL-6-mediated immunological processes and also the functions of miR-142 in the lymphoid system [53].

2.1.2.3. RUNX1 Connection with miRs-222/221, 17-5p, 20a, 106a and 27a

RUNX1 is implicated in the control of miR-222/221 gene cluster transcription as demonstrated by ChiP and luciferase reporter gene assays in U937 monocytic leukemia cells. Indeed, RUNX1 binds to two regions containing RUNX1-consensus sequences in the miR-222/221 gene promoter displaying four consensus sequences. Moreover, RUNX1 dose-dependently activates miR-222/221 promoter in the reporter gene constructs [54]. Expression of miR-222/221 was increased during GM-CSF-mediated myeloid differentiation of normal bone marrow CD133⁺ stem progenitor cells. These results correlate with the down-regulation of the stem cell factor (SCF) receptor KIT expression whose mRNA is targeted by miR-222/221. Conversely, the expression of miR-222/221 and miR-223 was lower in leukemia cells expressing RUNX1 fusion oncoproteins, in correlation with higher levels of KIT oncogene expression and inhibition of myeloid differentiation [54].

On the other side, RUNX1 binds to the promoters of miRNA 17-5p-92 and 106a-92 gene clusters leading to the inhibition of miRNA 17-5p, 20a and 106a transcription. These miRNAs are upregulated in undifferentiated CD34⁺ hematopoietic stem/progenitor cells (HSPCs) while their expression decreases during monocytic differentiation and maturation. This down-regulation is correlated to the increasingly expressed RUNX1 protein in monocytic cells. In turn, miRNAs 17-5p, 20a and 106a

inhibit RUNX1 translation by interacting with the 3' UTR generating a regulatory loop. RUNX1 was also shown to act as a regulator of miR-27 transcription, which is involved in erythro-megakaryocytic and granulocytic differentiation pathways. Generating a negative regulatory loop, miR-27 targets RUNX1 mRNA. This leads to RUNX1 down-regulation allowing granulocytic differentiation of myeloblasts [55]. Similarly, TPA (12-o-tetradecanoylphorbol-13-acetate)-induced megakaryocytic differentiation of K562 cells was concomitant with increased miR-27a expression and RUNX1 down-regulation [56]. RUNX1 appears thus as a crucial transcription factor in hematopoiesis, given its expression in the different cell lineages and its ability to directly regulate a wide panel of miRNAs as well as TFs including C/EBP and PU.1, which in turn target promoters of miRNA genes [57].

2.1.2.4. MiR-146a and miR-155 Genes Are Regulated by PU.1

During differentiation progress, TFs integrate the extensive regulatory network insuring hematopoietic homeostasis. GATA-1 and GATA-2 are the major TFs regulating erythro-megakaryocytic pathways whereas PU.1 and C/EBP control myelo-lymphoid differentiation. Nevertheless, GATA-1 and PU.1 proteins physically interact and inhibit each other. PU.1, a member of the Ets(E-twenty six) TF family, is involved in differentiation from HSCs to multipotent progenitors at different levels [58]. PU.1 is required for myeloid and lymphoid [59] differentiation and plays a role in cell fate decision. Its expression levels determine the fate of early T-cell progenitors since over-expression of PU.1 was shown to reorient differentiation to the myeloid lineage [60–63] depending on Notch. In the absence of Notch signaling, PU.1 promotes the myeloid pathway, whereas activation of Notch signaling is observed in T-cell lineage pathways. Del Real *et al.*, demonstrated that in absence of Notch signaling, PU.1 regulates expression of TF genes essential for T-cells, including Myb, Tcf7 and Gata3.

PU.1 also plays a pivotal role in lympho-myeloid development through its ability to regulate transcription of miRNA genes as reported by Ghani *et al.* These authors highlighted transcriptional regulation of miR-146a, miR-342, miR-338 and miR-155 genes by PU.1. Results revealed that miR-146a was the most robustly PU.1-induced miRNA. miR-146a and miR-155 up-regulation was independent of *de novo* protein synthesis suggesting that PU.1 is able to regulate transcription of these miRNA genes. Conversely, miR-342 and miR-338 regulation could require cooperative factors linked to PU.1 [64]. Authors also showed that miR-146a directed selective differentiation of HSCs into peritoneal macrophages during adult hematopoiesis and concluded that PU.1 temporally controls the expression of miRNAs required for correct HSCs differentiation.

PU.1 up-regulation was also shown to control human monocyte-macrophage differentiation through the activation of miR-424 [65]. The authors validated NFI-A mRNA as a true functional target of miR-424. Similarly to the C/EBP α -miR223-NFI-A regulatory circuit that regulates granulopoiesis, PU.1-miR-424-NFI-A regulates monocytic differentiation.

2.1.2.5. Erythroid Specific Expression of miRs-144 and 451

GATA factors play a central role in the network of hematopoiesis regulation through interactions with co-factors. Predominant expression of GATA-1 results in the erythroid differentiation. Moreover, PU.1 activity is also involved in the lineage-specific fate decisions of erythroid or myeloid differentiation through its physical interaction with GATA-1 leading to reciprocal inhibition of their

transcriptional activities. Especially, suppression of GATA1 activity by PU.1 leads to a shift in cell fate towards the myeloid-lymphoid lineages [66,67]. Upon myeloid differentiation, PU.1 is overexpressed in correlation with GATA-1 and -2 down-regulation. Using the G1ME hematopoietic cell line derived from *in vitro* differentiation of murine GATA-1⁻ embryonic stem cells (ESC) [68], Chou *et al.*, demonstrated that GATA factors act sequentially to control lineage determination during hematopoiesis, through modulation of repressive effects at key regulatory elements of the PU.1 gene [69]. So far, no link between PU.1 expression and the production of miRNAs regulated by GATA-1 has been demonstrated. GATA-1 was shown to regulate *miR-144/451* gene transcription. GATA-1 up-regulation induced *miR-144* and *miR-451* in G1E cells as well as in human CD34⁺ cells and murine erythroleukemia (MEL) cells in correlation with erythroid maturation [70]. The crucial role of miRs-451/144 in physiological erythropoiesis was also demonstrated in zebrafish and mouse models [71,72]. Expression of *miR-451* is restricted to the erythroid lineage and its induced silencing affected neither megakaryopoiesis nor granulopoiesis in zebrafish embryos. GATA-2, which is able to bind the *miR-144/451* locus in the absence of GATA-1, does not activate transcription. A couple of predicted miR-144 and miR-451 target genes, that were down-regulated after GATA-1 activation, were reported [70]. Interestingly, proliferation regulator c-myc, whose overexpression inhibits erythropoiesis, emerged as a predicted target gene of miR-451. Furthermore, it was shown in zebrafish that GATA-2 3'UTR was targeted by miR-451 but not by miR-144 [72]. *GATA-2* is known to preserve the immaturity of hematopoietic precursor cells [73] and induces overexpression of GATA-1, which in turn triggers GATA-2 down-regulation as a negative regulatory feedback loop, required for normal erythropoiesis. Considering that GATA-2 positively regulates PU.1 gene expression, it seems plausible that miR-451 could play a role in the GATA-1/GATA-2 balance and PU.1 down-regulation to warrant erythroid maturation according to a GATA-1 > miR451 > GATA-2 > PU.1 axis. Besides, this would be in agreement with decreased miR-155 expression observed during erythropoiesis since PU.1 regulates gene expression of this miRNA [31,74]. Mir-144 was reported to selectively regulate embryonic α -globin gene expression during primitive erythropoiesis in zebrafish through negative feedback regulation involving erythroid-specific Krüppel-like (KLF)-D TF, which selectively binds to promoters of both α -globin and miR-144 genes to activate their transcriptions [38,75].

Except for miRs-144/451 that are erythroid-specific miRNAs, most miRNAs are ubiquitously expressed at different stages of hematopoietic differentiation in correlation with TF expression and activities involved in the regulation of lineage specific genes.

2.2. LncRNA and Hematopoietic Lineage

2.2.1. LncRNA

Among all sub-categories of ncRNAs, long non-coding RNAs (lncRNAs) constitute the most recent and least characterized family. LncRNAs include all ncRNA larger than 200 nucleotides and not yet categorized in one of the other RNA families. In contrast to small ncRNAs, which are highly conserved among species and which are involved in transcriptional and posttranscriptional gene silencing, lncRNAs are poorly conserved [76] and their modes of regulation are diverse and not yet totally elucidated. The process of lncRNA transcription and maturation is similar to that of mRNA

even if their genes are not subject to the same histone modifications (H3K4me3 and H3K36me). As for mRNA, lncRNAs are transcribed by RNA polymerase II. Processing of lncRNA involves 3' poly(A) tailing and 5'-end capping as well as splicing. LncRNAs have small open reading frames without any protein-coding potential, but can be sometimes be associated with ribosomes in the cytoplasm, suggesting an additional role in mRNA metabolism.

Analysis of the human genome and transcriptome estimated that there are about 23,000 lncRNAs, comparable to the number of protein-coding RNAs and greatly exceeding the number of miRNAs (close to 2000) [77,78]. LncRNAs sequences are spread over the entire genome and can be found on all chromosomes [79]. During the last decade, the number of studies on this type of RNA has been increasing, but the general knowledge remains low given the large number of existing lncRNAs. However, it has already been shown that these lncRNAs are tissue specific and differently expressed under both normal and pathological conditions, implying that they may play important biological roles. Indeed they can regulate biological processes, including cell division, survival, and differentiation, as well as some processes related to cancer development. The first schemes of control or mechanism of action were also recently revealed. Several long ncRNA regulate gene expression by modifying chromatin structure. Different studies demonstrate a wide diversity of mechanisms by which a lncRNA regulate chromatin of a single promoter, a gene cluster, or an entire chromosome [6], in order to activate or silence genes in *cis* or in *trans* [80]. Until now, only a small number of lncRNAs were identified and fully characterized.

LncRNAs play different roles in transcription [81]. For example MALAT1 can regulate preRNA splicing by influencing distribution and phosphorylation of the serine/arginine (SR) splicing factor [82]. They can also increase mRNA stability by interaction with exonuclease XRN1, as shown in a recent study of the Moon [83]. Perfect hybridization between mRNA and lncRNA is another system to stabilize RNA: BACE1-AS fully hybridized to exon 6 of BACE1, which presents a binding site for miRNA-485-5p, thus leading to a competition between lncRNA and miRNA for this site [84]. On the opposite, it has also been shown that these ncRNAs can decrease mRNA stability by promoting degradation of mRNA [85]. LncRNA involvement in the regulation of translation was initially suggested by the observation of an association between lncRNA and ribosomes, even though lncRNA are not translated. This kind of regulation seems to be negative in some cases, for example when lncRNA-p21 annealed imperfectly but throughout the coding sequence of the mRNA of β -catenin or JunB, blocking their translation [86]. Positive regulation was also reported, for example in the case of lncRNA AS-Uchl1 that promoted polysome formation around Uchl1 mRNA [87]. In a less specific way, lncRNA BC1 was shown to inhibit the whole process of translation by preventing the assembly of translation initiation complex in neurons [88].

2.2.1.1. Expression and Regulation of lncRNA

As mentioned previously, lncRNA are tissue specific and differently expressed under normal and pathological conditions, implicating that their transcripts are regulated. While information about this regulation is constantly expanding, the way lncRNA abundance is controlled is still widely unknown. Intronic sequences represent 30% of the human genome and represent the major source for lncRNAs (intronic lncRNAs). In this case, regulation of non-coding RNA expression is usually linked to host

gene expression. LncRNA sequence can also be independent of genes and its expression is then under promoter dependence. This second subcategory is commonly called Long Intergenic Non-Coding RNA (lincRNA). In addition to their regulation at a transcriptional level, lincRNAs can also be regulated post-transcriptionally [81]. A study about the overall evaluation of ncRNA half-life (involving about 800 transcripts) in mouse neuronal cells showed that all lincRNA differed in stability with a minority of “unstable” lincRNA. Stability appears to depend on localization (intronic lincRNA have a lower stability compared to lincRNA), splicing (spliced non-coding transcripts are more stable than the non-spliced ones), and finally subcellular localization (nuclear lincRNA may have shorter half-lives compared to cytoplasmic ones) [89]. Most of the lincRNAs are stabilized with a poly(A) tail similar to mRNAs. Another type of stabilization involves a triple helix at the 3' end of lincRNAs protecting against 3'→5' exonucleolytic cleavage. This structure was recently shown for lincRNA-MALAT1 in multiple endocrine neoplasia β [90,91].

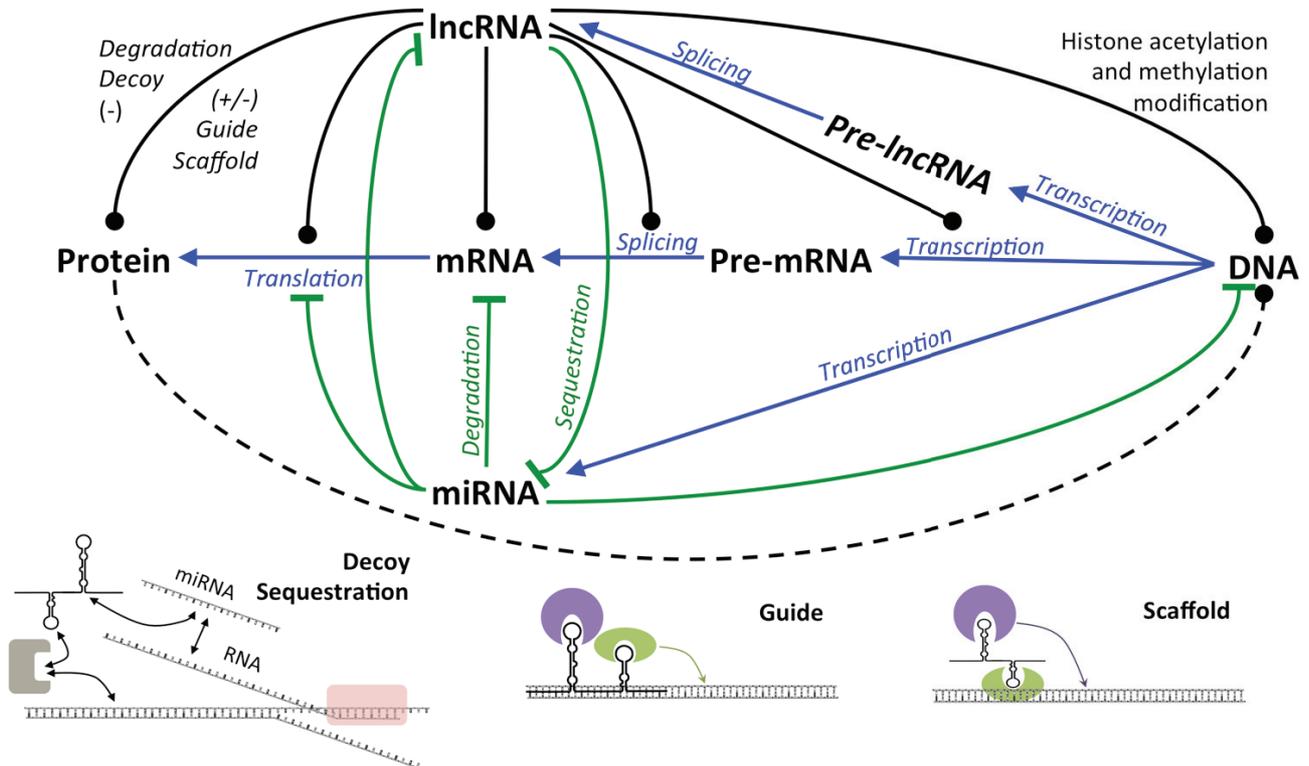
If lincRNAs present differential stability depending on their location or their origin, stability can also be modulated, thereby altering the turnover of these regulators. It has been recently shown that lincRNA-p21 could be destabilized by human antigen R (HuR)/Argonaute (Ago)-2 complex as well as by the miR Let-7b. Silencing of HuR or Ago2 increases lincRNA-p21 stability, whereas overexpression of let-7b induces destabilization. Degradation of RNA requires decapping and deadenylation, and these aspects of lincRNA degradation are still not well known and need to be further elucidated in order to improve the understanding of lincRNA abundance control [86].

2.2.1.2. Mechanisms

Unlike miRNAs, which follow a well-established mechanism of action, causing inhibition of translation or degradation of mRNA, lincRNAs act almost as diversely as proteins (Figure 4). LncRNAs can interact with proteins as well as with RNA or DNA. Interaction of lincRNA with other RNA occurs through total or partial hybridization of complementary sequences. This mechanism is common for anti-sense lincRNAs such as p15AS. It was recently published that interaction of lincRNA with miRNA happens with a preferential hybridization of the miRNA to the 3' end of the lincRNA [92]. RNA-RNA interaction can lead to translational regulation, but also to splicing or to inhibition of RNA function. LncRNA-DNA interaction can occur similarly to a RNA-RNA interaction, by the sequence complementarily forming a RNA-DNA duplex, or in a more complex configuration such as the hybridization of lincRNA and the dihydrofolate reductase (DHFR) promoter that forms a DNA-DNA-RNA triplex [93], leading to sequence specific transcriptional repression [94]. DNA-RNA interactions recruit DNA regulating factors that can affect transcription, histone or DNA modifications including methylation or acetylation. Finally, the ability of lincRNA to form secondary and tertiary structures allows them to establish complex interactions enabling them to catch and sequester proteins as well as to join distant areas.

In an effort to summarize the different modes of regulation involving lincRNAs, Wang and Chang propose four classes or archetypes of lincRNA regulation [95], further summarized in the recent publication of Da Sacco *et al.* [96]. LncRNAs usually follow several modes of action, and so most of them belong to several of these archetypes.

Figure 4. Scheme of regulations involving lncRNAs and miRNAs, influencing transcription, maturation or translation processes and example of interaction between lncRNAs, DNA, proteins and miRNAs. Blue arrows correspond to physiological maturation of RNA. Green lines correspond to inhibition processes and dotted black lines correspond to variable regulations (positive or negative depending on the mechanism). (–) Negative effect; (+/–) negative or positive effect.



Signal: Expression of these lncRNA can be used as a marker of intracellular signaling or response to stimuli, as it can induce responses *via* transcription of RNA without translation or posttranslational modifications. More a tool than a mechanism, these lncRNAs essentially serve as biomarkers. This archetype therefore concerns all lncRNA with a strong relationship in spatiotemporal response, such as rapid reaction to temperature changes regulated by lncRNAs COLDAIR and COOLAIR [97,98] or imprinting controlled by Xist.

Decoy: In molecular decoy, lncRNA compete with another nucleotide sequence or structure for the binding of a TF, chromatin modifier, or other regulatory factor such as miRNA thereby preventing the miRNA to inhibit translation of their targets [99,100]. These lncRNAs are therefore considered negative regulators of effectors.

Guide: The third archetype is the guide RNA, which can be described as a connection between lncRNAs and proteins, further guided by the lncRNA to a target. This pattern of regulation stems from the observation that lncRNAs are able to induce changes in gene expression in *cis* (neighboring) or *trans* (distant) genes. Target proteins recognized by the lncRNA may be repressor or activator complexes, or TFs, with the final goal of controlling the expression of a target gene, causing changes in the epigenome, whether the control takes place in *cis* or in *trans*.

Scaffold (protein linker): In the fourth archetype proposed by Wang *et al.*, the lncRNA becomes a docking structure for regulatory proteins, in order to assure accurate assembly of ribonucleoprotein complexes. In many biological signaling processes, this control feature is essential for specificity and dynamics of molecular interactions and signaling events. Until now, it was thought that specificity of such complexes was based essentially on intrinsic protein properties, but recently it was hypothesized in addition that lncRNA could also contribute. LncRNAs have specific RNA sequences that recognize individual protein effectors leading eventually to ribonucleoprotein complexes that combine the regulative properties of both RNA and protein components. The idea of protein assembly dependent on a nucleotidic scaffold provides novel insights into therapeutic strategies to artificially modify specificity or activity of these bipartite signaling complexes.

2.2.2. LncRNAs in Hematopoiesis

2.2.2.1. LncRNAs in Erythropoiesis

The knowledge about involvement of lncRNA in hematopoietic differentiation or self-renewal of hematopoietic stem cells remains poor, and a part of the observed current results are derived from murine cellular models. However, unlike miRNAs, lncRNAs are very poorly conserved between species so that results cannot readily be extrapolated to human. However, it may be possible that the overall operating mode is preserved and that it would be possible to transpose a process from one species to another, considering that functions of lncRNA are mainly mediated by their secondary and tertiary structure and not only through their sequence [101]. Currently, such a transposition has yet to be demonstrated, but the importance of the secondary structures is shown by the joint action of human Alu, and mouse B2 that both show no direct primary sequence similarity but are both able to sequester RNA polymerase II [93,102,103].

In murine Ter-199 cells, a direct link between differentiation of hematopoietic cells towards the erythroid pathway and the lncRNA EPS (erythroid prosurvival) regulates inhibition of apoptosis during differentiation [104,105] and terminal erythropoiesis. Interference with the expression of this lncRNA (even at 50%) induced a significant decrease in cell proliferation and increased annexin V expression. At least one of the targets of lncRNA-EPS could be Pycard, which is an antiapoptotic gene whose protein product is able to activate caspases during apoptosis. Many examples of Pycard regulation by lncRNA-EPS exist. First of all, expression of Pycard and EPS are inversely correlated during erythropoiesis. Moreover, overexpression of EPS leads essentially to repression of Pycard (among those studied by the team of Hu *et al.*). In addition, Pycard overexpression generates the same phenotype than inhibition of EPS during erythroid terminal differentiation: inhibition of proliferation, induction of cell death and inhibition of enucleation, the final step of differentiation. Finally, overexpression of Pycard abrogates anti-apoptotic effects induced by ectopic expression of EPS.

We mentioned earlier possible interactions between miRNAs and lncRNAs, so lncRNAs could modulate miRNAs *via* their transcription or their sequestration, and conversely, a miRNA can regulate lncRNA indirectly by acting at least on translation of the TF involved in the expression of this lncRNA. A recent study focused on the interaction between lncRNA and miRNA, using Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP) and

targeting Ago (Argonaute) proteins [92], the catalytic components of the RISC complex. Thus, the miRNA bound to Ago, guides it to its complementary sequence, allowing to identify the targets of miRNA and then to hypothesize the lncRNA involved in the action of the miRNA. This study mentioned four miRNAs targeting lncRNAs related to hematopoietic differentiation: miR-196A, miR-196B, miR-9-1 and miR-210. MiR196A and miR196B target HOXA5 [106] implicated in erythropoiesis and myelopoiesis [107,108]. MiR-9-1 and miR-210 [75] target ALK4, which can be related to erythropoiesis [109]. MiR-196A and B seem to interact with three ncRNAs referenced as ENST00000523790.1, ENST00000519935.1, ENST00000489695.1; miR-210 interacts with ENST00000525865.1 and miR-9-1 with ENST00000511014.1, ENST00000509783.1, ENST00000505030.1, ENST00000504246.1, and ENST00000500197.2. LncRNAs interact with miRNAs which themselves modulate genes involved in hematopoietic differentiation. This does not automatically mean that these lncRNA are involved in differentiation but could be targets for future investigations.

A better understanding of the interactions between miRNA and lncRNA will allow deducing the effects of an ncRNA when the effect of its interacting molecule is known. To simplify these analyses, the miRcode database was implemented [110], using GENCODE database and providing a comprehensive map of putative miRNA target sites across the GENCODE long non-coding transcriptome.

2.2.2.2. LncRNAs in Hematopoietic Differentiation

H19 was one of the first imprinted noncoding RNAs to be identified [111,112]. It is transcribed from chromosome 11, contains the sequence of miR-675 in its first exon, and is one of the few lncRNAs conserved between species. H19 is co-regulated with insulin-like growth factor (IGF)-2, which is expressed from the same locus. In hematopoietic cells, this pair of growth regulatory genes is expressed in precursor cells and downregulated during normal and pathological differentiation [113]. However, expression of this ncRNA is not restricted to hematopoietic tissue, as it seems also expressed in ES cells, fetal and adult tissues [114]. Recent results show that H19 RNA could play a role in *trans* repression of *Igf2* expression [115]. Abundant data about the implication of this lncRNA in cancer development show that H19 is essential for growth of some human tumor types, and it appears that its aberrant expression in cancer cells is due to modification of the methylation of imprinting control regions (ICR) of H19/IGF2 during differentiation. Recent data on bladder cancer documented how upregulated H19 increased cancer cell proliferation by increasing inhibitor of DNA binding 2 (ID2) expression [116].

In 2007, a study mentioned a new essential lncRNA implicated in hematopoietic differentiation [117]. Authors showed in several steps that a transcript is highly over-expressed during eosinophilic differentiation, and that the intronic region that encodes two transcript variants is highly conserved between human, mouse and chicken, which could reflect an evolutionary pressure leading to conservation of this potentially essential sequence. Moreover, the lack of a large ORF and paradoxically a poor amino acid conservation of small ORF strongly suggests that this transcript is an ncRNA. Results were confirmed by the lack of association of this transcript with ribosomes and absence of translation. An association with other types of proteins was proved nevertheless. This ncRNA was termed eosinophil Granule Ontogeny (EGO) and exists in two isoforms (EGO-A and EGO-B) of 535 and 1460 bases, respectively. Tissue-specific expression patterns suggest that EGOS

act on bone marrow hematopoietic cell development but not on lymphoid development. Its silencing proved its requirement for major basic protein (MBP) and eosinophil derived neurotoxin (EDN) mRNA expression, but not for GATA-1. Its precise mechanism of action thus remains to be elucidated.

It has been shown recently in mice that HOXA6 and HOXA7 genes are indirectly regulated by lncRNA Mistral (MIRA) *via* recruitment of Histone-lysine *N*-methyltransferase (MLL)1 to chromatin, allowing transcription of HOXA6 and 7 genes. Results obtained in the murine system demonstrate involvement of ncRNAs in regulation of genes involved in mouse embryonic stem cell (mESC) differentiation [118]. These results can be extended beyond regulation of mESC, as HOXA6 plays an important role in the regulation of HSC self-renewal in human, and its overexpression is also involved in leukemia [119]; HOXA7 acts as an intermediate in the regulation of granulocytic differentiation repressed by the Polycomb group RING finger protein 2 (PCGF2) [120].

The transcription factor TAL1 (SCF) is a major regulator of hematopoietic differentiation [121]. The locus containing this gene also encodes two sense and antisense lncRNAs. Transcriptional inhibition of a lincRNA can lead to decreased, transcriptionally independent, expression of neighboring protein-coding genes in multiple human loci. Orom *et al.* showed that the lincRNA ENST00000444042.2 (CYP4A22-AS1-001) encoded by a sequence downstream of TAL1, on the same DNA strand, is a positive regulator of this gene's expression [76]. The study shows that these effects are independent of lncRNA orientation towards their target sequences. Depletion of this lincRNA induces a specific, strong and significant decrease of TAL1 expression, but does not affect expression of other genes on the same locus. Depletion of ENST00000429328, a lincRNA located on the same locus, affects the expression of another closely located gene, deoxycytidylate kinase (CMPK1) but not the expression of TAL1. Possible mechanisms of action could be: interaction through sequence or structural homology with the encoding target gene, through the recruitment of transcriptional activator or basal TFs, or by the removal of a repressor. Finally, chromatin remodeling could also be involved.

Zhang *et al.* identified an intergenic transcriptional activity, located between the human HOXA1 and HOXA2 genes, which presents a myeloid-specific expression with specific up-regulation during granulocytic differentiation [122]. This lncRNA is called HOX antisense intergenic RNA myeloid 1 (HOTAIRM1). Its induction during RA-induced granulocytic differentiation acts through the RA receptor and depends on the expression of myeloid cell development factors targeted by RA signaling. Extinction of this gene attenuated the expression of HOXA1 and HOXA2 genes induced by RA, without affecting expression of more distal HOXA genes. This knock out also affects transcription of CD11b and CD18 involved in myeloid differentiation. Finally, authors suggest that HOTAIRM1 plays a role in myelopoiesis through modulation of HOXA cluster gene expression.

The lncRNA HOXA cluster antisense RNA 2 (HOXA-AS2) is transcribed from a gene located between genes HOXA3 and HOXA4, and is expressed in human peripheral blood neutrophils. This ncRNA plays an anti-apoptotic role in All trans RA (ATRA)-induced myeloid differentiation, protecting cells against ATRA-induced apoptosis probably through the inhibition of the TNF α -related apoptosis-inducing ligand (TRAIL) pathway [123]: in the promyelocytic leukemia cell model NB4, knockdown of HOXA-AS2 increases number of both apoptotic cells and TRAIL. Conversely, ATRA-induced NB4 cells treated with TRAIL show an increase in HOXA-AS2 expression.

The lncRNA Xist triggers X chromosome inactivation in female mammals and is expressed at an early embryonic stage. Xist is important for the development of mice embryos particularly between 3.5 and 12.5 days post-coitum. Moreover, the adults that already committed hematopoietic progenitors, rather than HSC, were also able to express this lncRNA [124]. Overexpression of Xist led to the loss of a majority of blood cell types besides pre-B and pre-T lymphocytes, which are the cell types able to reactivate Xist expression. It is interesting to note that this reactivation occurs in pre-B cells and pre-T lymphoid cells that undergo allelic exclusion of immunoglobulin heavy chain (*Igh*) or TCR β loci. This observation led to the hypothesis of a link between allelic exclusion of antigen receptor genes in lymphocytes and X chromosome inactivation in embryonic cells.

A few years ago, a meta-analysis was conducted by Gibb *et al.* on 272 human serial analyses of gene expression (SAGE) involving lncRNA sequences. These analyses included 26 different normal and 19 cancer tissue types in order to establish a first global profiling of lncRNA [79]. Results show tissue-specific lncRNA expression in normal tissues and a systematic abnormal lncRNA expression in human cancer. The paper essentially discusses the difference between normal and cancerous tissues (breast, brain, lung and blood) and discusses lncRNA expression that could become interesting for future studies. Four lncRNA have an expression pattern specific to white blood cells: ENSG00000232192 (Transcribed ENST00000446321, 810 bases) also named RP11-62I21.1 is an lncRNA transcribed by the antisense strand of the protein-coding gene KIF26B-005, coding for the kinesin family member 26B which is not implicated in any hematopoietic differentiation. The second lncRNA gene, ENSG00000246100 (CTC-774J1.2), is located in an intergenic region and is transcribed to five lncRNA variants of 3322, 982, 504, 693 and 995 bases. ENSG00000227712 (RP11-418J17.3), transcribes into an lncRNA of 898 bases. The most interesting one, ENSG00000256910 (AL034397.1), corresponds to a sequence close to the gene encoding the protein Z39Ig associated with monocyte and macrophage cells and linked to inflammatory reaction [125,126]. The AL034397.1 gene produces two transcripts by alternative splicing containing respectively, two and three exons: ENST00000540516 and ENST00000538676. For both, the last exon contains the sequence of miRNA-223, cited previously for its essential implication in hematopoietic differentiation. No links have yet been established between these two non-coding RNAs, but it is tempting to speculate about an interaction between these two non-coding RNA.

Altogether there are less than a dozen lncRNAs with a confirmed involvement in hematopoietic differentiation. For some others, involvement may be assumed. It is not yet possible to define a network of global interactions between lncRNAs, but we strongly believe that such a regulatory network involving TFs, miRNA and lncRNA exists and will be defined in the near future.

3. Conclusions

Long term studies of TFs involved in the regulation of genes driving hematopoietic cell fate decision, proliferation, survival, differentiation and death, designed a complex regulation network from HSC to differentiated blood cells. During the last decade, an additional level of regulation was added to this TF network with the discovery and understanding of a new class of regulatory non-coding RNAs, the miRNAs. Understanding of functional and reciprocal interactions with TFs significantly improved our knowledge about the molecular biology of hematopoiesis and

hematological diseases. Nowadays, the complexity of the regulatory network is even increasing through the emergence of lncRNAs as novel relevant regulatory elements in all biological processes. Even though the roles of lncRNAs remain partially undetermined, their ability to interact with miRNAs, regulatory proteins and DNA has been evidenced, including in the hematopoietic system. Together, TFs, miRNAs and lncRNAs most likely constitute a wide and complex regulatory network contributing to physiological hematopoietic development as well as to pathological alterations.

Acknowledgments

Sébastien Chateauvieux and Anthoula Gaigneaux were supported by fellowships from Télévie-Luxembourg. Research at LBMCC is financially supported by the Fondation de Recherche Cancer et Sang, the Recherches Scientifiques Luxembourg association, the Een Haerz fir kriibskrank Kanner association, the Action Lions Vaincre le Cancer association, the European Union (ITN “RedCat” 215009, Interreg Iva project “Corena”) and the Télévie Luxembourg. MD is supported by the National Research Foundation (NRF) by the MEST of Korea for Tumor Microenvironment Global Core Research Center (GCRC) grant, [grant number 2012-0001184]; by the Seoul National University Research grant and by the Research Settlement Fund for the new faculty of SNU.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Inaba, M.; Yamashita, Y.M. Asymmetric stem cell division: Precision for robustness. *Cell Stem Cell* **2012**, *11*, 461–469.
2. Lichtinger, M.; Ingram, R.; Hannah, R.; Muller, D.; Clarke, D.; Assi, S.A.; Lie, A.L.M.; Noailles, L.; Vijayabaskar, M.S.; Wu, M.; *et al.* RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis. *EMBO J.* **2012**, *31*, 4318–4333.
3. Papathanasiou, P.; Attema, J.L.; Karsunky, H.; Hosen, N.; Sontani, Y.; Hoyne, G.F.; Tunningley, R.; Smale, S.T.; Weissman, I.L. Self-renewal of the long-term reconstituting subset of hematopoietic stem cells is regulated by Ikaros. *Stem Cells* **2009**, *27*, 3082–3092.
4. Malinge, S.; Thiollier, C.; Chlon, T.M.; Dore, L.C.; Diebold, L.; Bluteau, O.; Mabilah, V.; Vainchenker, W.; Dessen, P.; Winandy, S.; *et al.* Ikaros inhibits megakaryopoiesis through functional interaction with GATA-1 and NOTCH signaling. *Blood* **2013**, *121*, 2440–2451.
5. Mattick, J.S. Non-coding RNAs: The architects of eukaryotic complexity. *EMBO Reports* **2001**, *2*, 986–991.
6. Mattick, J.S.; Makunin, I.V. Non-coding RNA. *Hum. Mol. Gen.* **2006**, *15*, R17–R29.
7. Pheasant, M.; Mattick, J.S. Raising the estimate of functional human sequences. *Genome Res.* **2007**, *17*, 1245–1253.
8. Fire, A.; Xu, S.; Montgomery, M.K.; Kostas, S.A.; Driver, S.E.; Mello, C.C. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **1998**, *391*, 806–811.

9. 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.
10. Hamilton, A.J.; Baulcombe, D.C. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **1999**, *286*, 950–952.
11. Elbashir, S.M.; Harborth, J.; Lendeckel, W.; Yalcin, A.; Weber, K.; Tuschl, T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **2001**, *411*, 494–498.
12. 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.
13. 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.
14. 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.
15. Lagos-Quintana, M.; Rauhut, R.; Lendeckel, W.; Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **2001**, *294*, 853–858.
16. Lee, R.C.; Ambros, V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **2001**, *294*, 862–864.
17. 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.
18. 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.
19. 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.
20. Wang, Z.; Yao, H.; Lin, S.; Zhu, X.; Shen, Z.; Lu, G.; Poon, W.S.; Xie, D.; Lin, M.C.; Kung, H.F. Transcriptional and epigenetic regulation of human microRNAs. *Cancer Lett.* **2013**, *331*, 1–10.
21. Kloosterman, W.P.; Plasterk, R.H. The diverse functions of microRNAs in animal development and disease. *Dev. Cell* **2006**, *11*, 441–450.
22. Yates, L.A.; Norbury, C.J.; Gilbert, R.J. The long and short of microRNA. *Cell* **2013**, *153*, 516–519.
23. Borchert, G.M.; Lanier, W.; Davidson, B.L. RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.* **2006**, *13*, 1097–1101.
24. Altuvia, Y.; Landgraf, P.; Lithwick, G.; Elefant, N.; Pfeffer, S.; Aravin, A.; Brownstein, M.J.; Tuschl, T.; Margalit, H. Clustering and conservation patterns of human microRNAs. *Nucleic Acids Res.* **2005**, *33*, 2697–2706.
25. 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.
26. Wu, L.; Fan, J.; Belasco, J.G. MicroRNAs direct rapid deadenylation of mRNA. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4034–4039.

27. Zardo, G.; Ciolfi, A.; Vian, L.; Starnes, L.M.; Billi, M.; Racanicchi, S.; Maresca, C.; Fazi, F.; Travaglini, L.; Noguera, N.; *et al.* Polycombs and microRNA-223 regulate human granulopoiesis by transcriptional control of target gene expression. *Blood* **2012**, *119*, 4034–4046.
28. Johnnidis, J.B.; Harris, M.H.; Wheeler, R.T.; Stehling-Sun, S.; Lam, M.H.; Kirak, O.; Brummelkamp, T.R.; Fleming, M.D.; Camargo, F.D. Regulation of progenitor cell proliferation and granulocyte function by microRNA-223. *Nature* **2008**, *451*, 1125–1129.
29. Fazi, F.; Rosa, A.; Fatica, A.; Gelmetti, V.; De Marchis, M.L.; Nervi, C.; Bozzoni, I. A minicircuitry comprised of microRNA-223 and transcription factors NFI-A and C/EBPalpha regulates human granulopoiesis. *Cell* **2005**, *123*, 819–831.
30. Fukao, T.; Fukuda, Y.; Kiga, K.; Sharif, J.; Hino, K.; Enomoto, Y.; Kawamura, A.; Nakamura, K.; Takeuchi, T.; Tanabe, M. An evolutionarily conserved mechanism for microRNA-223 expression revealed by microRNA gene profiling. *Cell* **2007**, *129*, 617–631.
31. Masaki, S.; Ohtsuka, R.; Abe, Y.; Muta, K.; Umemura, T. Expression patterns of microRNAs 155 and 451 during normal human erythropoiesis. *Biochem. Biophys. Res. Commun.* **2007**, *364*, 509–514.
32. 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.
33. 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.
34. Choong, M.L.; Yang, H.H.; McNiece, I. MicroRNA expression profiling during human cord blood-derived CD34 cell erythropoiesis. *Exp. Hematol.* **2007**, *35*, 551–564.
35. Felli, N.; Pedini, F.; Romania, P.; Biffoni, M.; Morsilli, O.; Castelli, G.; Santoro, S.; Chicarella, S.; Sorrentino, A.; Peschle, C.; *et al.* MicroRNA 223-dependent expression of LMO2 regulates normal erythropoiesis. *Haematologica* **2009**, *94*, 479–486.
36. Cline, M.S.; Smoot, M.; Cerami, E.; Kuchinsky, A.; Landys, N.; Workman, C.; Christmas, R.; Avila-Campilo, I.; Creech, M.; Gross, B.; *et al.* Integration of biological networks and gene expression data using Cytoscape. *Nat. Protoc.* **2007**, *2*, 2366–2382.
37. Zhou, K.; Yi, S.; Yu, Z.; Li, Z.; Wang, Y.; Zou, D.; Qi, J.; Zhao, Y.; Qiu, L. MicroRNA-223 expression is uniformly down-regulated in B cell lymphoproliferative disorders and is associated with poor survival in patients with chronic lymphocytic leukemia. *Leukemia Lymphoma* **2012**, *53*, 1155–1161.
38. Chen, M.J.; Yokomizo, T.; Zeigler, B.M.; Dzierzak, E.; Speck, N.A. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature* **2009**, *457*, 887–891.
39. Okuda, T.; Takeda, K.; Fujita, Y.; Nishimura, M.; Yagyu, S.; Yoshida, M.; Akira, S.; Downing, J.R.; Abe, T. Biological characteristics of the leukemia-associated transcriptional factor AML1 disclosed by hematopoietic rescue of AML1-deficient embryonic stem cells by using a knock-in strategy. *Mol. Cell. Biol.* **2000**, *20*, 319–328.
40. Wang, Q.; Stacy, T.; Binder, M.; Marin-Padilla, M.; Sharpe, A.H.; Speck, N.A. Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 3444–3449.

41. Yokomizo, T.; Ogawa, M.; Osato, M.; Kanno, T.; Yoshida, H.; Fujimoto, T.; Fraser, S.; Nishikawa, S.; Okada, H.; Satake, M.; *et al.* Requirement of Runx1/AML1/PEBP2alphaB for the generation of haematopoietic cells from endothelial cells. *Genes Cells* **2001**, *6*, 13–23.
42. Nishimura, M.; Fukushima-Nakase, Y.; Fujita, Y.; Nakao, M.; Toda, S.; Kitamura, N.; Abe, T.; Okuda, T. VWRPY motif-dependent and -independent roles of AML1/Runx1 transcription factor in murine hematopoietic development. *Blood* **2004**, *103*, 562–570.
43. Ran, D.; Shia, W.J.; Lo, M.C.; Fan, J.B.; Knorr, D.A.; Ferrell, P.I.; Ye, Z.; Yan, M.; Cheng, L.; Kaufman, D.S.; *et al.* RUNX1a enhances hematopoietic lineage commitment from human embryonic stem cells and inducible pluripotent stem cells. *Blood* **2013**, *121*, 2882–2890.
44. Fazi, F.; Racanicchi, S.; Zardo, G.; Starnes, L.M.; Mancini, M.; Travaglini, L.; Diverio, D.; Ammatuna, E.; Cimino, G.; Lo-Coco, F.; *et al.* Epigenetic silencing of the myelopoiesis regulator microRNA-223 by the AML1/ETO oncoprotein. *Cancer Cell* **2007**, *12*, 457–466.
45. Nimer, S.D.; Moore, M.A. Effects of the leukemia-associated AML1-ETO protein on hematopoietic stem and progenitor cells. *Oncogene* **2004**, *23*, 4249–4254.
46. Frank, R.; Zhang, J.; Uchida, H.; Meyers, S.; Hiebert, S.W.; Nimer, S.D. The AML1/ETO fusion protein blocks transactivation of the GM-CSF promoter by AML1B. *Oncogene* **1995**, *11*, 2667–2674.
47. Zhang, D.E.; Fujioka, K.; Hetherington, C.J.; Shapiro, L.H.; Chen, H.M.; Look, A.T.; Tenen, D.G. Identification of a region which directs the monocytic activity of the colony-stimulating factor 1 (macrophage colony-stimulating factor) receptor promoter and binds PEBP2/CBF (AML1). *Mol. Cell. Biol.* **1994**, *14*, 8085–8095.
48. Sun, W.; Shen, W.; Yang, S.; Hu, F.; Li, H.; Zhu, T.H. miR-223 and miR-142 attenuate hematopoietic cell proliferation, and miR-223 positively regulates miR-142 through LMO2 isoforms and CEBP-beta. *Cell Res.* **2010**, *20*, 1158–1169.
49. Warren, A.J.; Colledge, W.H.; Carlton, M.B.; Evans, M.J.; Smith, A.J.; Rabbitts, T.H. The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development. *Cell* **1994**, *78*, 45–57.
50. Wadman, I.A.; Osada, H.; Grutz, G.G.; Agulnick, A.D.; Westphal, H.; Forster, A.; Rabbitts, T.H. The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J.* **1997**, *16*, 3145–3157.
51. Starnes, L.M.; Sorrentino, A.; Ferracin, M.; Negrini, M.; Pelosi, E.; Nervi, C.; Peschle, C. A transcriptome-wide approach reveals the key contribution of NFI-A in promoting erythroid differentiation of human CD34(+) progenitors and CML cells. *Leukemia* **2010**, *24*, 1220–1223.
52. Starnes, L.M.; Sorrentino, A.; Pelosi, E.; Ballarino, M.; Morsilli, O.; Biffoni, M.; Santoro, S.; Felli, N.; Castelli, G.; de Marchis, M.L.; *et al.* NFI-A directs the fate of hematopoietic progenitors to the erythroid or granulocytic lineage and controls beta-globin and G-CSF receptor expression. *Blood* **2009**, *114*, 1753–1763.
53. Sun, Y.; Sun, J.; Tomomi, T.; Nieves, E.; Mathewson, N.; Tamaki, H.; Evers, R.; Reddy, P. PU.1-dependent transcriptional regulation of miR-142 contributes to its hematopoietic cell-specific expression and modulation of IL-6. *J. Immunol.* **2013**, *190*, 4005–4013.

54. Brioschi, M.; Fischer, J.; Cairoli, R.; Rossetti, S.; Pezzetti, L.; Nichelatti, M.; Turrini, M.; Corlazzoli, F.; Scarpati, B.; Morra, E.; *et al.* Down-regulation of microRNAs 222/221 in acute myelogenous leukemia with deranged core-binding factor subunits. *Neoplasia* **2010**, *12*, 866–876.
55. Feng, J.; Iwama, A.; Satake, M.; Kohu, K. MicroRNA-27 enhances differentiation of myeloblasts into granulocytes by post-transcriptionally downregulating Runx1. *Br. J. Haematol.* **2009**, *145*, 412–423.
56. Ben-Ami, O.; Pencovich, N.; Lotem, J.; Levanon, D.; Groner, Y. A regulatory interplay between miR-27a and Runx1 during megakaryopoiesis. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 238–243.
57. Rossetti, S.; Sacchi, N. RUNX1: A microRNA hub in normal and malignant hematopoiesis. *Int. J. Mol. Sci.* **2013**, *14*, 1566–1588.
58. Wontakal, S.N.; Guo, X.; Will, B.; Shi, M.; Raha, D.; Mahajan, M.C.; Weissman, S.; Snyder, M.; Steidl, U.; Zheng, D.; *et al.* A large gene network in immature erythroid cells is controlled by the myeloid and B cell transcriptional regulator PU.1. *PLoS Genet.* **2011**, *7*, e1001392.
59. Houston, I.B.; Kamath, M.B.; Schweitzer, B.L.; Chlon, T.M.; DeKoter, R.P. Reduction in PU.1 activity results in a block to B-cell development, abnormal myeloid proliferation, and neonatal lethality. *Exp. Hematol.* **2007**, *35*, 1056–1068.
60. Anderson, M.K.; Weiss, A.H.; Hernandez-Hoyos, G.; Dionne, C.J.; Rothenberg, E.V. Constitutive expression of PU.1 in fetal hematopoietic progenitors blocks T cell development at the pro-T cell stage. *Immunity* **2002**, *16*, 285–296.
61. Dionne, C.J.; Tse, K.Y.; Weiss, A.H.; Franco, C.B.; Wiest, D.L.; Anderson, M.K.; Rothenberg, E.V. Subversion of T lineage commitment by PU.1 in a clonal cell line system. *Dev. Biol.* **2005**, *280*, 448–466.
62. Laiosa, C.V.; Stadtfeld, M.; Xie, H.; de Andres-Aguayo, L.; Graf, T. Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBP alpha and PU.1 transcription factors. *Immunity* **2006**, *25*, 731–744.
63. Lefebvre, J.M.; Haks, M.C.; Carleton, M.O.; Rhodes, M.; Sinnathamby, G.; Simon, M.C.; Eisenlohr, L.C.; Garrett-Sinha, L.A.; Wiest, D.L. Enforced expression of Spi-B reverses T lineage commitment and blocks beta-selection. *J. Immunol.* **2005**, *174*, 6184–6194.
64. Ghani, S.; Riemke, P.; Schonheit, J.; Lenze, D.; Stumm, J.; Hoogenkamp, M.; Lagendijk, A.; Heinz, S.; Bonifer, C.; Bakkers, J.; *et al.* Macrophage development from HSCs requires PU.1-coordinated microRNA expression. *Blood* **2011**, *118*, 2275–2284.
65. Rosa, A.; Ballarino, M.; Sorrentino, A.; Sthandier, O.; de Angelis, F.G.; Marchioni, M.; Masella, B.; Guarini, A.; Fatica, A.; Peschle, C.; *et al.* The interplay between the master transcription factor PU.1 and miR-424 regulates human monocyte/macrophage differentiation. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19849–19854.
66. Rhodes, J.; Hagen, A.; Hsu, K.; Deng, M.; Liu, T.X.; Look, A.T.; Kanki, J.P. Interplay of pu.1 and gata1 determines myelo-erythroid progenitor cell fate in zebrafish. *Dev. Cell* **2005**, *8*, 97–108.
67. Zhang, P.; Zhang, X.; Iwama, A.; Yu, C.; Smith, K.A.; Mueller, B.U.; Narravula, S.; Torbett, B.E.; Orkin, S.H.; Tenen, D.G. PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding. *Blood* **2000**, *96*, 2641–2648.
68. Stachura, D.L.; Chou, S.T.; Weiss, M.J. Early block to erythromegakaryocytic development conferred by loss of transcription factor GATA-1. *Blood* **2006**, *107*, 87–97.

69. Chou, S.T.; Khandros, E.; Bailey, L.C.; Nichols, K.E.; Vakoc, C.R.; Yao, Y.; Huang, Z.; Crispino, J.D.; Hardison, R.C.; Blobel, G.A.; *et al.* Graded repression of PU.1/Sfp1 gene transcription by GATA factors regulates hematopoietic cell fate. *Blood* **2009**, *114*, 983–994.
70. Dore, L.C.; Amigo, J.D.; Dos Santos, C.O.; Zhang, Z.; Gai, X.; Tobias, J.W.; Yu, D.; Klein, A.M.; Dorman, C.; Wu, W.; *et al.* A GATA-1-regulated microRNA locus essential for erythropoiesis. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3333–3338.
71. Papapetrou, E.P.; Korkola, J.E.; Sadelain, M. A genetic strategy for single and combinatorial analysis of miRNA function in mammalian hematopoietic stem cells. *Stem Cells* **2010**, *28*, 287–296.
72. Pase, L.; Layton, J.E.; Kloosterman, W.P.; Carradice, D.; Waterhouse, P.M.; Lieschke, G.J. miR-451 regulates zebrafish erythroid maturation *in vivo* via its target gata2. *Blood* **2009**, *113*, 1794–1804.
73. Minegishi, N.; Suzuki, N.; Yokomizo, T.; Pan, X.; Fujimoto, T.; Takahashi, S.; Hara, T.; Miyajima, A.; Nishikawa, S.; Yamamoto, M. Expression and domain-specific function of GATA-2 during differentiation of the hematopoietic precursor cells in midgestation mouse embryos. *Blood* **2003**, *102*, 896–905.
74. Bruchova, H.; Yoon, D.; Agarwal, A.M.; Mendell, J.; Prchal, J.T. Regulated expression of microRNAs in normal and polycythemia vera erythropoiesis. *Exp. Hematol.* **2007**, *35*, 1657–1667.
75. Mizuno, Y.; Tokuzawa, Y.; Ninomiya, Y.; Yagi, K.; Yatsuka-Kanesaki, Y.; Suda, T.; Fukuda, T.; Katagiri, T.; Kondoh, Y.; Amemiya, T.; *et al.* miR-210 promotes osteoblastic differentiation through inhibition of AcvR1b. *FEBS Lett.* **2009**, *583*, 2263–2268.
76. 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.
77. 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.
78. Carninci, P.; Hayashizaki, Y. Noncoding RNA transcription beyond annotated genes. *Curr. Opin. Genetics Dev.* **2007**, *17*, 139–144.
79. 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.
80. Whitehead, J.; Pandey, G.K.; Kanduri, C. Regulation of the mammalian epigenome by long noncoding RNAs. *Biochim. Biophys. Acta* **2009**, *1790*, 936–947.
81. Yoon, J.H.; Abdelmohsen, K.; Gorospe, M. Posttranscriptional gene regulation by long noncoding RNA. *J. Mol. Biol.* in press, 2012.
82. 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.
83. Moon, S.L.; Anderson, J.R.; Kumagai, Y.; Wilusz, C.J.; Akira, S.; Khromykh, A.A.; Wilusz, J. A noncoding RNA produced by arthropod-borne flaviviruses inhibits the cellular exoribonuclease XRN1 and alters host mRNA stability. *RNA* **2012**, *18*, 2029–2040.

84. 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.
85. Kim, Y.K.; Furic, L.; Parisien, M.; Major, F.; DesGroseillers, L.; Maquat, L.E. Staufen1 regulates diverse classes of mammalian transcripts. *EMBO J.* **2007**, *26*, 2670–2681.
86. 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.
87. 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.
88. Kindler, S.; Wang, H.; Richter, D.; Tiedge, H. RNA transport and local control of translation. *Annu. Rev. Cell Dev. Biol.* **2005**, *21*, 223–245.
89. 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.
90. 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.
91. Wilusz, J.E.; Inbaptiste, 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.
92. Jalali, S.; Bhartiya, D.; Lalwani, M.K.; Sivasubbu, S.; Scaria, V. Systematic transcriptome wide analysis of lncRNA-miRNA interactions. *PLoS One* **2013**, *8*, e53823.
93. Hung, T.; Chang, H.Y. Long noncoding RNA in genome regulation: Prospects and mechanisms. *RNA Biol.* **2010**, *7*, 582–585.
94. 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.
95. Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914.
96. Da Sacco, L.; Baldassarre, A.; Masotti, A. Bioinformatics tools and novel challenges in long non-coding RNAs (lncRNAs) functional analysis. *Int. J. Mol. Sci.* **2012**, *13*, 97–114.
97. Heo, J.B.; Sung, S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* **2011**, *331*, 76–79.
98. 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.
99. 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.

100. Poliseno, L.; Salmena, L.; Riccardi, L.; Fornari, A.; Song, M.S.; Hobbs, R.M.; Sportoletti, P.; Varmeh, S.; Egia, A.; Fedele, G.; *et al.* Identification of the miR-106b~25 microRNA cluster as a proto-oncogenic PTEN-targeting intron that cooperates with its host gene MCM7 in transformation. *Sci. Signal.* **2010**, *3*, ra29.
101. Torarinsson, E.; Sawera, M.; Havgaard, J.H.; Fredholm, M.; Gorodkin, J. Thousands of corresponding human and mouse genomic regions unalignable in primary sequence contain common RNA structure. *Genome Res.* **2006**, *16*, 885–889.
102. Espinoza, C.A.; Allen, T.A.; Hieb, A.R.; Kugel, J.F.; Goodrich, J.A. B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. *Nat. Struct. Mol. Biol.* **2004**, *11*, 822–829.
103. 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.
104. Hu, J.; Xu, T.; Zhu, T.; Lou, Q.; Wang, X.; Wu, Y.; Huang, R.; Liu, J.; Liu, H.; Yu, F.; *et al.* Monoclonal antibodies against accumulation-associated protein affect EPS biosynthesis and enhance bacterial accumulation of *Staphylococcus epidermidis*. *PLoS One* **2011**, *6*, e20918.
105. Paralkar, V.R.; Weiss, M.J. A new ‘Linc’ between noncoding RNAs and blood development. *Genes Dev.* **2011**, *25*, 2555–2558.
106. McGlenn, E.; Yekta, S.; Mansfield, J.H.; Soutschek, J.; Bartel, D.P.; Tabin, C.J. *In ovo* application of antagomiRs indicates a role for miR-196 in patterning the chick axial skeleton through Hox gene regulation. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18610–18615.
107. Crooks, G.M.; Fuller, J.; Petersen, D.; Izadi, P.; Malik, P.; Pattengale, P.K.; Kohn, D.B.; Gasson, J.C. Constitutive HOXA5 expression inhibits erythropoiesis and increases myelopoiesis from human hematopoietic progenitors. *Blood* **1999**, *94*, 519–528.
108. Moore, M.A.; Dorn, D.C.; Schuringa, J.J.; Chung, K.Y.; Morrone, G. Constitutive activation of Flt3 and STAT5A enhances self-renewal and alters differentiation of hematopoietic stem cells. *Exp. Hematol.* **2007**, *35*, 105–116.
109. Wang, Q.; Huang, Z.; Xue, H.; Jin, C.; Ju, X.L.; Han, J.D.; Chen, Y.G. MicroRNA miR-24 inhibits erythropoiesis by targeting activin type I receptor ALK4. *Blood* **2008**, *111*, 588–595.
110. Jeggari, A.; Marks, D.S.; Larsson, E. miRcode: A map of putative microRNA target sites in the long non-coding transcriptome. *Bioinformatics* **2012**, *28*, 2062–2063.
111. Goshen, R.; Rachmilewitz, J.; Schneider, T.; de-Groot, N.; Ariel, I.; Palti, Z.; Hochberg, A.A. The expression of the H-19 and IGF-2 genes during human embryogenesis and placental development. *Mol. Reproduct. Dev.* **1993**, *34*, 374–379.
112. 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.
113. Tessema, M.; Langer, F.; Bock, O.; Seltsam, A.; Metzsig, 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.

114. Lustig, O.; Ariel, I.; Ilan, J.; Lev-Lehman, E.; De-Groot, N.; Hochberg, A. Expression of the imprinted gene H19 in the human fetus. *Mol. Reproduct. Dev.* **1994**, *38*, 239–246.
115. Borensztein, M.; Monnier, P.; Court, F.; Louault, Y.; Ripoché, M.A.; Tiret, L.; Yao, Z.; Tapscott, S.J.; Forne, T.; Montarras, D.; *et al.* Myod and H19-Igf2 locus interactions are required for diaphragm formation in the mouse. *Development* **2013**, *140*, 1231–1239.
116. Luo, M.; Li, Z.; Wang, W.; Zeng, Y.; Liu, Z.; Qiu, J. Upregulated H19 contributes to bladder cancer cell proliferation by regulating ID2 expression. *FEBS J.* **2013**, *280*, 1709–1716.
117. 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.
118. 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.
119. Jacob, E.; Hod-Dvorai, R.; Ben-Mordechai, O.L.; Boyko, Y.; Avni, O. Dual function of polycomb group proteins in differentiated murine T helper (CD4+) cells. *J. Mol. Signal.* **2011**, *6*, 5.
120. Jo, S.; Lee, H.; Kim, S.; Hwang, E.M.; Park, J.Y.; Kang, S.S.; Chung, H. Inhibition of PCGF2 enhances granulocytic differentiation of acute promyelocytic leukemia cell line HL-60 via induction of HOXA7. *Biochem. Biophys. Res. Commun.* **2011**, *416*, 86–91.
121. Porcher, C.; Swat, W.; Rockwell, K.; Fujiwara, Y.; Alt, F.W.; Orkin, S.H. The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell* **1996**, *86*, 47–57.
122. 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.
123. 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.
124. Savarese, F.; Flahndorfer, K.; Jaenisch, R.; Busslinger, M.; Wutz, A. Hematopoietic precursor cells transiently reestablish permissiveness for X inactivation. *Mol. Cell. Biol.* **2006**, *26*, 7167–7177.
125. Ahn, J.H.; Lee, Y.; Jeon, C.; Lee, S.J.; Lee, B.H.; Choi, K.D.; Bae, Y.S. Identification of the genes differentially expressed in human dendritic cell subsets by cDNA subtraction and microarray analysis. *Blood* **2002**, *100*, 1742–1754.
126. Lee, M.Y.; Kim, W.J.; Kang, Y.J.; Jung, Y.M.; Kang, Y.M.; Suk, K.; Park, J.E.; Choi, E.M.; Choi, B.K.; Kwon, B.S.; *et al.* Z39Ig is expressed on macrophages and may mediate inflammatory reactions in arthritis and atherosclerosis. *J. Leukocyte Biol.* **2006**, *80*, 922–928.

Reprinted from *IJMS*. Cite as: Rosa, A.; Brivanlou, A.H. Regulatory Non-Coding RNAs in Pluripotent Stem Cells. *Int. J. Mol. Sci.* **2013**, *14*, 14346-14373.

Review

Regulatory Non-Coding RNAs in Pluripotent Stem Cells

Alessandro Rosa ^{1,2} and Ali H. Brivanlou ^{2,*}

¹ Department of Biology and Biotechnology “Charles Darwin”, Sapienza University of Rome, Rome 00185, Italy; E-Mail: alessandro.rosa@uniroma1.it

² Laboratory of Molecular Vertebrate Embryology, The Rockefeller University, New York, NY 10065, USA

* Author to whom correspondence should be addressed; E-Mail: brvnlou@rockefeller.edu; Tel.: +1-212-327-8656; Fax: +1-212-327-8685.

Received: 6 May 2013; in revised form: 25 June 2013 / Accepted: 2 July 2013 /

Published: 11 July 2013

Abstract: The most part of our genome encodes for RNA transcripts are never translated into proteins. These include families of RNA molecules with a regulatory function, which can be arbitrarily subdivided in short (less than 200 nucleotides) and long non-coding RNAs (ncRNAs). MicroRNAs, which act post-transcriptionally to repress the function of target mRNAs, belong to the first group. Included in the second group are multi-exonic and polyadenylated long ncRNAs (lncRNAs), localized either in the nucleus, where they can associate with chromatin remodeling complexes to regulate transcription, or in the cytoplasm, acting as post-transcriptional regulators. Pluripotent stem cells, such as embryonic stem cells (ESCs) or induced pluripotent stem cells (iPSCs), represent useful systems for modeling normal development and human diseases, as well as promising tools for regenerative medicine. To fully explore their potential, however, a deep understanding of the molecular basis of stemness is crucial. In recent years, increasing evidence of the importance of regulation by ncRNAs in pluripotent cells is accumulating. In this review, we will discuss recent findings pointing to multiple roles played by regulatory ncRNAs in ESC and iPSCs, where they act in concert with signaling pathways, transcriptional regulatory circuitries and epigenetic factors to modulate the balance between pluripotency and differentiation.

Keywords: embryonic stem cells; induced Pluripotent Stem Cells; microRNA; long non-coding RNA; pluripotency; reprogramming

1. Introduction

Pluripotency is defined, in a broad sense, as the ability of a cell to give rise to derivatives of the three germ layers, ectoderm, mesoderm and endoderm, and the germ line. In mammals, this property is restricted to embryonic cells until the stage of blastocyst, in which a small number of cells that constitute the Inner Cell Mass (ICM) are still pluripotent. Such pluripotent cells exist very transiently during embryonic development. Upon gastrulation and during subsequent embryonic development, they lose pluripotency and progressively acquire a specialized character. Embryonic stem cells (ESCs) can be derived from the ICM of human and mouse blastocysts; are pluripotent and can self-renew *in vitro* indefinitely.

ESC pluripotency is tightly regulated. Amongst several signaling pathways, the TGF- β pathway has been shown to play a central role [1]. Interestingly, the two branches of the pathway play different roles in human and in mouse. More precisely, in human ESCs (hESCs) the Nodal/Activin branch is both necessary and sufficient to sustain pluripotency [2,3], whereas in mouse ESC (mESCs) the bone morphogenetic protein (BMP) branch is required for maintaining self-renewal and prevent differentiation [4]. Downstream of signaling pathways, the maintenance of ESCs pluripotency is ensured by a regulatory circuitry including three main core transcription factors (TFs), Oct4, Sox2 and Nanog [5–7]. The three core TFs co-occupy a conspicuous set of target promoters and have a dual role. They can activate transcription of genes involved in the maintenance of pluripotency, including their own genes. At the same time, in association with Polycomb Repressive Complexes (PRC1 and PRC2), they silence a subset of lineage-specific genes that play a role in development [8,9]. In ESCs, the promoters of these genes present peculiar bivalent chromatin domains, in which epigenetic histone modifications normally associated with silent genes co-exist with marks of active transcription [10]. Such unique epigenetic profiles are required to keep key developmental genes “poised” in a repressed state that can be quickly turned on.

Pluripotent cells exist in two different states, defined as naïve and primed [11]. Mouse ESCs are considered to be in a naïve ground state of pluripotency that corresponds to the preimplantation epiblast. *In vitro*, they require Leukemia Inhibitory Factor (LIF) and BMP signaling for self-renewal and differentiate in presence of basic Fibroblast Growth Factor (bFGF). They express a typical set of markers and female cells have two active X chromosomes. Functionally, they are able to form chimeric embryos upon injection in a recipient blastocyst. Conversely, mouse Epiblast Stem Cells (EpiSCs) are in a primed state, which corresponds to the post-implantation epiblast. EpiSCs require bFGF and Nodal/Activin signaling for self-renew, present X chromosome inactivation in female cells, and fulfill some criteria for pluripotency, such as the ability to form teratomas, but cannot generate chimeric mice. There is some plasticity in pluripotent cells, as mESCs can be induced to differentiate into EpiSCs by culturing in presence of Activin and bFGF and EpiSCs can be converted to the ground state by transfection with Klf4 or Nanog and culturing in the presence of LIF and BMP [12,13]. Human ESCs share many features with EpiSCs but differ from mESCs in terms of morphology, culture requirements, expression markers and X chromosome inactivation. This evidence led to the hypothesis that hESCs are in a primed state of pluripotency [14].

The seminal work by Shinya Yamanaka showed that pluripotency could be induced in mouse somatic cells by ectopic expression of a defined set of reprogramming factors (RFs) [15]. Since then,

reprogramming of somatic cells into induced Pluripotent Stem Cells (iPSCs) has been achieved in human as well [16,17]. Alternative sets of RFs that contain well-known pluripotency factors can be used. The most common are Oct4, Sox2, Klf4 and c-Myc (Yamanaka RFs) and Oct4, Sox2, Nanog and Lin28 (Thomson RFs). iPSCs share with ESCs the pluripotency, meaning that they could virtually be differentiated *in vitro* into all adult cell types. Moreover, they can be derived from human patients as patient-specific iPSCs (PS-iPSCs) that hold the same disease-causing genetic alteration [18]. The mechanisms underlying reprogramming have been deeply investigated and involve a profound change in cell identity. During reprogramming, the epigenetic landscape of the somatic cell of origin shifts to a state proper of the embryonic stem cell, including erasure of repressive marks on the chromatin of pluripotency genes and establishing of bivalent domains on lineage-specific genes [19].

The ability to differentiate into multiple tissues makes ESCs and iPSCs promising tools for regenerative medicine and cell-replacement therapy approaches [20]. However, to fully exploit their potential, the molecular basis of pluripotency must be deeply characterized. Non-coding RNA (ncRNA) molecules, previously regarded to exert only passive roles in the cell, are conversely primary players to define the cell identity. Rather than the coding portion of the genome, it is now clear that its non-coding counterpart is correlated with the greater complexity of higher eukaryotes [21]. Recently, ncRNAs are also emerging as new regulatory factors in pluripotent cells. Among small non-coding RNAs (<200 nucleotides), microRNAs (miRNAs) are now considered major regulators of development, metabolism, differentiation and homeostasis in all multicellular organisms [22–26]. miRNAs are also involved in several human diseases, including cancer [27]. Biogenesis of miRNAs requires a multistep process [23]. miRNAs are generally transcribed by RNA polymerase II as part of introns of mRNA genes, or from intergenic regions. The miRNA primary precursor (pri-miRNA) is then processed in the nucleus by the Microprocessor complex [28,29], comprised of the cleavage enzymes Drosha, DGCR8 and other factors, releasing a stem-loop precursor (pre-miRNA). The pre-miRNA is then exported in the cytoplasm and cleaved by the RNase III enzyme Dicer, which is also involved in the maturation of short interfering RNAs (siRNAs) [30,31]. The mature miRNA is finally incorporated as a single strand in the RNA Induced Silencing Complex (RISC) [32]. Guided by the miRNA, RISC binds the 3'UTR and/or the coding sequences of target mRNAs resulting in inhibition of translation and/or degradation. A single miRNA can inhibit several targets and a single mRNA can be targeted by multiple miRNAs in a combinatorial way [33]. For target recognition by miRNAs, base pairing of a short “seed” sequence, located at the 5'end of the mature miRNA, is required [34]. Families of miRNAs comprise members with identical seed sequences and all miRNAs belonging to a family are thought to share the same targets. Multiple miRNAs can be excised from a single, multicistronic, pri-miRNA transcript. Such clusters can comprise multiple members of a miRNA family as well as unrelated miRNAs.

Despite lacking an open reading frame (ORF), long (>200 nucleotides) non-coding RNAs (lncRNAs) are transcribed by RNA polymerase II and have a structure that resemble protein-coding mRNAs. They are generally spliced from multi-exonic precursors, have a 5'cap and are polyadenylated [35,36]. Some lncRNAs are transcribed from loci which overlap or are very close to protein-coding genes, while others are intergenic (long intergenic non-coding RNAs, lincRNAs). Globally, lncRNA expression levels are only slightly lower than protein-coding transcripts [37,38]. Similar to protein-coding and miRNA genes, and differently from other structural non-coding RNAs such as ribosomal RNAs, the expression of lncRNAs is characterized by tissue specificity, and is dynamically regulated during

differentiation [39–42]. Moreover, many lncRNA are conserved in terms of sequence and predicted secondary structure [37,38,43]. All together, these elements suggest that lncRNAs may represent a new, still largely unexplored class of functional molecules, potentially involved in multiple biological processes [44–46]. In the last years, the continuous development of powerful sequencing technologies and bioinformatics tools boosted the discovery of thousands novel lncRNAs. According to some estimates, the total number of human lincRNAs would be around 4500 [47]. So far, however, a function has been only assigned to a limited number of lncRNAs. Intracellular localization is often used as a predictive element to get insights into lncRNA molecular mechanisms [48]. Nuclear lncRNAs can regulate gene expression either in cis (on neighboring genes) or in trans (on distant genes). Some of them are able to modulate the activity of chromatin modifiers. A paradigmatic example for this class of lncRNAs is HOTAIR, which is required for Polycomb Repressive Complex 2 (PRC2) occupancy and histone H3 lysine-27 trimethylation of the HOXD locus [49]. Other nuclear lncRNAs act as antisense transcripts or as decoy for splicing factors [50,51]. Cytoplasmic lncRNAs may conversely function as endogenous “sponges” for miRNAs, thus releasing miRNA repression on target genes [52,53].

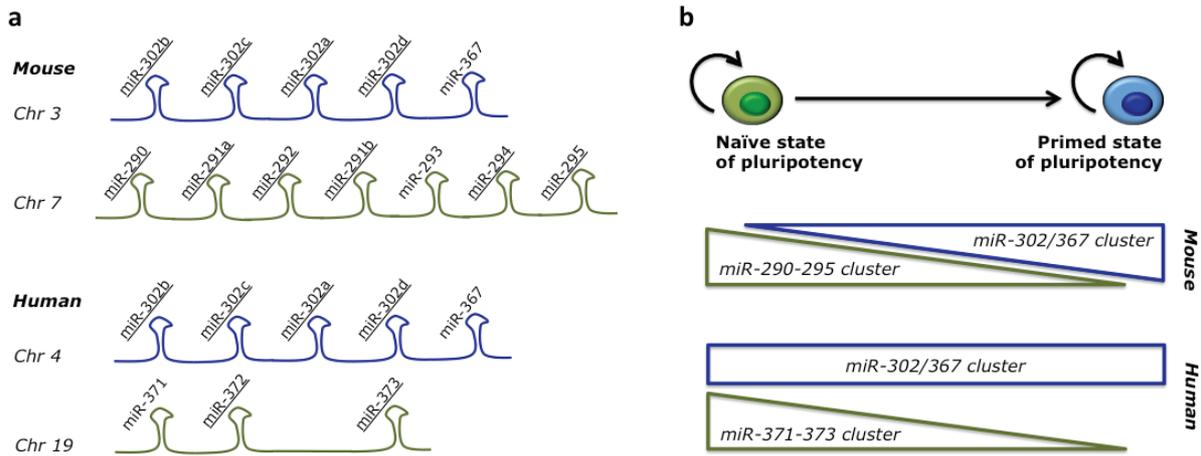
Here we will review recent findings that point to regulatory non-coding RNAs as important players in the maintenance of ESCs pluripotency and cell fate choices during differentiation. We will also discuss the roles played by ncRNAs during somatic cells reprogramming.

2. MicroRNAs in Embryonic Stem Cells

Among all miRNAs expressed in embryonic stem cells, a single family of miRNAs with the AAGUGC seed sequence is the most highly expressed and has the most functional impact. Members of this family are organized in two major clusters (Figure 1a). The conserved miR-302/367 cluster comprises four AAGUGC seed-containing miRNAs (miR-302a, miR-302b, miR-302c and miR-302d) and the unrelated miR-367. The second cluster is less conserved. In the mouse it is commonly referred to as the miR-290-295 cluster and includes six miRNAs with the AAGUGC seed (miR-290, miR-291a, miR-291b, miR-292, miR-294 and miR-295) and miR-293. The human orthologue comprises miR-371 and the AAGUGC seed-containing miR-372 and miR-373 (miR-371-373 cluster). Therefore, multiple members with the AAGUGC seed are present in each cluster, as well as other miRNAs with a different seed [25]. For simplicity, in this review we will refer to miRNAs containing the AAGUGC seed as miR-302 family.

The miR-302 family miRNAs are abundantly expressed in undifferentiated ESCs and decline upon differentiation [54–58]. According to the miRNA expression atlas [59], miR-302 family members are specifically expressed in embryonic cells in both mouse and human. In mouse, high-throughput sequencing indicated that miRNAs belonging to the miR-290-295 cluster represent almost one third of all miRNAs expressed in undifferentiated mESCs [60]. Similarly, undifferentiated human ESCs are dominated by the miR-302 cluster, which accounts for more than 60% of all expressed miRNAs [61]. miRNAs with the AAGUGC seed are also conserved in other vertebrates, such as Zebrafish, Xenopus and Chicken, where they seem to be specifically expressed during early embryonic development [25].

Figure 1. Clusters of miRNAs specifically expressed in embryonic stem cells (ESCs). (a) Schematic representations of the two ESC-specific miRNA clusters in mouse and human. miRNAs with the AAGUGC seed sequence within each cluster are underlined. The miR-302/367 cluster is highly conserved in mammals, whereas the other cluster is less conserved; (b) The two clusters are differentially expressed in ESCs during the conversion between the naïve and primed states of pluripotency.



The specific expression of miR-302 family miRNAs is ensured by their regulation, at the transcriptional level, by the core ESC transcriptional regulatory circuitry [8,9,62,63]. The promoters of the miR-302 and the miR-290-295 clusters are bound by Oct4, Nanog, Sox2 and Tcf3, that also promote transcription of other unrelated miRNAs. The same transcription factors, in cooperation with Polycomb group proteins, repress transcription of lineage-specific miRNAs, such as the neural miRNAs miR-9 and miR-124 and the mesodermal miRNA miR-155 [9].

Interestingly, the conversion from a naïve to a primed state in mESCs correlates with a switch between the miR-290-295 and the miR-302/367 clusters in terms of miRNA abundance [64] (Figure 1b). mESCs express high levels of miR-290-295 that decline after conversion to EpiSCs and are replaced by an increase of miR-302/367. The total levels of miRNAs with the AAGUGC seed are maintained. Since miRNAs with the same seed should share the same targets, the significance of this switch remains obscure. In human ESCs, which correspond to a primed state of pluripotency, the levels of miR-302/367 are much higher than the levels of miR-371-373 and the switch to an earlier developmental state led to an increase of the levels of miR-371-373 [65–67].

2.1. miRNAs Regulate ESC Cell Cycle

Inactivation of genes involved in the processing of miRNAs represents a useful approach to study the global function of miRNAs. Dicer $-/-$ pre-gastrulation embryos showed lack of Oct4-positive epiblast cells and could not undergo gastrulation [68], indicating that the miRNA pathway is essential for establishing the pool of pluripotent embryonic cells. As expected, mESCs could not be derived from Dicer mutant embryos. Surprisingly, Dicer $-/-$ mESCs obtained by conditional gene targeting are viable, hold an appropriate morphology, and express normal levels of pluripotency markers [69]. However, the miRNA pathway is crucial for mESC differentiation both *in vitro* and *in vivo*, as assessed by defects in lineage markers expression and teratoma formation. Moreover, inactivation of

Dicer impairs the proliferation potential of mESCs [70]. As Dicer is involved in the processing of both miRNAs and other endogenous small RNAs, the observed phenotype could be due to the loss of both classes of Dicer targets. However, comparative analysis of Dicer $+/+$ and Dicer $-/-$ mESCs by high-throughput sequencing indicated that very few short non-miRNA transcripts that are dependent on Dicer exist in these cells [60]. Moreover, a similar phenotype, albeit less severe, was observed in DGCR8 $-/-$ mESCs [71], indicating that the defects detected in Dicer $-/-$ mESCs are mainly due to the loss of miRNAs. A comparative transcriptome analysis revealed that in cells lacking all miRNAs there was a significant increase of transcripts harboring a GCACUU motif in the 3'UTR [72]. This sequence is complementary to the AAGUGC seed of miR-302 family miRNAs, thus confirming their prominent role in pluripotent stem cells. Accordingly, miR-302 family members could rescue the proliferation defects of DGCR8 mutant mESCs [73]. For their role in the regulation of the characteristic cell cycle of ESCs, these miRNAs have been called “ES-cell-specific cell-cycle-regulating” (ESCC) miRNAs [74]. Specifically, they target the Cyclin Dependent Kinase (CDK) inhibitors p21 (Cdkn1a), Rbl2 and Lats2, which are present at low levels in wt mESCs and overexpressed upon disruption of the miRNA pathway [72,73,75]. Other miRNAs may contribute to the silencing of cell cycle inhibitors in mESCs. For example, miR-320 and miR-702 target p21 and p57 [76]. Interestingly, these are non-canonical miRNAs, which require Dicer, but not DGCR8, for their biogenesis [77]. The activity of such non-canonical miRNAs might explain the more severe phenotype observed in mESCs upon loss of Dicer as compared to loss of DGCR8 [76].

Similar to mESCs, also hESCs are characterized by a shortened cell cycle, accumulating in the S phase at the expense of the G1 phase [78]. Regulation of embryonic stem cell cycle by miRNAs is evolutionary conserved, as knockdown of Dicer or Drosha impairs human ESCs proliferation [79]. This defect could be partially rescued by miR-302 family members and the unrelated miR-195. These miRNAs play distinct roles in regulating cell cycle progression. miR-195 facilitates the G2/M transition by targeting the WEE1 kinase, a negative regulator of the Cycline B/CDK complex. The miR-302 family targets p21 and promotes G1/S transition [79,80]. Another miRNA, miR-92b, participates in the regulation of the G1/S transition by targeting the CDK inhibitor p57 (Cdkn1c) [81]. Moreover, at least other two components of the cell cycle machinery, Cyclins D1 and D2, are under the control of miR-302 family members in hESCs [63,82]. Therefore, the miRNA pathway plays a major role in maintaining the peculiar cell cycle of pluripotent embryonic cells.

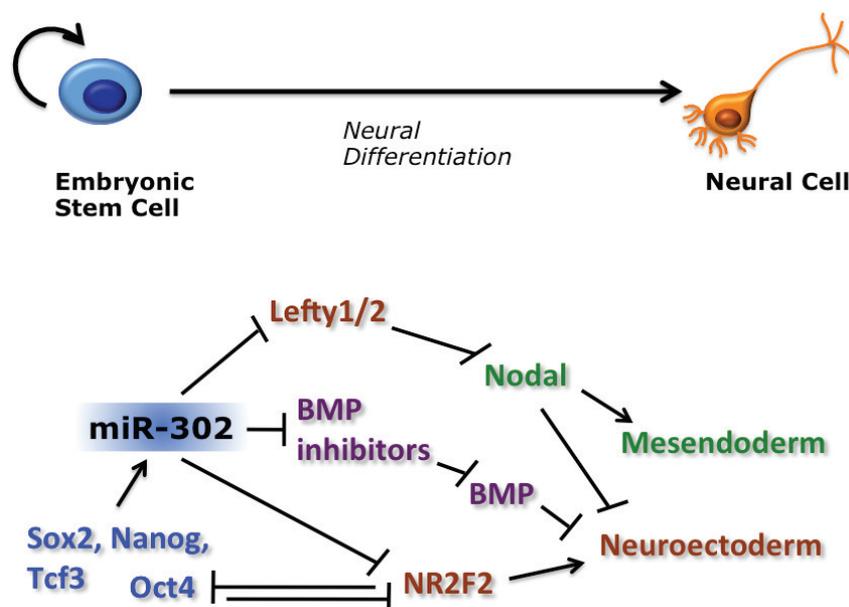
2.2. miRNAs Regulate ESC Pluripotency and Differentiation

A genome-wide approach has been recently performed in hESCs to systematically identify bona fide miRNA targets [61]. PAR-CLIP (photoactivatable ribonucleoside-enhanced cross-linking and immunoprecipitation) [83] for the RISC component AGO2, combined with perturbation of miRNA levels, led to the identification of a set of 146 high-confidence miR-302-367 direct targets. These included previously validated targets, as well as novel genes regulated by this cluster. This gene set is enriched in factors involved in the maintenance of pluripotency, such as regulators of cell cycle and proliferation, chromatin modification, metabolism and signaling [84].

Despite holding normal morphology and expression of pluripotency markers, both mESCs and hESCs depleted of miRNAs cannot properly differentiate [69,71,79]. The miR-302 family has been

involved in the TGF- β /BMP signaling pathway, which regulates embryonic stem cells pluripotency and differentiation. During hESC differentiation, Nodal activity inhibits neuroectodermal specification and promotes mesendodermal formation, whereas overexpression of the Nodal inhibitor, Lefty, results in the opposite effect [85,86]. We have shown that, by directly inhibiting Lefty, miR-302 is necessary for proper mesoderm and endoderm specification [87]. The regulation of Nodal signaling by miR-302 seems evolutionary conserved, as the *Xenopus* and Zebrafish orthologues target Lefty during early embryogenesis [87,88]. In addition to the inhibition of Nodal antagonists, miR-302 could promote BMP signaling by targeting the inhibitors DAZAP2, SLAIN1, and TOB2 [61]. Double blockage of the two branches of the TGF- β pathway induces a neural fate in ESCs [89,90]. By inhibiting inhibitors of both branches, miR-302 plays a central role in negatively regulating neural induction in pluripotent stem cells (Figure 2). It has recently been shown that BMP signaling down-regulates the miR-302/367 cluster in human primary pulmonary artery smooth muscle cells (PASMCs), mouse mesenchymal cells and embryonic carcinoma p19 cells [91]. In a feedback loop, miR-302 targets the type II BMP receptor (BMPRII) in PASMCs, thus inhibiting BMP signaling. Further studies are necessary to address whether this reciprocal inhibition between BMP signaling and miR-302 also exist in ESCs.

Figure 2. Role of miR-302 during human ESCs (hESC) neural differentiation. Transcription from the miR-302 locus is activated by the ESC core transcriptional regulatory circuitry. miR-302 post-transcriptionally inhibits NR2F2 which in turn is an inhibitor of Oct4. In a negative feedback loop, Oct4 inhibits NR2F2 transcription. NR2F2 expression is necessary for proper activation of neuroectoderm genes. miR-302 also targets inhibitors of both branches of the TGF β pathway: the Nodal inhibitors, Lefty1 and Lefty 2, and the bone morphogenetic protein (BMP) inhibitors, DAZAP2, SLAIN1 and TOB2. According to the neural default model, inhibition of both branches of the TGF β pathway leads to neural induction. Thus, by targeting inhibitors of both branches miR-302 has a negative effect on neural induction.



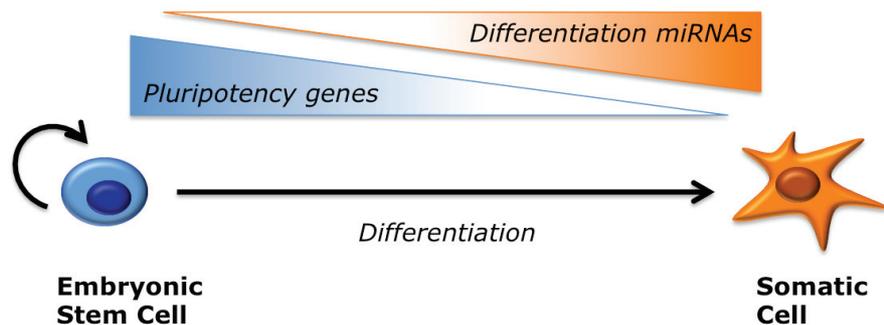
Interestingly, a regulatory loop comprising an unrelated miRNA, miR-125, and components of the BMP signaling has recently been described in mouse embryonic stem cells [92]. Such regulatory circuitry sets mESC sensitivity to BMP4. Therefore, the interplay between the TGF- β /BMP pathway and miRNAs may represent a more general regulatory mechanism modulating the response of ESC to extracellular stimuli. The interaction between miRNAs and signaling pathways underlying pluripotency might not be limited to TGF- β . In cancer cells, the Wnt/ β -catenin pathway causes aberrant expression of the miR-371-373 cluster and these miRNAs, in turn, repress the activity of Wnt inhibitors, such as DKK1 [93]. It would be interesting to assess whether such interplay between the Wnt signaling and the miR-371-373 cluster has also a role in ESC pluripotency.

miR-302 is deeply integrated in the core transcriptional regulatory circuitry of ESCs. As mentioned above, the miR-302 host gene is under the control of Oct4, Nanog and Sox2, which ensure high miRNA levels in undifferentiated ESCs [9,62,63]. We have shown that both Oct4 (at the transcriptional level) and miR-302 (post-transcriptionally) repress a common target, NR2F2 (also known as COUP-TFII) [94]. NR2F2 in turn is an inhibitor of Oct4. It is activated during early neural ectoderm induction and is necessary for the proper expression of neural genes upon hESC differentiation [94]. Therefore, miR-302 and the two transcription factors, NR2F2 and Oct4, form a feedback regulatory circuitry that regulates hESC exit from pluripotency and neural fate specification (Figure 2). Individual pluripotent cell lines have different propensity to differentiate along specific lineages. Interestingly, miR-371-373 expression levels negatively correlate with the neurogenic differentiation propensity of hESC and hiPSC lines [67].

Epigenetic silencing of the Oct4 locus is necessary to ensure proper differentiation upon exit from pluripotency. This is achieved by de novo DNA methylation by DNMT3 factors. It has been shown that miR-290-295 play a role in this process by inhibiting Rbl2, that in turn is an inhibitor of DNMT3 [72]. In the absence of miR-290-295 there would be incomplete silencing of the Oct4 locus during differentiation. This would in part explain the differentiation defects observed in *Dicer* $-/-$ cells, which maintain high levels of pluripotency factors and fail to activate lineage specific gene programs [69]. Other miRNAs may contribute to this phenotype, inhibiting the expression of pluripotency genes during mESC differentiation (Figure 3). For instance, miR-134, miR-296 and miR-470 target the core transcription factor trio, Oct4, Sox2 and Nanog [95]. Sox2 and Klf4 are also repressed by miR-200c, miR-183 and miR-203 [96]. It is interesting to notice that miR-203 has also a role in skin stem cell terminal differentiation by inhibiting p63 [97]. Pluripotency factors are under the control of miRNAs also in human ESC. An interesting regulatory loop has been shown for miR-145, which targets Oct4, Sox2 and Klf4 during differentiation, and in turn is repressed at the transcriptional level by Oct4 in undifferentiated cells [98]. Recently, it has been proposed that another role for miRNAs during ESC differentiation is to modulate the activity of chromatin modifiers. An interesting switch between different variants of the Polycomb Repressive Complex 1 (PRC1) has been shown during ESC differentiation. In particular, an ES-specific PRC1, containing the Cbx7 subunit, is replaced by a differentiation-specific PRC1, containing the subunits Cbx2/4/8 [99,100]. This switch is crucial, as Cbx proteins confer distinct target selectivity to the PRC1 complex. The different Cbx variants inhibit each other. In pluripotency conditions, Cbx7 negatively regulates the other Cbx that are conversely induced during differentiation, and repress Cbx7 transcription. miRNAs of the miR-125 and miR-181 families contribute to the downregulation of Cbx7 during mESC differentiation by

directly targeting its 3'UTR [100]. Overexpression of these miRNAs in mESCs caused loss of pluripotency markers and increased expression of a subset of PRC1 target genes involved in lineage specification.

Figure 3. miRNAs target pluripotency genes. The activity of pluripotency genes, including core transcription factors, must be shut down as embryonic stem cells differentiate. Several miRNAs that are induced upon exit from pluripotency directly target these genes in a combinatorial way. Examples are shown in the table. M: mouse; H: human.

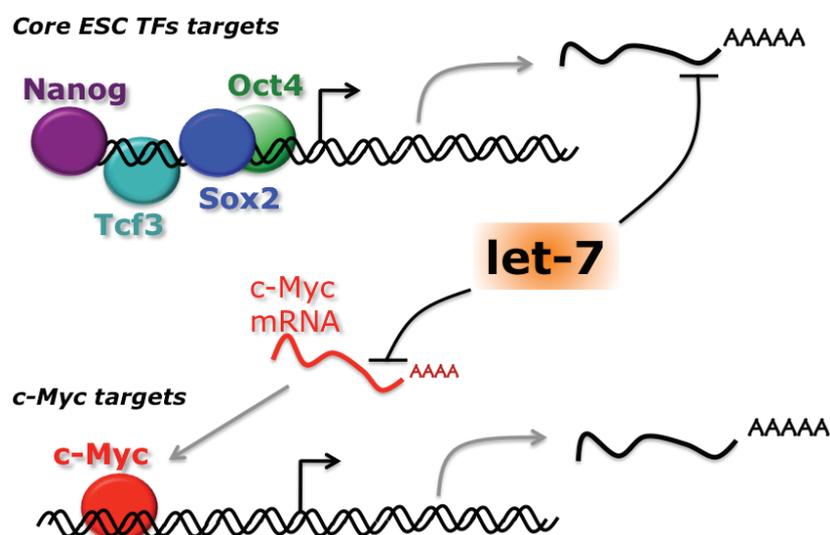


| <i>miRNAs</i> | <i>Targets</i> | <i>Species</i> | <i>References</i> |
|----------------------------|--------------------|----------------|-------------------------|
| miR-134, miR-296, miR-470 | Oct4, Sox2, Nanog | M | Tay et al., 2008 |
| miR-200c, miR-183, miR-203 | Sox2, Klf4 | M | Wellner et al., 2009 |
| miR-145 | Oct4, Sox2, Klf4 | H | Xu et al., 2009 |
| let-7 | cMyc, Lin28, Sall4 | M | Melton et al., 2010 |
| miR-125, miR-181 | Cbx7 | M, H | O'Loughlen et al., 2012 |

Among other miRNAs with a role in ESC differentiation, the let-7 family plays a prominent role. This is a highly conserved family with orthologues in all metazoa. In mammals it comprises several let-7 species (let-7a to let-7i) and other miRNAs, such as miR-98 and miR-202 [101]. Members of the let-7 family are induced during development and differentiation, with a parallel reduction of their targets, and have a role in cancer [102]. In ESCs, levels of let-7 miRNAs are regulated at the post-transcriptional level. Whereas the primary transcript and the hairpin precursor accumulate in these cells, the production of the mature miRNA is blocked [103]. The underlying molecular mechanism relies on the recognition of the terminal loop of let-7 precursors by the RNA binding protein Lin28 [104–106]. Lin28 recruits the terminal uridylyl transferase TUT4, which in turn adds a poly-U tail, targeting the miRNA for degradation [107,108]. A feedback regulatory loop exists, in which let-7 negatively regulates Lin28 [109]. Lin28 is highly expressed in undifferentiated ESCs and declines during differentiation, when levels of mature let-7 increase. This switch is ensured by their mutual repression. As mentioned above, DGCR8 $-/-$ ESCs maintain high levels of pluripotency genes expression. Upon transfection of let-7 family miRNAs, expression of Oct4, Sox2 and Nanog is inhibited, suggesting an anti-pluripotency activity for let-7. However, if miR-302 family miRNAs are co-transfected, they impair this activity and restore the levels of pluripotency markers [110].

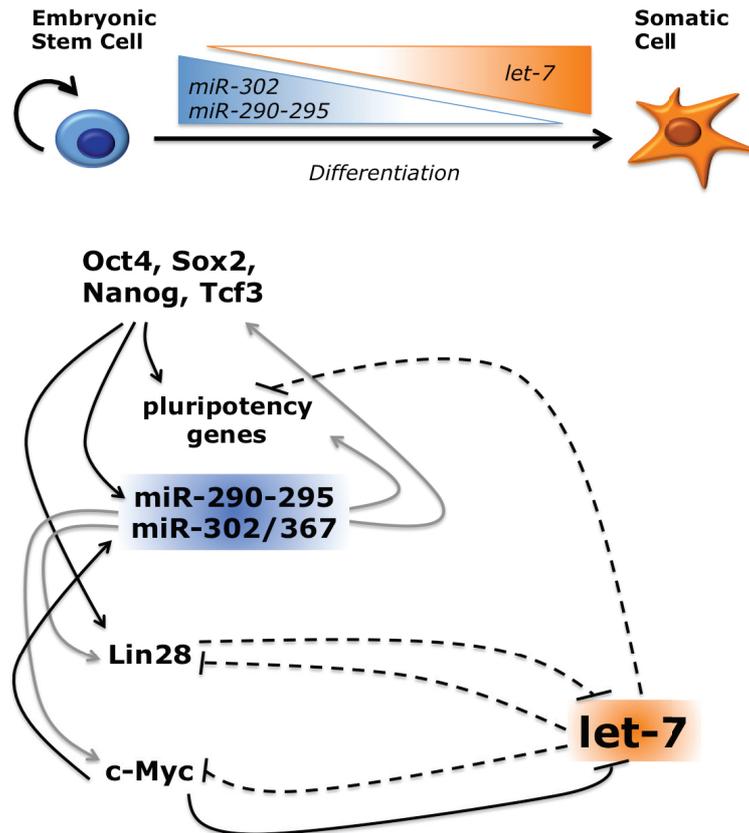
Comparative transcriptome analysis showed that let-7 introduction in DGCR8 $-/-$ mESCs led to the inhibition of genes that are transcriptionally activated by Myc factors. Both c-Myc and N-Myc are direct targets of let-7 [110]. Myc genes play a well-established role in ESC self-renewal [111]. Moreover, let-7 directly targets transcripts that are induced by other pluripotency factors, Oct4, Sox2, Nanog and Tcf3 [110]. Inhibition of the function of pluripotency factors, direct for Myc and indirect for other factors, explains the anti-pluripotency activity of let-7 (Figure 4). Conversely, Myc-activated genes and c-Myc itself are enriched among transcript that are upregulated in presence of miR-302 family members, suggesting that the pro-pluripotency activity of these miRNAs may be mediated, at least in part, by the indirect increase of Myc [110].

Figure 4. let-7 antagonizes pluripotency networks. Let-7 antagonizes indirectly the activity of the ESC core transcriptional regulatory circuitry by targeting multiple genes induced by the core transcription factors (TFs). Moreover, let-7 directly inhibits c-Myc, reducing transcription of its target genes.



For both let-7 and the miR-302 families feedback loops with Myc are in play. The promoter of the miR-290-295 cluster is directly bound and activated by c-Myc and N-Myc [7], and the miR-302 cluster is also induced by Myc [112], establishing positive feedback loops. Conversely, in cancer cells c-Myc has been shown to bind the promoters and repress transcription of several let-7 genes [113]. Other pluripotency factors that are directly inhibited by let-7, and indirectly activated by miR-302 family, are Lin28 and Sall4, suggesting that these miRNAs exert their function via multiple pathways [110]. Therefore, according to the model depicted in Figure 5, the miR-302 and let-7 miRNA families play opposite, crucial roles in regulating ESC pluripotency and differentiation.

Figure 5. Opposing roles for the let-7 and miR-302 families. During ESC differentiation, the levels of pluripotency miRNAs belonging to the miR-302 family are decreased. In parallel, let-7 family members are induced. The miR-302/367 and miR-290-295 clusters are induced by core ESC TFs, and in a positive feedback loop indirectly promote their expression and transcription of their targets. Conversely, let-7 is post-transcriptionally inhibited by Lin-28. In cancer cells c-Myc inhibits let-7 at the transcriptional level. In a negative feedback loop, let-7 targets its inhibitors. Black arrows indicate direct activation; grey arrows indicate indirect activation; dashed lines indicate post-transcriptional inhibition.



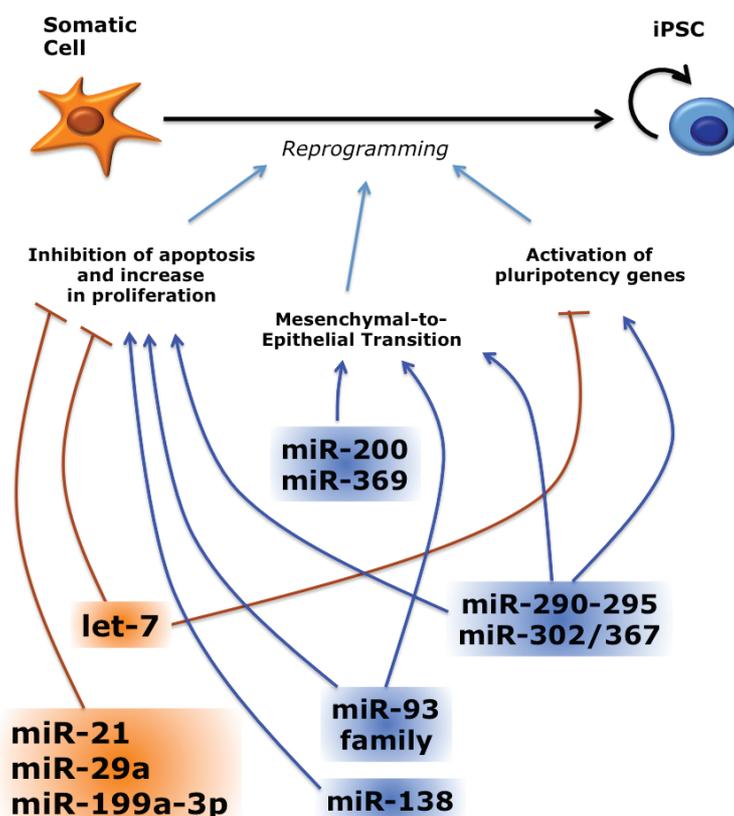
3. MicroRNAs and Reprogramming to iPSCs

ESCs and iPSCs express a similar signature group of miRNAs, including the miR-302 family, with small differences between the two cell types [114–116]. When genetically identical mouse ESCs and iPSCs were analyzed, these differences were circumscribed to miRNAs encoded in an imprinted locus on chromosome 12qF1 [117]. Imprinted loci are transcribed in a parental specific manner and contain clusters of protein-coding and noncoding genes [118]. The 12qF1 locus includes several maternally expressed genes, including two miRNA clusters, that are aberrantly silenced in many, but not all, iPSC lines [117]. Interestingly, iPSC lines in which this imprinted locus was silenced were unable to generate all-iPSC mice upon tetraploid blastocyst complementation, due to arrest of embryonic development around mid-gestation. Conversely, iPSC lines with normal expression of the imprinted locus generated viable all-iPSC mice. Therefore, the transcriptional status of imprinted genes in the 12qF1 locus, including miRNAs, marks the development potential of different iPSC clones

3.1. Several miRNAs Promote or Inhibit Reprogramming

Recently, several miRNAs have been directly involved in the reprogramming process (Figure 6). Knockdown of RISC components globally impairs miRNA activity and leads to a dramatic decrease of reprogramming activity [119]. Much work concerned the miRNAs belonging to the miR-302 family. When the minimal Oct4/Sox2/Klf4 (OSK) cocktail was provided to mouse embryonic fibroblasts (MEFs) in combination with these miRNAs, the efficiency of iPSCs generation was substantially increased [119,120]. Interestingly, in presence of the three factors plus c-Myc the reprogramming enhancement by the miR-302 family members was strongly reduced [120,121]. Other miRNAs may also promote reprogramming in the mouse. For instance, miR-93 and 106b (that belong to the same family and share 5/6 of the miR-302 seed) and the unrelated miR-138 enhance iPSC generation [118,122]. Similar to the mouse system, reprogramming of human fibroblasts was also enhanced when miR-302 family members are provided along with the reprogramming factors [123].

Figure 6. Role of miRNAs during reprogramming. The reprogramming process can be divided in three phases. Early events include inhibition of apoptosis and an increase in proliferation. A mesenchymal-to-epithelial transition (MET) occurs in the intermediate phase. Activation of the endogenous pluripotency program occurs as a late event. Several miRNAs have been shown to promote (blue) or inhibit (orange) reprogramming by facilitating or hampering the completion of these events. The miR-93 family includes miR-93 and miR-106b.



As expected, let-7 has an opposite role during reprogramming. MEFs express high levels of let-7 and this anti-pluripotency miRNA family must be shut down during conversion into iPSCs. In the

presence of let-7 inhibitors, reprogramming by the OSK cocktail was enhanced [110]. When OSK plus c-Myc were used, this increase of efficiency by let-7 inhibition was much less pronounced. Again, let-7 inhibits c-Myc, so the inhibition of let-7 might just increase c-Myc activity. Lin28, which is a component of the Thomson reprogramming cocktail and is induced during reprogramming with the Yamanaka factors, is another target of let-7. As in the case of c-Myc, addition of Lin28 to the reprogramming cocktail accelerates proliferation [124]. However, there is not an increase in MEF proliferation upon let-7 inhibition [110], suggesting that other targets besides c-Myc and Lin28 contribute to the effects of let-7 on the reprogramming process. Besides let-7, other miRNAs negatively affect reprogramming. For instance, inhibition of miR-21, miR-29a and miR-199a-3p enhances reprogramming [125,126].

3.2. miRNAs Reprogram Somatic Cells in the Absence of Protein-Coding Reprogramming Factors

Given their multiple roles in ESCs, the increase of reprogramming efficiency by miRNAs in combination with reprogramming factors was not unexpected. More surprisingly, recent work has shown that cocktails of miRNAs, without canonical reprogramming factors, can be sufficient to reprogram both mouse and human somatic cells. The Morrisey lab demonstrated that the miR-302-367 cluster induced pluripotency with two orders of magnitude more efficiency than standard methods [127]. Importantly iPSCs generated with miR-302-367 fulfilled stringent criteria that define bona fide pluripotent stem cells, such as the ability to contribute to chimeras with germ line transmission (mouse iPSCs) and to form teratomas (human iPSCs). In these experiments, Valproic Acid (VPA)-mediated inhibition of Hdac2 was required for mouse, but not human, miRNA-mediated reprogramming. The miR-302-367 cluster contains four miR-302 members, with the AAGUGC seed, and the unrelated miR-367. Both were required for reprogramming, as miR-302 alone is not able to give rise to iPSCs in the absence of miR-367. Interestingly, when used in combination with reprogramming factors to enhance reprogramming, miR-302 alone was almost as effective as the intact miR-302-367 cluster, whereas miR-367 alone had no effect [121]. Moreover, miR-367, but not miR-302, can be substituted by other miRNAs in alternative reprogramming cocktails. For instance, it has been shown that the combination of miR-302, miR-200c and miR-369 can reprogram both human and mouse somatic cells [128]. In this case, instead of delivering miRNA genes via viral vectors, repeated transfections of mature synthetic miRNAs were used. This approach reduces the reprogramming efficiency but might be useful towards therapeutic applications of iPSCs, which would require cells devoid of exogenous reprogramming genes integration in the genome [18]. In contrast to the finding that combinations of multiple miRNA families are necessary for reprogramming [123,127,128], Lin *et al.* have reported that miR-302 alone could convert skin cancer and hair follicle cells into iPSCs in the absence of other miRNAs or RFs [129,130]. It remains unclear whether other somatic cell types can be reprogrammed under these conditions.

3.3. miRNAs Regulate Reprogramming by Multiple Pathways

The mechanisms underlying miRNA-mediated reprogramming have been only partially clarified. As mentioned before, introduction of miR-302 family members in DGCR8 $-/-$ mESCs increases expression of endogenous c-Myc and N-Myc downstream genes [110]. However, it is unlikely that the

pro-reprogramming activity of the miR-302 is merely mediated by downstream activation of c-Myc. In fact, the mechanisms by which these miRNAs and c-Myc enhance the reprogramming efficiency seem different, since the miRNAs, unlike c-Myc, do not accelerate doubling time of MEFs and produce a more homogenous population of fully reprogrammed cells [120]. Multiple pathways may be affected by overexpression of these miRNAs. For instance, it has been proposed that during reprogramming miR-302 regulates multiple genes involved in cell cycle regulation, epigenetic regulation, vesicular transport, cell signaling and mesenchymal-to-epithelial transition [123].

Other miRNAs that affect reprogramming are integrated in the p53 pathway. p53 activation elicits apoptosis and cell cycle arrest, representing a roadblock to reprogramming. Consistently, p53 inhibition promotes reprogramming [131]. The anti-reprogramming miR-21 and miR-29a sustain p53 activity. As in the case of let-7, c-Myc negatively regulates these miRNAs [125]. p53 positively regulates several anti-reprogramming miRNAs: it promotes transcription of miR-34a and miR-145, which target pluripotency genes [132], and upregulates at the post-transcriptional level miR-199a-3p that inhibits cell proliferation [126]. Conversely, miR-138 enhances reprogramming by directly targeting p53 [119,122].

A mesenchymal-to-epithelial transition (MET) is one of the early events occurring during fibroblasts reprogramming to iPSCs [133,134]. Since iPSCs have an epithelial character, MET is necessary for reprogramming of mesenchymal cells, such as fibroblasts, but not for keratinocytes and other somatic cells with an epithelial character. MET is under the control of the TGF β signaling and several reprogramming factors and miRNAs are involved in this process [135]. In particular, TGF β promotes expression of Snail, a mediator of epithelial-to-mesenchymal transition (EMT), whereas Sox2 and Oct4 inhibit Snail transcription. c-Myc downregulates TGF β 1 and TGF β receptor 2 (TGFBR2), which is also a target of the miR-302 and miR-93 families [119,121,123,133]. Conversely, BMP signaling stimulates expression of the miR-200 family and miR-205, which repress other EMT factors, ZEB1 and SIP1 (ZEB2) [133,134,136]. In turn, ZEB1 can inhibit transcription of the miR-200 family in a negative feedback loop [137,138]. ZEB2 is also a predicted target of miR-369 that, together with miR-200c, is a component of a miRNA-only reprogramming cocktail [128]. Another gene involved in EMT, RHOC, is a direct target of miR-302 [123].

miR-302 overexpression sustains pluripotency markers in differentiating hESCs [87,139]. Similarly, during reprogramming, miR-302 may also indirectly promote the activation of endogenous core pluripotency genes by targeting their inhibitors. For instance, methyl-DNA binding domain protein 2 (MBD2), an epigenetic suppressor of Nanog, is a direct target of miR-302 [140]. Downregulation of MBD2 by miR-302 is necessary to achieve a fully reprogrammed iPSC state. We have previously demonstrated that miR-302 targets NR2F2, which in turn is a transcriptional inhibitor of Oct4 [94]. Recently, it has been shown that NR2F2 knockdown enhances reprogramming efficiency, thus mimicking miR-302 overexpression [141].

4. Long Non-Coding RNAs in Embryonic Stem Cells

A number of differentially expressed lncRNAs were detected in undifferentiated mESCs and upon induction of differentiation by microarray analysis [37]. They could be classified as pluripotency, early mesoderm, and hematopoietic lncRNAs. New intergenic transcripts can be predicted by taking

advantage of chromatin immunoprecipitation followed by deep sequencing (ChIP-Seq). Genes actively transcribed by Pol II are characterized by a distinctive chromatin signature, consisting in trimethylation of lysine 4 of histone H3 in the promoter combined with trimethylation of lysine 36 of histone H3 in the transcribed region (K4-K36 domain). A K4-K36 domain located outside of known protein-coding loci would predict a putative novel long intergenic ncRNA. In mESCs and somatic cells, this approach led to the identification of over a thousand of novel lincRNAs that were further validated by microarray and northern blot [142]. Bioinformatics analysis showed that both the promoters and the transcribed sequences of these novel lincRNAs are conserved in mammals, suggesting that they may have biological functions. More recently, the number of mESC lincRNAs has been further expanded by taking advantage of a computational method that allows the reconstruction of the whole transcriptome (Scripture) from massive cDNA sequencing (RNA-Seq) data [38]. Overall, lincRNAs expressed in ESC show high evolutionary conservation and tissue specificity, and their expression levels are generally comparable to protein coding genes. Interestingly, Scripture identified hundreds of noncoding RNAs that are partially overlapped with protein-coding genes but transcribed in the opposite orientation. Such antisense transcripts are less conserved than long intergenic ncRNAs and likely represent a distinct class [38]. Analysis of intergenic K4-K36 domains indicated that lincRNAs are present also in human ESCs [47].

Many mESC lincRNA genes are regulated by core transcription factors that bind directly their promoters [142]. In particular, about 10% of Nanog and Oct4 binding sites are associated with lincRNA genes [143]. Knockdown of individual core TFs affects the expression of about 60% of lincRNAs [144]. As expected, mESC lincRNAs levels decrease upon differentiation, as the levels of mESC TFs decline. Also for some human ESC lincRNAs there is evidence for a direct regulation by core TFs [145,146].

lincRNAs Maintain Pluripotency in ESC

A biological role for mESC lincRNAs, initially postulated on the basis of their evolutionary conservation, has been recently experimentally demonstrated. RNAi against the lincRNA AK028326 (also known as Gomafu/Miat) resulted in morphological differentiation, reduced proliferation, decreased levels of Oct4 and other pluripotency markers and increase of trophoblast markers [143], suggesting that this lincRNA has a role in maintaining mESC pluripotency. AK028326 is activated by Oct4, establishing a positive feedback loop. RNAi against another mESC lincRNA, AK141205, also resulted in a decrease of Oct4, but in this case Nanog levels, morphology and proliferation were unaffected and differentiation genes were not induced [143]. Regulation of Oct4 by lincRNAs is conserved in human. Detailed analysis of transcripts generated by the Oct4 locus and Oct4 pseudogenes loci in human breast adenocarcinoma cells (MCF-7) has shown the existence of two transcripts, which were antisense to Oct4 (asOct4) and Oct4 pseudogene 5 (asOct4-pg5) [147]. It has been proposed that asOct4-pg5 is a functional lincRNA in MCF-7 cells, as it is able to recruit chromatin modifiers to the Oct4 promoter, leading to transcriptional silencing of the Oct4 gene by methylation of H3K27 and H3K9 by Ezh2 and G9a, respectively. It would be interesting to assess whether this regulatory mechanism has also a role in human pluripotent stem cells, which rely on Oct4 for the maintenance of pluripotency. Recent work has shown that three lincRNAs, lincRNA_ES1, lincRNA_ES2 and

lincRNA_ES3, were specifically expressed in undifferentiated hESCs, and absent in somatic tissues [145]. For two of them, direct regulation by Oct4 and Nanog has been proposed. Functional analysis showed that the three ES lincRNAs are necessary to maintain pluripotency in hESCs, as their individual knockdown by RNAi resulted in downregulation of pluripotency markers and induction of lineage genes. Nuclear localization and association with PRC2 and Sox2 suggest that these lincRNAs may function by regulating gene expression at the transcriptional level [145].

Large-scale analysis suggests that the role of lincRNAs in pluripotent stem cells may be greater than expected. Loss-of-function by RNA interference was systematically achieved for 147 mESC lincRNAs [144]. Strikingly, for nearly all lincRNAs tested (93%), this produced a significant impact on global gene expression. On average, the number of genes affected by lincRNA knockdown was comparable to that obtained by the knockdown of previously characterized key ESC regulatory factors, suggesting that lincRNAs play major roles in mESCs. This was confirmed by further analysis, showing that knockdown of a subset of lincRNAs affected expression of pluripotency markers, including Oct4, Sox2 and Nanog, and led to a change in mESC morphology. Moreover, loss of function of many lincRNAs triggered the activation of specific neuroectoderm, endoderm, mesoderm, or trophectoderm genes [144]. These results indicate that lincRNAs are necessary for two crucial aspects of mESC pluripotency: the maintenance of the specific genetic program associated with the undifferentiated state, and the repression of genes involved in lineage differentiation.

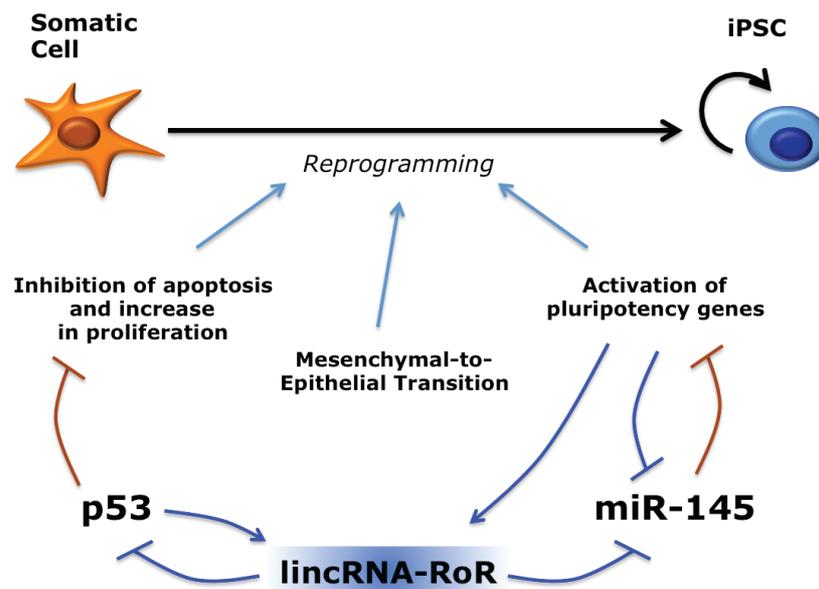
Mechanistically, mESC lincRNAs may exert their function by modulating the activity of chromatin-modifying factors. Both “readers”, “writers” and “erasers” histone modifiers and the chromatin-associated DNA binding protein Yy1 have been found in association with mESCs lincRNAs [37,144]. For example, the lincRNAs *Evx1as* and *Hoxb5/6as*, which are transcribed antisense to the homeotic genes *Evx1* and *Hoxb5-Hoxb6*, are associated with the Trithorax protein MLL1. Interestingly, many lincRNAs are bound by multiple chromatin regulatory complexes, which form consistent combinations of readers, writers and erasers. This evidence suggests that in mESCs lincRNAs may function as a bridge to tether distinct functionally related complexes, as proposed in other systems [148]. Such molecular scaffolding function may play a crucial role in maintaining the epigenetic state of pluripotent stem cells.

5. Long Non-Coding RNAs in Induced Pluripotent Stem Cells

A subset of human lincRNAs demonstrated expression at higher levels in iPSCs with respect to both somatic cells of origin and hESCs [146]. Among them, there is at least one with a role in reprogramming, named lincRNA-RoR (or linc-RoR) for “Regulator of Reprogramming” (Figure 7). Knockdown of linc-RoR resulted in a significant decrease of iPSC colonies formation, whereas overexpression enhanced reprogramming efficiency. This pro-reprogramming effect can be explained by recent work showing that linc-RoR is a negative regulator of p53 [149]. Mechanistically, linc-RoR inhibits translation of p53 mRNA in the cytoplasm and this inhibitory activity is dependent on the interaction with heterogeneous nuclear ribonucleoprotein I (hnRNP I). Interestingly, linc-RoR transcription is induced by p53, establishing an autoregulatory feedback loop [149]. Another function recently proposed for linc-RoR may underlie its pro-reprogramming activity [150]. By sequestering a

set of miRNAs that target Oct4, Nanog and Sox2, Linc-RoR may protect pluripotency factors from miRNA inhibition by acting as a competing endogenous RNA (ceRNA) [52].

Figure 7. Role of lincRNA-RoR during reprogramming. LincRNA-RoR may promote reprogramming by two different mechanisms. It inhibits p53, which in turn negatively affects the initial steps of reprogramming by inducing apoptosis. It also acts as a competing endogenous RNA and releases from repression the endogenous pluripotency genes Oct4, Sox2 and Klf4, which are targeted by miR-145. MiR-145 is also inhibited by Oct4 in a negative feedback loop. Other miRNA may be “sponged” by lincRNA-RoR (not shown).



6. Conclusions and Future Perspectives

In ESCs and iPSCs, pluripotency is the result of: (i) the interplay between multiple signaling pathways; (ii) a core transcriptional regulatory circuitry; (iii) epigenetic control of gene expression by specific chromatin modification patterns; (iv) the maintenance of a peculiar cell cycle. Recent work suggests that regulatory ncRNAs have a role in modulating all these biological pathways. Similar to the core set of transcription factors, ESCs are characterized by a core set of miRNAs, belonging to the miR-302 family. These miRNAs are hierarchically on top of multiple targets, which participate in the maintenance of self-renewal, and regulate exit from pluripotency and cell fate choice. Moreover, miR-302 family members are required for efficient reprogramming of somatic cells and, in combination with other miRNAs, are sufficient for iPSC generation in the absence of canonical reprogramming factors. A similar role has been shown for other miRNAs in different systems. For instance, neural-specific miRNAs can induce trans-differentiation of fibroblasts into induced neurons without other factors [151]. Since miRNAs can be provided as mature RNA molecules, this approach could lead to the generation of therapeutic relevant cells devoid of exogenous genes integrations and genomic lesions.

The functions of lncRNAs in pluripotent cells are less characterized. So far, large datasets generated by RNA-seq projects have been followed only by a limited number of functional studies. Such are hampered by the fact that unlike protein-coding genes and miRNA the function of a lncRNA cannot be normally predicted by the analysis of its sequence [152]. However, the few examples in which the

function of lncRNAs has been investigated suggest that they might be primary regulators of pluripotency, differentiation and reprogramming. In particular, since lncRNAs have been show to act as molecular scaffolds of chromatin regulators, they may have a huge impact in the context of stem cells where the maintenance of a characteristic epigenetic state of the chromatin is central for pluripotency and differentiation. So far, ESC lncRNAs have been associated with maintenance of pluripotency factors, repression of differentiation, and inhibition of apoptosis. Their possible role in other crucial pathways underlying pluripotency has not yet been investigated. For instance, it is unknown whether lncRNAs could be involved in tuning signaling pathways, whereas evidence exists for miRNAs. Finally, as cytoplasmic lncRNAs may work as inhibitors of miRNAs, they could add another layer of complexity to the regulatory circuitries underlying pluripotency.

Acknowledgments

We are grateful to A. Fatica, V. de Turris, J. Lenzi, R. De Santis and M. Heke for critical reading of the manuscript. A.H.B. is supported by NIH grant #5R01GM097615.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Warmflash, A.; Arduini, B.L.; Brivanlou, A.H. The molecular circuitry underlying pluripotency in embryonic stem cells. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2012**, *4*, 443–456.
2. James, D.; Levine, A.J.; Besser, D.; Hemmati-Brivanlou, A. TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development* **2005**, *132*, 1273–1282.
3. Vallier, L.; Alexander, M.; Pedersen, R.A. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J. Cell Sci.* **2005**, *118*, 4495–4509.
4. Ying, Q.L.; Nichols, J.; Chambers, I.; Smith, A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **2003**, *115*, 281–292.
5. Boyer, L.A.; Lee, T.I.; Cole, M.F.; Johnstone, S.E.; Levine, S.S.; Zucker, J.P.; Guenther, M.G.; Kumar, R.M.; Murray, H.L.; Jenner, R.G.; *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **2005**, *122*, 947–956.
6. Kim, J.; Chu, J.; Shen, X.; Wang, J.; Orkin, S.H. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **2008**, *132*, 1049–1061.
7. Chen, X.; Xu, H.; Yuan, P.; Fang, F.; Huss, M.; Vega, V.B.; Wong, E.; Orlov, Y.L.; Zhang, W.; Jiang, J.; *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **2008**, *133*, 1106–1117.
8. Boyer, L.A.; Plath, K.; Zeitlinger, J.; Brambrink, T.; Medeiros, L.A.; Lee, T.I.; Levine, S.S.; Wernig, M.; Tajonar, A.; Ray, M.K.; *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **2006**, *441*, 349–353.

9. 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.
10. Spivakov, M.; Fisher, A.G. Epigenetic signatures of stem-cell identity. *Nat. Rev. Genet.* **2007**, *8*, 263–271.
11. Nichols, J.; Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **2009**, *4*, 487–492.
12. Guo, G.; Yang, J.; Nichols, J.; Hall, J.S.; Eyres, I.; Mansfield, W.; Smith, A. Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* **2009**, *136*, 1063–1069.
13. Silva, J.; Nichols, J.; Theunissen, T.W.; Guo, G.; van Oosten, A.L.; Barrandon, O.; Wray, J.; Yamanaka, S.; Chambers, I.; Smith, A. Nanog is the gateway to the pluripotent ground state. *Cell* **2009**, *138*, 722–737.
14. Pera, M.F.; Tam, P.P.L. Extrinsic regulation of pluripotent stem cells. *Nature* **2010**, *465*, 713–720.
15. Takahashi, K.; Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **2006**, *126*, 663–676.
16. Takahashi, K.; Tanabe, K.; Ohnuki, M.; Narita, M.; Ichisaka, T.; Tomoda, K.; Yamanaka, S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **2007**, *131*, 861–872.
17. Yu, J.; Vodyanik, M.A.; Smuga-Otto, K.; Antosiewicz-Bourget, J.; Frane, J.L.; Tian, S.; Nie, J.; Jonsdottir, G.A.; Ruotti, V.; Stewart, R.; *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **2007**, *318*, 1917–1920.
18. Stadtfeld, M.; Hochedlinger, K. Induced pluripotency: History, mechanisms, and applications. *Genes Dev.* **2010**, *24*, 2239–2263.
19. Papp, B.; Plath, K. Reprogramming to pluripotency: Stepwise resetting of the epigenetic landscape. *Cell Res.* **2011**, *21*, 486–501.
20. Cohen, D.E.; Melton, D. Turning straw into gold: Directing cell fate for regenerative medicine. *Nat. Rev. Genet.* **2011**, *12*, 243–252.
21. Mattick, J.S. The central role of RNA in human development and cognition. *FEBS Lett.* **2011**, *585*, 1600–1616.
22. Bartel, D.P. MicroRNAs: Target recognition and regulatory functions. *Cell* **2009**, *136*, 215–233.
23. Bushati, N.; Cohen, S.M. microRNA functions. *Annu. Rev. Cell Dev. Biol.* **2007**, *23*, 175–205.
24. Fatica, A.; Rosa, A.; Ballarino, M.; de Marchis, M.L.; Rasmussen, K.D.; Bozzoni, I. Role of microRNAs in myeloid differentiation. *Biochem. Soc. Trans.* **2008**, *36*, 1201–1205.
25. Rosa, A.; Brivanlou, A.H. microRNAs in early vertebrate development. *Cell Cycle* **2009**, *8*, 3531–3520.
26. Rottiers, V.; Näär, A.M. MicroRNAs in metabolism and metabolic disorders. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 239–250.
27. Iorio, M.V.; Croce, C.M. MicroRNA dysregulation in cancer: Diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO Mol. Med.* **2012**, *4*, 143–159.
28. Denli, A.M.; Tops, B.B.J.; Plasterk, R.H.A.; Ketting, R.E.F.; Hannon, G.J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **2004**, *432*, 231–235.

29. 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.
30. Hutvagner, G.; Zamore, P.D. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **2002**, *297*, 2056–2060.
31. Hammond, S.M. Dicing and slicing: The core machinery of the RNA interference pathway. *FEBS Lett.* **2005**, *579*, 5822–5829.
32. Gregory, R.I.; Chendrimada, T.P.; Cooch, N.; Shiekhattar, R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **2005**, *123*, 631–640.
33. Rajewsky, N. microRNA target predictions in animals. *Nat. Genet.* **2006**, *38*, S8–S13.
34. 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.
35. 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.
36. Sone, M.; Hayashi, T.; Tarui, H.; Agata, K.; Takeichi, M.; Nakagawa, S. The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. *J. Cell Sci.* **2007**, *120*, 2498–2506.
37. Dinger, M.E.; Amaral, P.P.; Mercer, T.R.; Pang, K.C.; Bruce, S.J.; Gardiner, B.B.; Askarian-Amiri, M.E.; Ru, K.; Soldà, G.; Simons, C.; *et al.* Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res.* **2008**, *18*, 1433–1445.
38. 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.
39. Kapranov, P.; Cheng, J.; Dike, S.; Nix, D.A.; Duttagupta, R.; Willingham, A.T.; Stadler, P.F.; Hertel, J.; Hackermüller, J.; Hofacker, I.L.; *et al.* RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **2007**, *316*, 1484–1488.
40. Nakaya, H.I.; Amaral, P.P.; Louro, R.; Lopes, A.; Fachel, A.A.; Moreira, Y.B.; El-Jundi, T.A.; da Silva, A.M.; Reis, E.M.; Verjovski-Almeida, S. Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. *Genome Biol.* **2007**, *8*, R43.
41. Mercer, T.R.; Dinger, M.E.; Sunken, 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.
42. 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.
43. Ponjavic, J.; Ponting, C.P.; Lunter, G. Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. *Genome Res.* **2007**, *17*, 556–565.
44. 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.
45. Mattick, J.S. A new paradigm for developmental biology. *J. Exp. Biol.* **2007**, *210*, 1526–1547.

46. Carninci, P.; Hayashizaki, Y. Noncoding RNA transcription beyond annotated genes. *Curr. Opin. Genet. Dev.* **2007**, *17*, 139–144.
47. 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.
48. Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914.
49. 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.
50. Beltran, M.; Puig, I.; Peña, C.; García, J.M.; Alvarez, A.B.; Peña, 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.
51. 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.
52. 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.
53. 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.
54. Houbaviy, H.B.; Murray, M.F.; Sharp, P.A. Embryonic stem cell-specific microRNAs. *Dev. Cell* **2003**, *5*, 351–358.
55. Laurent, L.C.; Chen, J.; Ulitsky, I.; Mueller, F.-J.; Lu, C.; Shamir, R.; Fan, J.-B.; Loring, J.F. Comprehensive microRNA profiling reveals a unique human embryonic stem cell signature dominated by a single seed sequence. *Stem Cells* **2008**, *26*, 1506–1516.
56. Bar, M.; Wyman, S.K.; Fritz, B.R.; Qi, J.; Garg, K.S.; Parkin, R.K.; Kroh, E.M.; Bendoraitis, A.; Mitchell, P.S.; Nelson, A.M.; *et al.* MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. *Stem Cells* **2008**, *26*, 2496–2505.
57. 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.
58. Ciaudo, C.; Servant, N.; Cognat, V.; Sarazin, A.; Kieffer, E.; Viville, S.; Colot, V.; Barillot, E.; Heard, E.; Voinnet, O. Highly dynamic and sex-specific expression of microRNAs during early ES cell differentiation. *PLoS Genet.* **2009**, *5*, e1000620.
59. 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.
60. Calabrese, J.M.; Seila, A.C.; Yeo, G.W.; Sharp, P.A. RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 18097–18102.

61. Lipchina, I.; Elkabetz, Y.; Hafner, M.; Sheridan, R.; Mihailovic, A.; Tuschl, T.; Sander, C.; Studer, L.; Betel, D. Genome-wide identification of microRNA targets in human ES cells reveals a role for miR-302 in modulating BMP response. *Genes Dev.* **2011**, *25*, 2173–2186.
62. Barroso-delJesus, A.; Romero-López, C.; Lucena-Aguilar, G.; Melen, G.J.; Sanchez, L.; Ligeró, G.; Berzal-Herranz, A.; Menendez, P. Embryonic stem cell-specific miR302-367 cluster: Human gene structure and functional characterization of its core promoter. *Mol. Cell Biol.* **2008**, *28*, 6609–6619.
63. Card, D.A.G.; Hebbar, P.B.; Li, L.; Trotter, K.W.; Komatsu, Y.; Mishina, Y.; Archer, T.K. Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. *Mol. Cell Biol.* **2008**, *28*, 6426–6438.
64. Jouneau, A.; Ciaudo, C.; Sismeiro, O.; Brochard, V.; Jouneau, L.; Vandormael-Pournin, S.; Coppée, J.-Y.; Zhou, Q.; Heard, E.; Antoniewski, C.; *et al.* Naive and primed murine pluripotent stem cells have distinct miRNA expression profiles. *RNA* **2012**, *18*, 253–264.
65. Stadler, B.; Ivanovska, I.; Mehta, K.; Song, S.; Nelson, A.; Tan, Y.; Mathieu, J.; Darby, C.; Blau, C.A.; Ware, C.; *et al.* Characterization of microRNAs involved in embryonic stem cell states. *Stem Cells Dev.* **2010**, *19*, 935–950.
66. Hanna, J.; Cheng, A.W.; Saha, K.; Kim, J.; Lengner, C.J.; Soldner, F.; Cassady, J.P.; Muffat, J.; Carey, B.W.; Jaenisch, R. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 9222–9227.
67. Kim, H.; Lee, G.; Ganat, Y.; Papapetrou, E.P.; Lipchina, I.; Socci, N.D.; Sadelain, M.; Studer, L. miR-371-3 expression predicts neural differentiation propensity in human pluripotent stem cells. *Cell Stem Cell* **2011**, *8*, 695–706.
68. 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.
69. Kanellopoulou, C.; Muljo, S.A.; Kung, A.L.; Ganesan, S.; Drapkin, R.; Jenuwein, T.; Livingston, D.M.; Rajewsky, K. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev.* **2005**, *19*, 489–501.
70. Murchison, E.P.; Partridge, J.F.; Tam, O.H.; Cheloufi, S.; Hannon, G.J. Characterization of Dicer-deficient murine embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12135–12140.
71. Wang, Y.; Medvid, R.; Melton, C.; Jaenisch, R.; Blelloch, R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat. Genet.* **2007**, *39*, 380–385.
72. Sinkkonen, L.; Hugenschmidt, T.; Berninger, P.; Gaidatzis, D.; Mohn, F.; Artus-Revel, C.G.; Zavolan, M.; Svoboda, P.; Filipowicz, W. microRNAs control de novo DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.* **2008**, *15*, 259–267.
73. Wang, Y.; Baskerville, S.; Shenoy, A.; Babiarz, J.E.; Baehner, L.; Blelloch, R. Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. *Nat. Genet.* **2008**, *40*, 1478–1483.
74. Wang, Y.; Blelloch, R. Cell cycle regulation by microRNAs in embryonic stem cells. *Cancer Res.* **2009**, *69*, 4093–4096.

75. Benetti, R.; Gonzalo, S.; Jaco, I.; Muñoz, P.; Gonzalez, S.; Schoeftner, S.; Murchison, E.; Andl, T.; Chen, T.; Klatt, P.; *et al.* A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. *Nat. Struct. Mol. Biol.* **2008**, *15*, 268–279.
76. Kim, B.-M.; Choi, M.Y. Non-canonical microRNAs miR-320 and miR-702 promote proliferation in Dgcr8-deficient embryonic stem cells. *Biochem. Biophys. Res. Commun.* **2012**, *426*, 183–189.
77. Babiarz, J.E.; Ruby, J.G.; Wang, Y.; Bartel, D.P.; Bilello, R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev.* **2008**, *22*, 2773–2785.
78. White, J.; Dalton, S. Cell cycle control of embryonic stem cells. *Stem Cell Rev.* **2005**, *1*, 131–138.
79. Qi, J.; Yu, J.-Y.; Shcherbata, H.R.; Mathieu, J.; Wang, A.J.; Seal, S.; Zhou, W.; Stadler, B.M.; Bourgin, D.; Wang, L.; *et al.* microRNAs regulate human embryonic stem cell division. *Cell Cycle* **2009**, *8*, 3729–3741.
80. Dolezalova, D.; Mraz, M.; Barta, T.; Plevova, K.; Vinarsky, V.; Holubcova, Z.; Jaros, J.; Dvorak, P.; Pospisilova, S.; Hampl, A. microRNAs regulate p21(Waf1/Cip1) protein expression and the DNA damage response in human embryonic stem cells. *Stem Cells* **2012**, *30*, 1362–1372.
81. Sengupta, S.; Nie, J.; Wagner, R.J.; Yang, C.; Stewart, R.; Thomson, J.A. microRNA 92b controls the G1/S checkpoint gene p57 in human embryonic stem cells. *Stem Cells* **2009**, *27*, 1524–1528.
82. Lee, N.S.; Kim, J.S.; Cho, W.J.; Lee, M.R.; Steiner, R.; Gompers, A.; Ling, D.; Zhang, J.; Strom, P.; Behlke, M.; *et al.* miR-302b maintains “stemness” of human embryonal carcinoma cells by post-transcriptional regulation of Cyclin D2 expression. *Biochem. Biophys. Res. Commun.* **2008**, *377*, 434–440.
83. Hafner, M.; Landthaler, M.; Burger, L.; Khorshid, M.; Hausser, J.; Berninger, P.; Rothballer, A.; Ascano, M.; 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.
84. Lipchina, I.; Studer, L.; Betel, D. The expanding role of miR-302-367 in pluripotency and reprogramming. *Cell Cycle* **2012**, *11*, 1517–1523.
85. Smith, J.R.; Vallier, L.; Lupo, G.; Alexander, M.; Harris, W.A.; Pedersen, R.A. Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm. *Dev. Biol.* **2008**, *313*, 107–117.
86. Vallier, L.; Reynolds, D.; Pedersen, R.A. Nodal inhibits differentiation of human embryonic stem cells along the neuroectodermal default pathway. *Dev. Biol.* **2004**, *275*, 403–421.
87. Rosa, A.; Spagnoli, F.M.; Brivanlou, A.H. The miR-430/427/302 family controls mesendodermal fate specification via species-specific target selection. *Dev. Cell* **2009**, *16*, 517–527.
88. Choi, W.-Y.; Giraldez, A.J.; Schier, A.F. Target protectors reveal dampening and balancing of Nodal agonist and antagonist by miR-430. *Science* **2007**, *318*, 271–274.
89. Chambers, S.M.; Fasano, C.A.; Papapetrou, E.P.; Tomishima, M.; Sadelain, M.; Studer, L. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat. Biotechnol.* **2009**, *27*, 275–280.

90. Ozair, M.Z.; Noggle, S.; Warmflash, A.; Krzyspiak, J.E.; Brivanlou, A.H. SMAD7 Directly converts human embryonic stem cells to telencephalic fate by a default mechanism. *Stem Cells* **2013**, *31*, 35–47.
91. Kang, H.; Louie, J.; Weisman, A.; Sheu-Gruttadauria, J.; Davis-Dusenbery, B.N.; Lagna, G.; Hata, A. Inhibition of microRNA-302 (miR-302) by bone morphogenetic protein 4 (BMP4) facilitates the BMP signaling pathway. *J. Biol. Chem.* **2012**, *287*, 38656–38664.
92. Parisi, S.; Battista, M.; Musto, A.; Navarra, A.; Tarantino, C.; Russo, T. A regulatory loop involving *Dies1* and miR-125a controls BMP4 signaling in mouse embryonic stem cells. *FASEB J.* **2012**, *26*, 3957–3968.
93. Zhou, A.-D.; Diao, L.-T.; Xu, H.; Xiao, Z.-D.; Li, J.-H.; Zhou, H.; Qu, L.-H. β -Catenin/LEF1 transactivates the microRNA-371-373 cluster that modulates the Wnt/ β -catenin-signaling pathway. *Oncogene* **2012**, *31*, 2968–2978.
94. Rosa, A.; Brivanlou, A.H. A regulatory circuitry comprised of miR-302 and the transcription factors OCT4 and NR2F2 regulates human embryonic stem cell differentiation. *EMBO J.* **2010**, *30*, 237–248.
95. 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.
96. Wellner, U.; Schubert, J.; Burk, U.C.; Schmalhofer, O.; Zhu, F.; Sonntag, A.; Waldvogel, B.; Vannier, C.; Darling, D.; zur Hausen, A.; *et al.* The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nat. Cell Biol.* **2009**, *11*, 1487–1495.
97. Yi, R.; Poy, M.N.; Stoffel, M.; Fuchs, E. A skin microRNA promotes differentiation by repressing “stemness”. *Nature* **2008**, *452*, 225–229.
98. Xu, N.; Papagiannakopoulos, T.; Pan, G.; Thomson, J.; Kosik, K. microRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell* **2009**, *137*, 647–658.
99. Morey, L.; Pascual, G.; Cozzuto, L.; Roma, G.; Wutz, A.; Benitah, S.A.; Di Croce, L. Nonoverlapping functions of the polycomb group cbx family of proteins in embryonic stem cells. *Cell Stem Cell* **2012**, *10*, 47–62.
100. O’Loghlen, A.; Muñoz-Cabello, A.M.; Gaspar-Maia, A.; Wu, H.-A.; Banito, A.; Kunowska, N.; Racek, T.; Pemberton, H.N.; Beolchi, P.; Laval, F.; *et al.* MicroRNA regulation of *Cbx7* mediates a switch of polycomb orthologs during ESC differentiation. *Cell Stem Cell* **2012**, *10*, 33–46.
101. Ambros, V. The evolution of our thinking about microRNAs. *Nat. Med.* **2008**, *14*, 1036–1040.
102. Büssing, I.; Slack, F.J.; Großhans, H. let-7 microRNAs in development, stem cells and cancer. *Trends Mol. Med.* **2008**, *14*, 400–409.
103. 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.
104. Viswanathan, S.R.; Daley, G.Q.; Gregory, R.I. Selective blockade of microRNA processing by *Lin28*. *Science* **2008**, *320*, 97–100.
105. Viswanathan, S.R.; Daley, G.Q. *Lin28*: A microRNA regulator with a macro role. *Cell* **2010**, *140*, 445–449.

106. 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.
107. 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.
108. 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.
109. Rybak, A.; Fuchs, H.; Smirnova, L.; Brandt, C.; Pohl, E.E.; Nitsch, R.; Wulczyn, F.G. 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.
110. Melton, C.; Judson, R.L.; Blelloch, R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature* **2010**, *463*, 621–626.
111. Cartwright, P.; McLean, C.; Sheppard, A.; Rivett, D.; Jones, K.; Dalton, S. LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development* **2005**, *132*, 885–896.
112. Lin, C.-H.; Jackson, A.L.; Guo, J.; Linsley, P.S.; Eisenman, R.N. Myc-regulated microRNAs attenuate embryonic stem cell differentiation. *EMBO J.* **2009**, *28*, 3157–3170.
113. Chang, T.-C.; Yu, D.; Lee, Y.-S.; Wentzel, E.A.; Arking, D.E.; West, K.M.; Dang, C.V.; Thomas-Tikhonenko, A.; Mendell, J.T. Widespread microRNA repression by Myc contributes to tumorigenesis. *Nat. Genet.* **2008**, *40*, 43–50.
114. Chin, M.H.; Mason, M.J.; Xie, W.; Volinia, S.; Singer, M.; Peterson, C.; Ambartsumyan, G.; Aimiwu, O.; Richter, L.; Zhang, J.; *et al.* Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* **2009**, *5*, 111–123.
115. Wilson, K.D.; Venkatasubrahmanyam, S.; Jia, F.; Sun, N.; Butte, A.J.; Wu, J.C. microRNA profiling of human-induced pluripotent stem cells. *Stem Cells Dev.* **2009**, *18*, 749–757.
116. Neveu, P.; Kye, M.J.; Qi, S.; Buchholz, D.E.; Clegg, D.O.; Sahin, M.; Park, I.-H.; Kim, K.-S.; Daley, G.Q.; Kornblum, H.I.; *et al.* MicroRNA profiling reveals two distinct p53-related human pluripotent stem cell states. *Cell Stem Cell* **2010**, *7*, 671–681.
117. Stadtfeld, M.; Apostolou, E.; Akutsu, H.; Fukuda, A.; Follett, P.; Natesan, S.; Kono, T.; Shioda, T.; Hochedlinger, K. Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* **2010**, *465*, 175–181.
118. Peters, J.; Robson, J.E. Imprinted noncoding RNAs. *Mamm. Genome* **2008**, *19*, 493–502.
119. Li, Z.; Yang, C.-S.; Nakashima, K.; Rana, T.M. Small RNA-mediated regulation of iPS cell generation. *EMBO J.* **2011**, *30*, 823–834.
120. Judson, R.; Babiarz, J.; Venere, M.; Blelloch, R. Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nat. Biotechnol.* **2009**, *27*, 459–461.
121. Liao, B.; Bao, X.; Liu, L.; Feng, S.; Zovoilis, A.; Liu, W.; Xue, Y.; Cai, J.; Guo, X.; Qin, B.; *et al.* MicroRNA cluster 302-367 enhances somatic cell reprogramming by accelerating a mesenchymal-to-epithelial transition. *J. Biol. Chem.* **2011**, *286*, 17359–17364.
122. Ye, D.; Wang, G.; Liu, Y.; Huang, W.; Wu, M.; Zhu, S.; Jia, W.; Deng, A.-M.; Liu, H.; Kang, J. MiR-138 promotes induced pluripotent stem cell generation through the regulation of the p53 signaling. *Stem Cells* **2012**, *30*, 1645–1654.

123. Subramanyam, D.; Lamouille, S.; Judson, R.L.; Liu, J.Y.; Bucay, N.; Derynck, R.; Blecloch, R. Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nat. Biotechnol.* **2011**, *29*, 443–448.
124. Hanna, J.; Saha, K.; Pando, B.; van Zon, J.; Lengner, C.J.; Creighton, M.P.; van Oudenaarden, A.; Jaenisch, R. Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **2009**, *462*, 595–601.
125. Yang, C.-S.; Li, Z.; Rana, T.M. microRNAs modulate iPS cell generation. *RNA* **2011**, *17*, 1451–1460.
126. 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.
127. Anokye-Danso, F.; Trivedi, C.M.; Juhr, D.; Gupta, M.; Cui, Z.; Tian, Y.; Zhang, Y.; Yang, W.; Gruber, P.J.; Epstein, J.A.; Morrissey, E.E. Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell* **2011**, *8*, 376–388.
128. Miyoshi, N.; Ishii, H.; Nagano, H.; Haraguchi, N.; Dewi, D.L.; Kano, Y.; Nishikawa, S.; Tanemura, M.; Mimori, K.; Tanaka, F.; *et al.* Reprogramming of mouse and human cells to pluripotency using mature microRNAs. *Cell Stem Cell* **2011**, *8*, 633–638.
129. Lin, S.-L.; Chang, D.C.; Chang-Lin, S.; Lin, C.-H.; Wu, D.T.S.; Chen, D.T.; Ying, S.-Y. Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state. *RNA* **2008**, *14*, 2115–2124.
130. Lin, S.-L.; Chang, D.C.; Lin, C.-H.; Ying, S.-Y.; Leu, D.; Wu, D.T.S. Regulation of somatic cell reprogramming through inducible mir-302 expression. *Nucleic Acids Res.* **2011**, *39*, 1054–1065.
131. Menendez, S.; Camus, S.; Izpisua Belmonte, J.C. p53: Guardian of reprogramming. *Cell Cycle* **2010**, *9*, 3887–3891.
132. Jain, A.K.; Allton, K.; Iacovino, M.; Mahen, E.; Milczarek, R.J.; Zwaka, T.P.; Kyba, M.; Barton, M.C. p53 regulates cell cycle and microRNAs to promote differentiation of human embryonic stem cells. *PLoS Biol.* **2012**, *10*, e1001268.
133. Li, R.; Liang, J.; Ni, S.; Zhou, T.; Qing, X.; Li, H.; He, W.; Chen, J.; Li, F.; Zhuang, Q.; *et al.* A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* **2010**, *7*, 51–63.
134. Samavarchi-Tehrani, P.; Golipour, A.; David, L.; Sung, H.-K.; Beyer, T.A.; Datti, A.; Woltjen, K.; Nagy, A.; Wrana, J.L. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* **2010**, *7*, 64–77.
135. Massagué, J. TGF β signalling in context. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 616–630.
136. Gregory, P.A.; Bert, A.G.; Paterson, E.L.; Barry, S.C.; Tsykin, A.; Farshid, G.; Vadas, M.A.; Khew-Goodall, Y.; Goodall, G.J. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat. Cell Biol.* **2008**, *10*, 593–601.
137. 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.
138. 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.

139. Barroso-delJesus, A.; Lucena-Aguilar, G.; Sanchez, L.; Ligerro, G.; Gutierrez-Aranda, I.; Menendez, P. The Nodal inhibitor Lefty is negatively modulated by the microRNA miR-302 in human embryonic stem cells. *FASEB J.* **2011**, *25*, 1497–1508.
140. Lee, M.R.; Prasain, N.; Chae, H.-D.; Kim, Y.-J.; Mantel, C.; Yoder, M.C.; Broxmeyer, H.E. Epigenetic regulation of nanog by MiR-302 cluster-MBD2 completes induced pluripotent stem cell reprogramming. *Stem Cells* **2012**, *31*, 666–681.
141. Hu, S.; Wilson, K.D.; Ghosh, Z.; Han, L.; Wang, Y.; Lan, F.; Ransohoff, K.J.; Burridge, P.; Wu, J.C. microRNA-302 increases reprogramming efficiency via repression of NR2F2. *Stem Cells* **2013**, *31*, 259–268.
142. 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.
143. 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.
144. 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.
145. Ng, S.-Y.; Johnson, R.; Stanton, L.W. Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. *EMBO J.* **2011**, *31*, 522–533.
146. 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.
147. 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.
148. Tsai, M.-C.; Manor, O.; Wan, Y.; Mosammamarast, 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.
149. 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.* **2012**, *23*, 340–350.
150. Wang, Y.; Xu, Z.; Jiang, J.; Xu, C.; Kang, J.; Xiao, L.; Wu, M.; Xiong, J.; Guo, X.; Liu, H. Endogenous miRNA sponge lincRNA-RoR regulates Oct4, nanog, and Sox2 in human embryonic stem cell self-renewal. *Dev. Cell* **2013**, *25*, 69–80.
151. Yoo, A.S.; Sun, A.X.; Li, L.; Shcheglovitov, A.; Portmann, T.; Li, Y.; Lee-Messer, C.; Dolmetsch, R.E.; Tsien, R.W.; Crabtree, G.R. microRNA-mediated conversion of human fibroblasts to neurons. *Nature* **2011**, *476*, 228–231.
152. Guttman, M.; Rinn, J.L. Modular regulatory principles of large non-coding RNAs. *Nature* **2012**, *482*, 339–346.

Reprinted from *IJMS*. Cite as: Battista, M.; Musto, A.; Navarra, A.; Minopoli, G.; Russo, T.; Parisi, S. miR-125b Regulates the Early Steps of ESC Differentiation through Dies1 in a TGF-Independent Manner. *Int. J. Mol. Sci.* **2013**, *14*, 13482-13496.

Article

miR-125b Regulates the Early Steps of ESC Differentiation through Dies1 in a TGF-Independent Manner

Marica Battista^{1,2,3}, Anna Musto^{1,2}, Angelica Navarra^{1,2}, Giuseppina Minopoli^{1,2}, Tommaso Russo^{1,2} and Silvia Parisi^{1,2,*}

¹ Department of Molecular Medicine and Medical Biotechnology, University of Naples, “Federico II”, Via Sergio Pansini 5, 80131 Naples, Italy;

E-Mails: battistam@ceinge.unina.it (M.B.); anna.musto@unina.it (A.M);
angelica.navarra@unina.it (A.N.); minopoli@dbbm.unina.it (G.M.);
tommaso.russo@unina.it (T.R.)

² Ceinge Biotechnologie Avanzate, Via Gaetano Salvatore 486, 80145 Naples, Italy

³ European School of Molecular Medicine (SEMM), 80145 Naples, Italy

* Author to whom correspondence should be addressed; E-Mail: silvia.parisi@unina.it;
Tel.: +39-08-1373-7878; Fax: +39-08-1373-7808.

Received: 30 May 2013; in revised form: 13 June 2013 / Accepted: 19 June 2013 /

Published: 27 June 2013

Abstract: Over the past few years, it has become evident that the distinctive pattern of miRNA expression seen in embryonic stem cells (ESCs) contributes to important signals in the choice of the cell fate. Thus, the identification of miRNAs and their targets, whose expression is linked to a specific step of differentiation, as well as the modulation of these miRNAs, may prove useful in the learning of how ESC potential is regulated. In this context, we have studied the expression profile of miRNAs during neural differentiation of ESCs. We have found that miR-125b is upregulated in the first steps of neural differentiation of ESCs. This miRNA targets the BMP4 co-receptor, Dies1, and, in turn, regulates the balance between BMP4 and Nodal/Activin signaling. The ectopic expression of miR-125b blocks ESC differentiation at the epiblast stage, and this arrest is rescued by restoring the expression of Dies1. Finally, opposite to miR-125a, whose expression is under the control of the BMP4, miR-125b is not directly regulated by Transforming Growth Factor beta (TGF β) signals. These results highlight a new important role of miR-125b in the regulation of the transition from ESCs to the epiblast stage and add a new level of control on TGF β signaling in ESCs.

Keywords: miRNAs; TGF β ; embryonic stem cells; differentiation; epiblast stem cells; signaling

1. Introduction

Embryonic stem cells (ESCs) have the unique characteristics of self-renewing in culture and generating a wide range of specialized cell types upon differentiation cues. For their characteristics, they represent an appealing system to study embryonic development and, more importantly, to obtain functional cells for replacement therapy. Of course, a precise and detailed knowledge of the gene programs regulating ESC fate is necessary to allow the use of these cells. Among the signals that regulate ESC fate, there are extrinsic signals, such as Leukemia Inhibitory Factor (LIF) and Transforming Growth Factor beta (TGF β). The LIF pathway is already well characterized, and it is required to maintain the undifferentiated state of ESCs [1]. TGF β pathways, *i.e.*, Bone Morphogenetic Protein 4 (BMP4) and Nodal/Activin, have multiple functions, both in stemness maintenance and during the differentiation [2–6], and they seem to work in a perfectly balanced way, so that a misregulation of one of these pathways has relevant consequences on the other ones [7,8]. Another important and widely studied class of regulators of ESC gene programs are the transcription factors (TFs), which have the important characteristic that one TF may control the expression of numerous genes to execute whole differentiation programs [9,10]. Another emerging cohort of molecules holding the same ability is that of miRNAs, the non-coding RNA assigned to regulate post-transcription gene expression. Like TFs, the temporal expression of miRNAs is highly regulated and responsive to changes, depending on the stem cell status. With the aim of deeply understanding the biology of ESCs and to learn more in the control of their potential, these cells have been profiled using different methods to identify miRNAs that have potential roles in stemness and differentiation [11,12]. These studies have revealed several miRNA families that are highly expressed, specifically, in undifferentiated cells [13], as well as the miRNAs that are specifically expressed when differentiation occurs, such as the well-known family of let-7 [14]. Over the past few years, it has become apparent that the distinctive pattern of miRNA expression seen in ESCs contributes to many of the unique phenotypic properties of these cells. Indeed, the switch from pluripotent to lineage-specific cells is marked by the downregulation of pluripotency markers and the activation of lineage-specific gene expression, which are accompanied by changes in the expression of many miRNAs. Whereas some miRNAs function in promoting exit from the pluripotent state by targeting pluripotency factors, other miRNAs stabilize the pluripotent state [15–18]. The introduction or depletion of miRNAs involved in regulating ESC fate may be useful in inducing differentiation along a particular lineage. Recent discoveries have revealed a model in which miRNA regulatory events are linked with transcription factor and signaling networks that control cell fate and differentiation, modulating their activity through positive and negative feedback loops to modulate stem cell fate decisions [19–21]. We have performed a miRNA profiling in ESCs undergoing neural differentiation, and we have demonstrated that some miRNAs are able to control ESC differentiation by targeting important regulators of chromatin remodeling [22]. Furthermore, we have recently illustrated a model in which miR-125a is

linked to the signaling of BMP4 and Nodal/Activin, modulating their activity through negative feedback loops to regulate ESC fate decisions [8]. Another miRNA belonging to the same family of miR-125a is miR-125b. We have analyzed the function of this miRNA, which shows an interesting expression profile during ESC differentiation. In this paper, we show that the modulation of this miRNA results in significant changes of the ESC state. These effects are due to the modulation of BMP4 and Nodal signaling during the first step of ESC differentiation. Moreover, we found that miR-125b, opposite to that observed for miR-125a, does not convey the control from TGF β signaling.

2. Results and Discussion

2.1. miR-125b Overexpression Blocks ESC at the Epiblast Stage

We screened for miRNAs differentially regulated during ESC neural differentiation and found many miRNAs that are specifically expressed in undifferentiated or differentiated cells [22]. Among these miRNAs, we found that miR-125b is expressed at low levels in ESCs, but its expression increases during the first steps of differentiation. At later time points, miR-125b expression reaches a higher level in differentiated cells (Figure 1A), and in agreement, it is highly expressed in many adult mouse tissues (Figure 1B). We have previously demonstrated that miR-125b, as well as miR-125a, are able to regulate the expression of the BMP4 co-receptor, *Dies1* [8], in ESCs, but we still don't know whether miR-125b may have a role in the early phases of ESC differentiation, when BMP4 regulates the differentiation fate of these cells. The analysis of undifferentiated markers clearly indicated that miR-125b overexpression (Figure S1) does not impair the undifferentiated state of ESCs (Figure 2A); thus, we explored whether miR-125b affects the differentiation program. ESCs transfected with pre-miR-125b were induced to differentiate as serum-free embryoid bodies (SFEBs), which mainly give rise to neuroectoderm derivatives at four days of differentiation. We found that the overexpression of miR-125b blocks ESC differentiation. Indeed, we found a decrease of neuroectodermal markers, whereas the expression of stemness markers is maintained at a high level (Figure 2B,C and Figure S2). These effects are accompanied by an impairment of ERK activation that suggests the failure of proper differentiation (Figure 2D). Considering that SFEB differentiation favors the transition through the epiblast stage that then leads to the formation of neuroectoderm, we performed an analysis to see if the block of differentiation observed upon miR-125b overexpression occurs before or after the epiblast transition. We found that the epiblast markers, *Fgf5*, *Cerberus* and *Dnmt3b* (Figure 2E), were significantly high at four days of differentiation, indicating that the cells are blocked in the epiblast stage. To verify that miR-125b overexpression is able to maintain the epiblast stem cell (EpiSC) phenotype, rather than simply slowing down the differentiation, we analyzed the methylation state of epiblast marker genes. We found that differentiated cells at day 4 upon miR-125b overexpression had epigenetic markers similar to those of EpiSCs (Figure 2F), thus suggesting, again, that these cells may still be pluripotent. To verify this hypothesis, we decided to test the pluripotency of this epiblast cell obtained upon miR-125b overexpression *in vivo*. Thus, we injected, into immunodeficient mice, cells transfected with miR-125b or with a control miR and pre-differentiated *in vitro* for three days. We found that miR-125b overexpressing cells differentiated for three days are still able to form an extensive differentiated teratoma (Figure 2G) in four out of the five mice injected with the cells. The control cells induced the formation of a small, not completely differentiated tumor (data not shown)

only in one mouse over five injected. These results demonstrated that miR-125b overexpression is able to sustain the undifferentiated phenotype, even three days after the induction of differentiation and that a large fraction of these cells maintains the pluripotency.

Figure 1. miR-125b expression in embryonic stem cells (ESCs) and mouse tissues. (A) miR-125b expression levels were analyzed by qPCR in undifferentiated ESCs and during neural differentiation through serum-free embryoid bodies (SFEBS) formation; (B) Analysis of miR-125b expression in mouse adult tissues. The data were normalized to the U6 internal control (* $p < 0.05$).

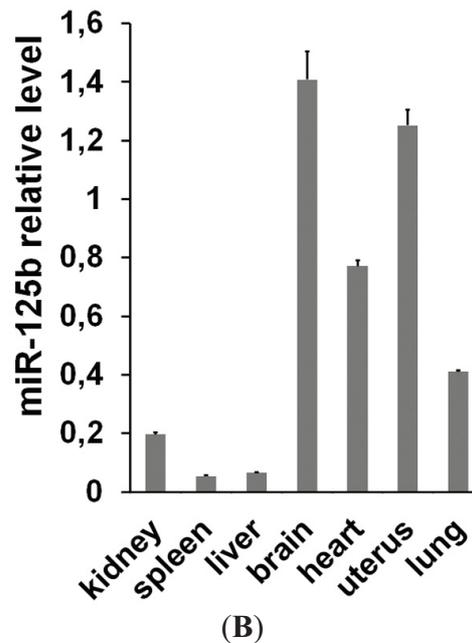
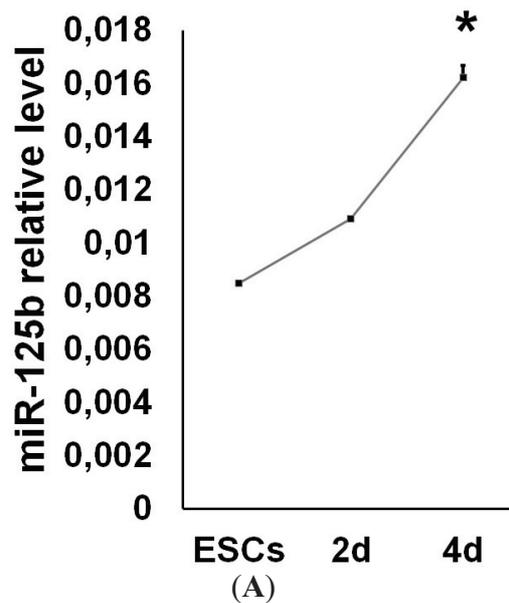


Figure 2. Effects of miR-125 ectopic expression *in vitro* and *in vivo*. **(A)** Analysis of stemness markers (Oct3/4, Nanog, Klf2 and Klf5) in undifferentiated ESCs transfected with pre-miR-125b or with control pre-miR (pre-miR-ctrl). The fold change is calculated by assigning the arbitrary value, one, to the amount found in cells transfected with control pre-miRNA; **(B)** Immunofluorescence analysis of four-day differentiated SFEBs upon miR-125b overexpression. Markers of pluripotency (Oct3/4, Nanog) and neuroectoderm (Sox1) are shown. Scale bar: 50 μ m; **(C)** qPCR analysis of stemness (Oct3/4 and Nanog) and neuroectodermal (Pax6) markers in differentiating ESCs upon miR-125b overexpression. The fold change is calculated by assigning the arbitrary value, one, to the time point showing the highest amount of the indicated mRNA; **(D)** The level of active ERK (P-ERK) were analyzed by Western blot in cells transfected with pre-miR-125b or the control pre-miR after four days of differentiation—the experiment shown in the Figure is representative of two independent experiments; **(E)** The level of the epiblast marker, Fgf5, was measured by qPCR in undifferentiated ESCs and during differentiation upon pre-miR transfection. The epiblast markers, Cerberus and Dnmt3b, were measured at four days of SFEB differentiation in the cells transfected with the indicated pre-miR. The fold change is calculated as indicated in **(C)**; **(F)** ChIP-qPCR analysis was performed on chromatin from ESCs transfected with the indicated pre-miR and induced differentiation for four days as SFEBs. The graphs show the methylation state of histone H3 on the promoters of pluripotency (Nanog and Klf2) and epiblast (Fgf5) markers. Data are expressed as fold enrichment relative to the control; **(G)** Immunodeficient mice were injected with ESCs transfected with the pre-miR-125b (right side) and ctrl pre-miR (left side) after three days of differentiation *in vitro* (left panel). Teratomas generated by ESCs overexpressing miR-125b were explanted after one month, and the tissues were analyzed after eosin-hematoxylin staining (right panels) (* $p < 0.05$).

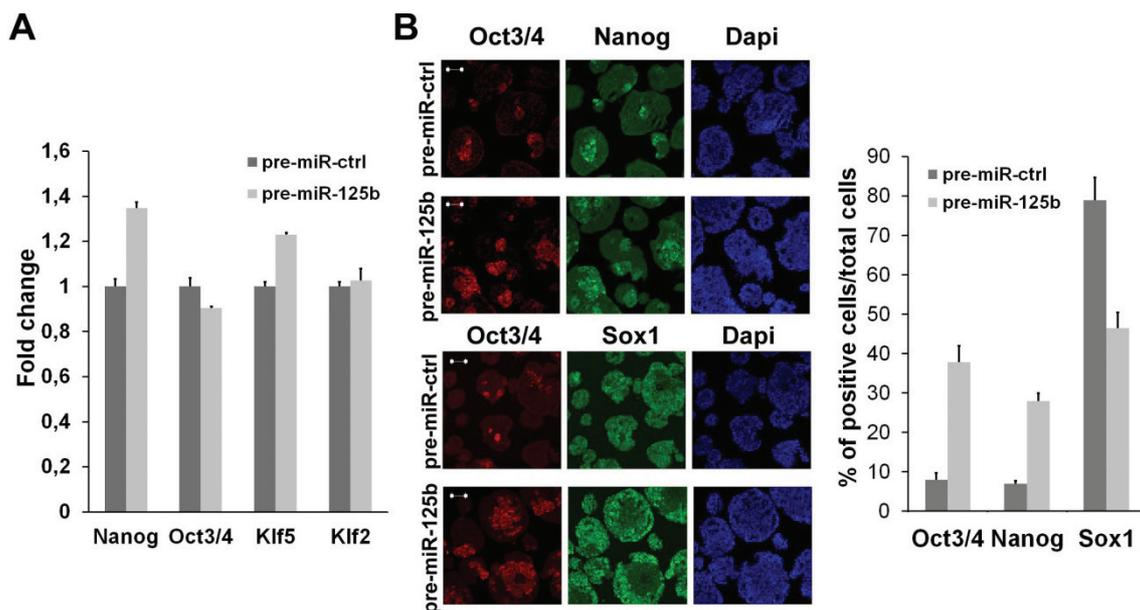
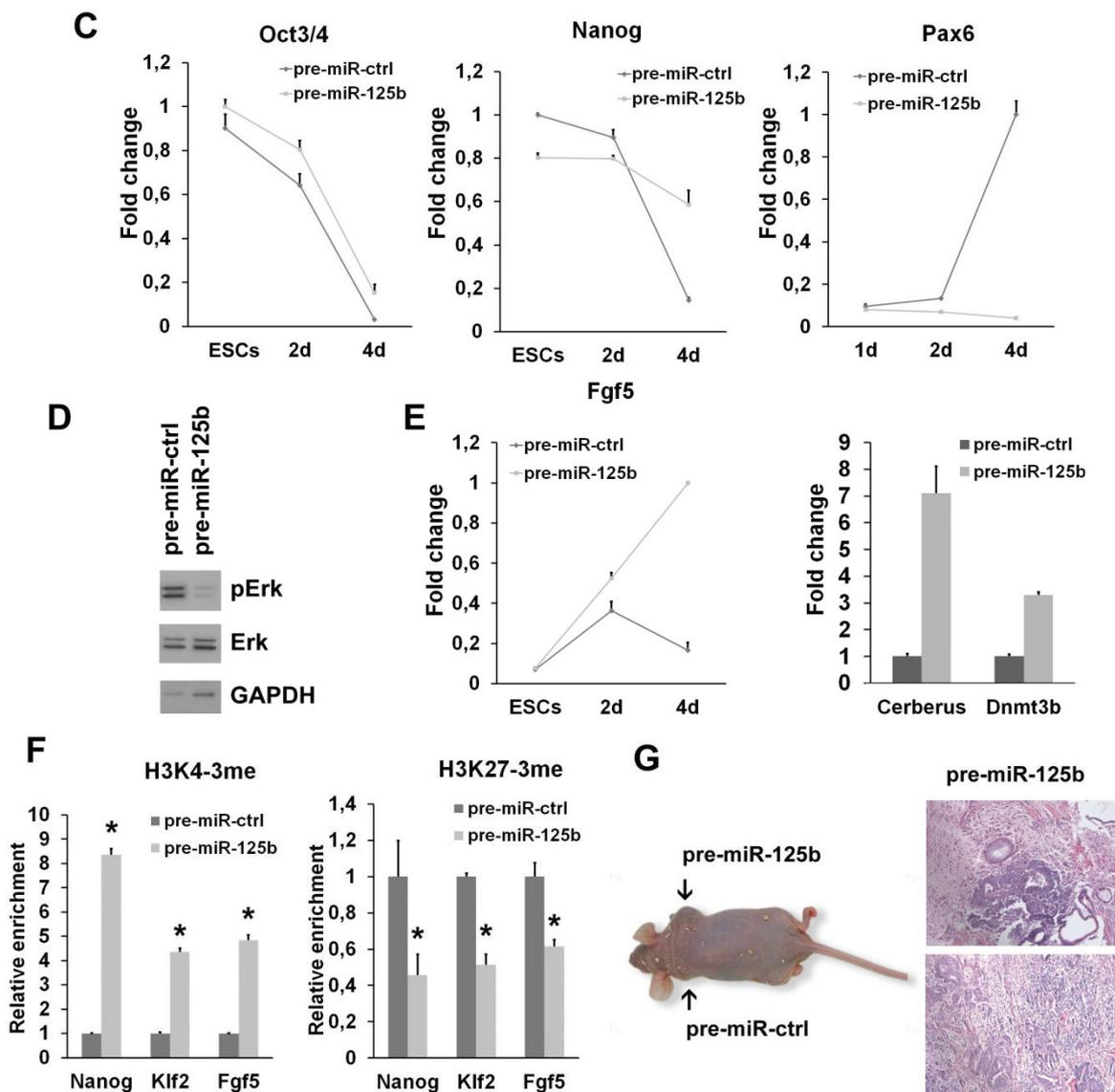


Figure 2. *Cont.*

2.2. miR-125b Effects on the ESC-Epiblast Transition Are Due to *Dies1*

Based on the evidence that miR-125b is able to control the expression of *Dies1* [8], the BMP4 co-receptor, we analyzed whether miR-125b overexpression alters the signaling of BMP4. miR-125b overexpression induced a significant decrease of BMP4 targets during ESC differentiation, accompanied by an evident increases of Nodal/Activin targets (Figure 3A). To address whether these effects of miR-125b overexpression are due to the suppression of *Dies1*, we tried to rescue the proper differentiation and the proper balance between BMP4 and Nodal pathways by re-expressing a form of *Dies1* insensitive to the miR-125b. We found that *Dies1* is able to fully rescue the block at the epiblast stage induced by miR-125b (Figure 3B,C). Moreover, this rescue corresponds to the restoration of the proper expression levels of BMP4 and Nodal/Activin targets (Figure 3D). A recent paper has indicated that miR-125b targets *Lin28* in ESCs to regulate mesendodermal differentiation [23]. To explore this point, we analyzed the mRNA and protein levels of *Lin28* upon miR-125b overexpression during SFEB differentiation. Interestingly, we found that *Lin28* expression is not impaired in this context (Figure 3E), indicating that the block at the epiblast stage induced by miR-125b overexpression is not

due to the repression of Lin28. Instead, we found that miR-125a overexpression induces a slight decrease of Lin28 protein level, suggesting that these two miRNAs can act during the early phases of ESC differentiation by regulating different subsets of targets.

Figure 3. The effects of miR-125b overexpression on TGF β signaling are mediated by the BMP4 co-receptor, Dies1. **(A)** qPCR analysis of the expression levels of BMP4 (Id1, Id3) and Nodal/Activin (Nodal, Cripto, Lefty1, Lefty2) target genes upon miR-125b overexpression; **(B)** Analysis of the phenotype of ESCs co-transfected with the indicated pre-miR and with the vector expressing Dies1 lacking its 3'UTR or with the empty vector (mock). The expression of stemness (Oct3/4) and neuroectodermal (Sox1) markers was analyzed by immunostaining in cells differentiated as SFEBs for four days. Scale bar: 20 μ m; **(C)** q-PCR analysis of the effects of Dies1 re-expression in ESCs transfected with the indicated pre-miR. After four days of differentiation, the expression of stemness (Oct3/4, Nanog) and epiblast (Fgf5) markers was analyzed; **(D)** q-PCR analysis of the expression of BMP4 (Id1) and Nodal/Activin (Lefty1 and Lefty2) targets in four-day differentiated ESCs re-expressing, or not, Dies1 upon miR-125b overexpression. Data in **(C)** and **(D)** are shown as fold changes relative to cognate controls; **(E)** The effects of miR-125a and b overexpression on Lin28 level were analyzed in four-day differentiated SFEBs by means of q-PCR (left panel) and Western blot (right panel). Data are expressed as fold change relative to the control (* $p < 0.05$).

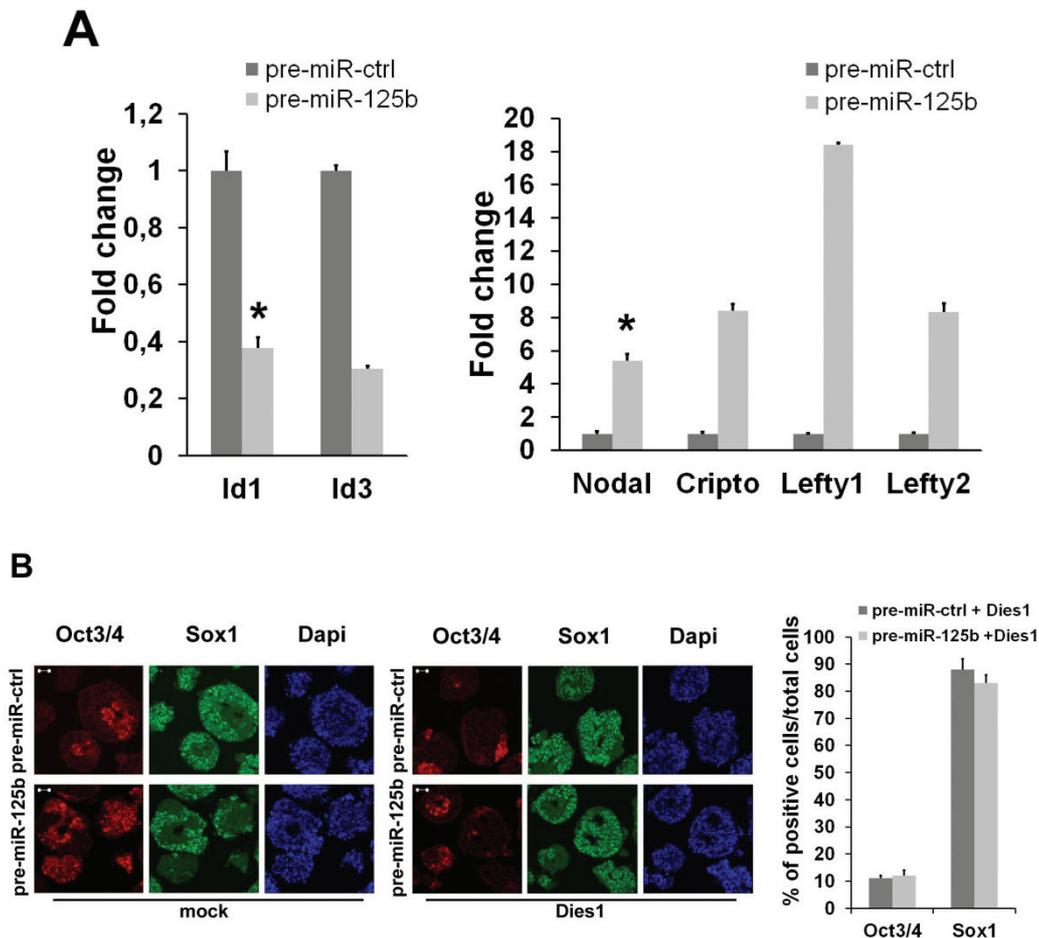
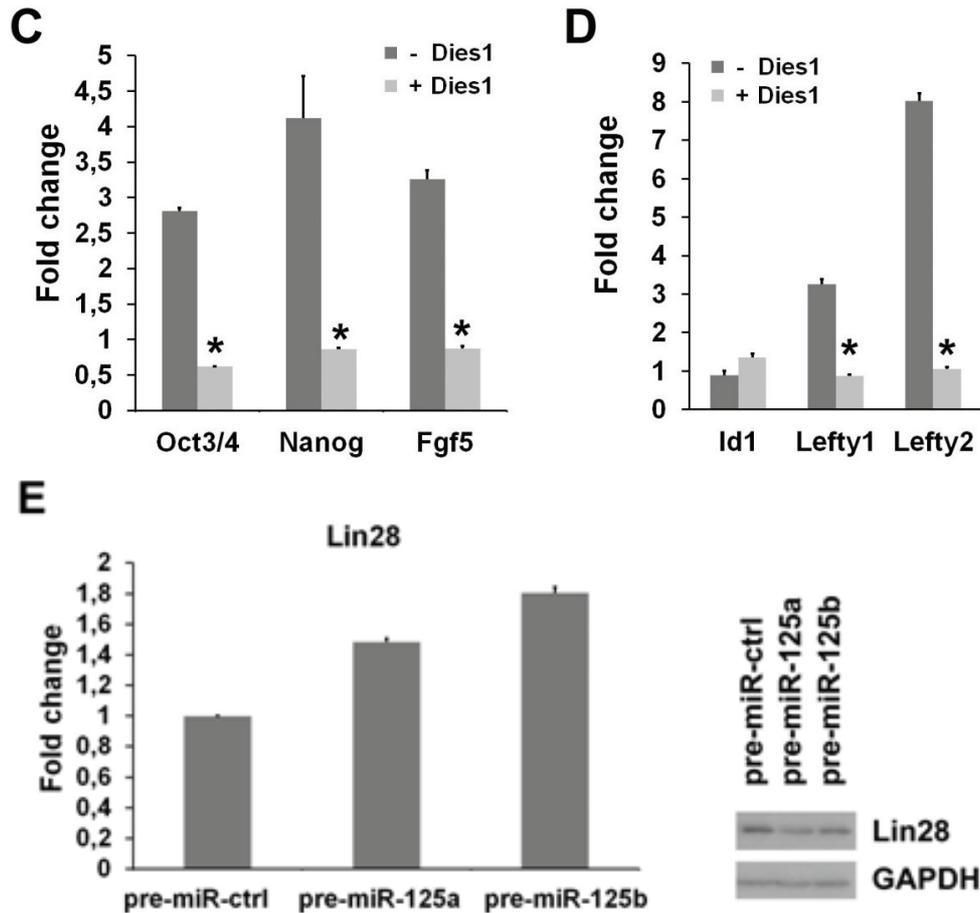
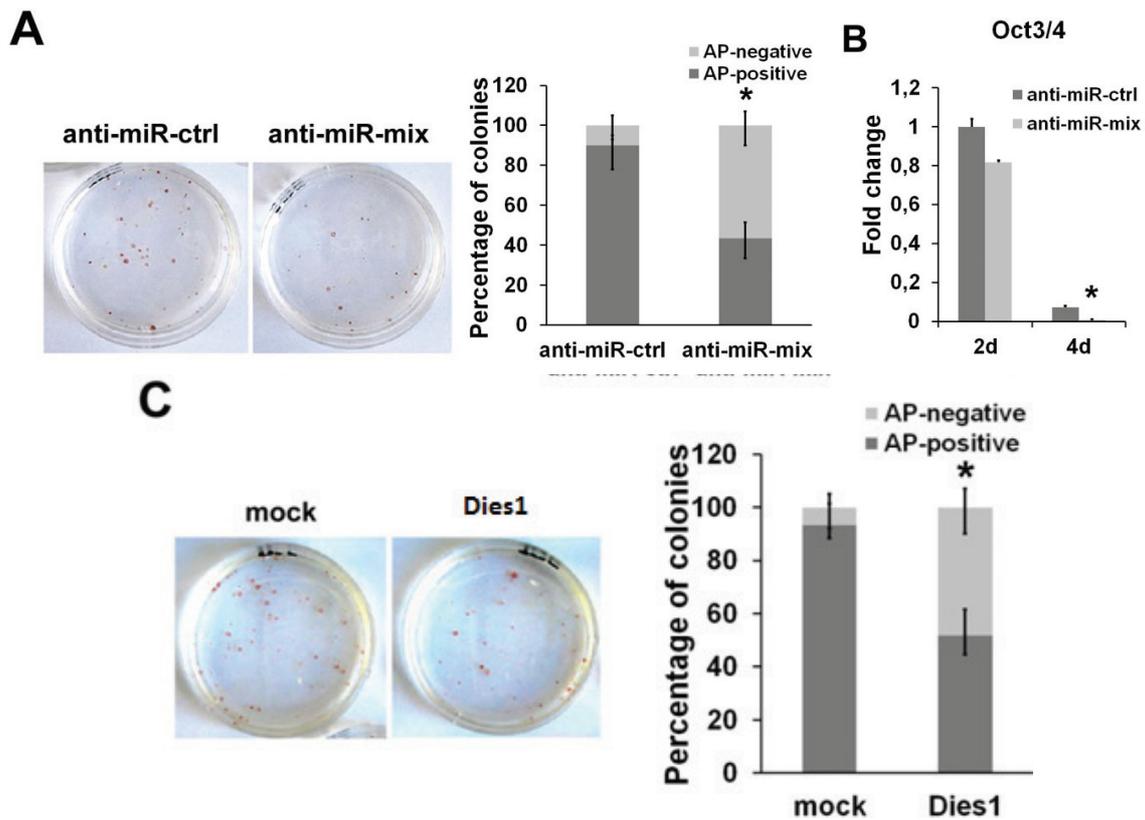


Figure 3. *Cont.*

2.3. miR-125 Suppression Promotes ESC Differentiation

Considering the evident phenotype induced by miR-125b overexpression, we asked if the suppression of this miR may have an effect on ESCs. To analyze the effects of the depletion of miR-125b in ESC differentiation, we suppressed the endogenous miRNAs by transfecting a mix of anti-miR-125a and anti-miR-125b to avoid the levels of endogenous miR-125a from being able to be substituted for the absence of miR-125b. An anti-miR with no complementarity to any known miRNAs was used as negative control. After anti-miR transfection, we cultured ESCs at low density, and after seven days, we performed an alkaline phosphatase (AP) staining to assay their ability to maintain an undifferentiated phenotype. As shown in Figure 4A, the cells transfected with anti-miR-125 showed a reduction in the number of AP-positive colonies compared to the control, indicating that miRNA suppression causes the loss of the undifferentiated state, even in the presence of LIF. Moreover, we found a rapid decrease of the stemness marker, Oct3/4, upon miR-125 suppression during differentiation, already at day two of SFEB formation, which became more evident at day four (Figure 4B). To better define the correlation between miR-125b and Dies1 in this context, we analyzed the effects of Dies1 ectopic expression in ESCs in the same conditions used for miRNA suppression. We found that Dies1 forced expression induces the decrease of the number of AP-positive colonies, thus resembling the phenotype observed upon miR-125 suppression (Figure 4C). These results support the idea that the levels of the two miR-125 are functionally correlated to that of Dies1 in ESCs.

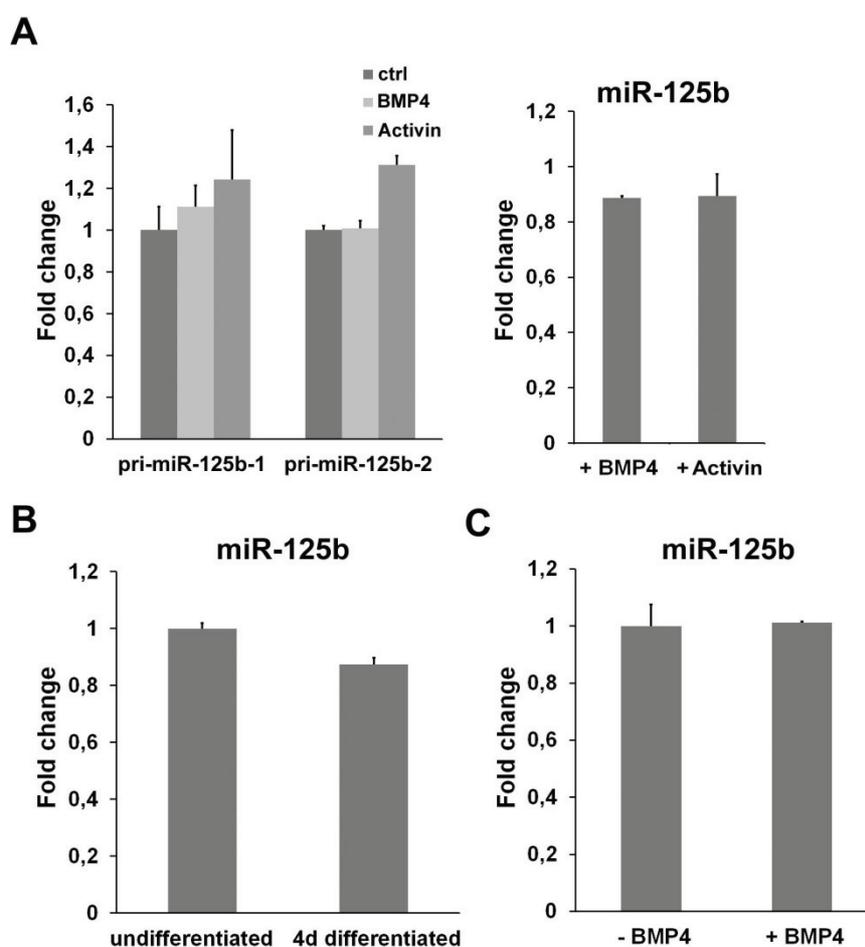
Figure 4. mir-125 suppression induces ESC differentiation. (A) Alkaline phosphatase (AP) staining was performed on cells transfected with the mix of anti-miR-125a and b (anti-miR-mix) and with the control anti-miR (anti-miR-ctrl) and cultured for seven days at clonal density in the presence of Leukemia Inhibitory Factor (LIF). The histogram represents the number of AP-positive and -negative colonies; (B) q-PCR analysis showing the expression of Oct3/4 in differentiating ESCs upon suppression of miR-125a and b. Data are expressed as fold change relative to the control; (C) AP staining of ESCs overexpressing Dies1 after seven days of culture at clonal density in the presence of LIF. The number of AP-positive and -negative colonies is reported in the graph (* $p < 0.05$).



2.4. miR-125b Is Not Regulated by a TGF β Signaling

We have shown that miR-125b overexpression impairs both BMP4 and Nodal/Activin pathways. To understand if the miR-125b can be directly regulated by one of these signalings, we exposed ESCs to BMP4 or Activin and analyzed the level of the two pri-miR-125b transcribed from the two miR-125b genes and the mature miR-125b. As shown in Figure 5A, we did not find any significant changes in the expression of miR-125b upon BMP4 or Activin treatment, indicating that this miR is not directly regulated by these two pathways. To verify that this independence of miR-125b from TGF β is not context-dependent, we analyzed the changes in the possible expression of miR-125b in a different experimental setting. To this aim, we used C2C12, a system in which both miR-125b and BMP4 are able to regulate the differentiation [24,25] and in which miR-125b is highly expressed (Figure 5B). After the exposure of these cells to BMP4, we didn't find any changes in the expression level of miR-125b (Figure 5C). All these data suggest that miR-125b is able to impair the TGF β pathway, but is not directly controlled by these molecules.

Figure 5. miR-125b expression is independent from TGF β . (A) Analysis of the level of miR-125b after the exposure of ESCs to BMP4 or Activin. The expression of pri-miRs (left panel) was measured after 1 h of treatment with the indicated molecules by using specific primers that distinguish between the transcripts deriving from the two miR-125b genes (pri-miR-125b-1 and pri-miR-125b-2). The level of mature miR-125b was measured after 24 h of BMP4 or Activin treatment; (B) Expression level of miR-125b in undifferentiated and four-day differentiated C2C12 cells; (C) q-PCR analysis of the expression of miR-125b in C2C12 cells after 24 h of treatment with BMP4. All the data are expressed as fold change relative to the control.



3. Experimental Section

3.1. Cell Culture, Transfection and Differentiation

E14Tg2a mouse ESCs (BayGenomics, San Francisco, CA, USA) were maintained on feeder-free, gelatin-coated plates in the following medium: Glasgow Minimum Essential Medium (GMEM, Sigma, St. Louis, MO, USA) supplemented with 2 mM glutamine (Invitrogen, Carlsbad, CA, USA), 1 mM sodium pyruvate (Invitrogen, Carlsbad, CA, USA), 1 \times nonessential amino acids (Invitrogen, Carlsbad, CA, USA), 0.1 mM β -mercaptoethanol (Sigma, St. Louis, MO, USA), 10% FBS (Hyclone, Waltham, MA, USA) and 10³ U/mL Leukemia Inhibitory Factor (LIF, Millipore, Billerica, MA, USA). Serum-free embryoid body (SFEB) differentiation was performed by plating 1 \times 10⁶ ESCs in 100 mm Petri dishes (BD Biosciences, San Jose, CA, USA) in the following differentiation medium: GMEM

supplemented with 2 mM glutamine, 1 mM sodium pyruvate, 1× nonessential amino acids, 0.1 mM β -mercaptoethanol and 10% Knock-out Serum Replacement (KSR, Invitrogen, Carlsbad, CA, USA). In this condition, after 2 days, the cells express a high level of epiblast markers that strongly decrease at 4 days of SFEB differentiation. EpiSCs can be derived with high yield at 2 days of SFEB differentiation, whereas at 3 and 4 days of differentiation, the yield of EpiSC derivation strongly decreases [5].

C2C12 myoblasts were maintained in Dulbecco's Modified Eagle Medium (Sigma, St. Louis, MO, USA) supplemented with 2 mM glutamine (Invitrogen, Carlsbad, CA, USA), 100 U/mL penicillin/streptomycin (Invitrogen, Carlsbad, CA, USA) and 10% FBS (GIBCO, Life Technologies Italia, Monza, Italy). For differentiation to myotubes, 20×10^3 cells/cm² were plated, and the following day, the medium was substituted with the following differentiation medium: DMEM with 2 mM glutamine (Invitrogen, Carlsbad, CA, USA) and 2% horse serum (Sigma, St. Louis, MO, USA). After 4 days of differentiation, the cells were collected.

Transfection of pre-miRs, anti-miRs (both from Ambion, Austin, TX, USA) and plasmid for *Dies1* expression was performed using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA), following the manufacturer's instructions.

3.2. Cell Treatment and Alkaline Phosphatase Staining

For cell treatment, ESCs and C2C12 were grown overnight in KSR containing medium with LIF or DMEM plus 1% FBS, respectively and, then, treated for the indicated time with 20 ng/mL of BMP4 or Activin (both from R&D, Minneapolis, MN, USA).

Alkaline phosphatase staining was performed culturing ESCs at clonal density (20–50 cells/cm²). After 7 days, the cells were fixed in 10% cold Neutral Formalin Buffer (10% formalin, 110 mM Na₂HPO₄, 30 mM NaH₂PO₄ in H₂O) for 15 min and, then, rinsed in distilled water for 15 min. The staining was obtained by incubation for 45 min at room temperature with the following staining solution: 0.1 M Tris-HCl, 0.01% naphthol AS MX-PO₄ (Sigma, St. Louis, MO, USA), 0.4% *N,N*-dimethylformamide (Sigma, St. Louis, MO, USA) and 0.06% Red Violet LB salt (Sigma, St. Louis, MO, USA).

3.3. RNA Isolation, q-PCR and TaqMan Analysis

Total RNA from ESCs and C2C12 was extracted by using TRI-Reagent (Sigma, St. Louis, MO, USA). For the first-strand cDNA synthesis, the manufacturer's instructions (M-MLV RT, New England BioLabs, Ipswich, MA, USA) were followed. Q-PCR was carried out on an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA, USA) using Power SYBR Green PCR Master mix (Applied Biosystems, Foster City, CA, USA). The housekeeping GAPDH mRNA was used to normalize the samples, using the $2^{-\Delta\Delta C_t}$ method. The gene-specific primers used are listed in Table S1.

For the measurement of mature miRNA, RNA was extracted with a mirVana microRNA Isolation kit (Ambion, Austin, TX, USA), according to the manufacturer's instructions. Single-stranded cDNA was synthesized from 10 ng of total RNA combined with the specific primer for miR-125a or miR-125b or U6 as internal control by using a TaqMan MicroRNA reverse transcription kit (Applied Biosystems, Foster City, CA, USA). miRNA levels were measured by using a TaqMan MicroRNA

detection kit (Applied Biosystems, Foster City, CA, USA) with the 7500 Real Time PCR System instrument and the Sequence Detection Systems (SDS) software version 1.4 (Applied Biosystems, Foster City, CA, USA) [26].

3.4. Northern Blot Analysis

For Northern blot analysis total RNA was isolated by using the TRI Reagent (Sigma, St. Louis, MO, USA), according to the manufacturer's instructions. For each sample, 20 µg of total RNA were fractionated on 15% TBE-Urea gel (Criterion precast Gel, Bio-Rad, Segrate, Italy), stained with EtBr (Sigma, St. Louis, MO, USA) for loading control, then transferred to a Hybond membrane (Amersham Pharmacia, Milan, Italy) and, finally, fixed by UV cross-linking in a Stratalinker (Stratagene, La Jolla, CA), according to the manufacturer's instruction. The membrane was hybridized with 10 pmol of miRCURY LNA Detection Probe, digoxigenin labeled (Exiqon, Vedbaek, Denmark), according to the manufacturer's instruction. The signal detection was obtained by an anti-digoxigenin-alkaline phosphatase antibody (Roche, Basel, Switzerland) and a chemiluminescent substrate (CDP-Star, Roche, Basel, Switzerland), according to the manufacturer's instruction.

3.5. Protein Extraction and Western Blot Analysis

For protein extracts, cells were lysed in a buffer containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1% Triton, 1% sodium deoxycholate and protease inhibitor cocktail (Sigma, St. Louis, MO, USA). The antibodies used for blotting were: Anti-Lin28 (1:700, Abcam, Cambridge, UK); anti-phospho-Erk1 (1:1000, Cell Signaling, Danvers, MA, USA); anti-Erk1 (1:1000, Santa Cruz, Santa Cruz, CA, USA); anti Oct3/4 (1:1000, Santa Cruz, Santa Cruz, CA, USA); and anti-GAPDH (1:1000, Santa Cruz, Santa Cruz, CA, USA).

3.6. Immunostaining

For immunostaining, 4-day differentiated SFEBs were fixed in 4% paraformaldehyde and processed, as described in Parisi *et al.*, 2012 [12]. The following primary antibodies were used: anti-Oct3/4 (1:200, Santa Cruz, Santa Cruz, CA, USA), anti-Nanog (1:500, Calbiochem, San Diego, CA, USA) and anti-Sox1 (1:100, Santa Cruz, Santa Cruz, CA, USA). The appropriate secondary antibodies were used (1:400, Alexa Molecular Probes, Invitrogen, Carlsbad, CA, USA). Confocal microscopy was performed with an LSM 510 Meta microscope (Zeiss, Milan, Italy) using LSM 510 Meta software [27] and the LSM Image Browser (Zeiss, Milan, Italy). The brightness, contrast and color balance of the images were adjusted in Photoshop CS2 (Adobe Systems, Agrate Brianza, Italy).

3.7. Teratoma Formation

ESCs transfected with pre-miR-125b or pre-miR-ctrl were differentiated as SFEBs for 3 days. Then, SFEBs were dissociated, and 2×10^6 cells were used for subcutaneous injection in nude mice. Four weeks after the injection, tumors were surgically dissected from the mice. Samples were fixed in 4% paraformaldehyde and embedded in paraffin. Sections were stained with hematoxylin and eosin.

3.8. Chromatin Immunoprecipitation (ChIP)-qPCR Analysis

Chromatin immunoprecipitation was performed as described previously [28]. To immunoprecipitate soluble chromatin extracts, anti-H3K4-3me (Millipore, Billerica, MA, USA) and anti-H3K27-3me (Millipore, Billerica, MA, USA) antibodies were used. Appropriate IgGs were used as negative control. Supernatant obtained without antibody was used as input control. After q-PCR, the amount of precipitated DNA was calculated relative to the total input chromatin and expressed as the percentage of total chromatin, according to the formula $2^{\Delta Ct}$, where Ct represents the cycle threshold and $\Delta Ct = Ct(\text{input}) - Ct(\text{immunoprecipitation})$. The gene-specific primers used are listed in Table S1.

3.9. Statistics

Data are presented as the means \pm SD of at least three independent experiments. Whenever necessary, the statistical significance of the data was analyzed using the Student's t -test (* $p < 0.05$).

4. Conclusions

We demonstrated that miR-125b is essential for the proper differentiation of ESCs, based on the relevant effects of its overexpression and suppression. We have previously demonstrated that miR-125a had similar effects on ESC differentiation [8], thus suggesting that these two miRNAs work cooperatively through the suppression of *Dies1*, the co-receptor of BMP4 [29]. This relationship is an elegant example of how multiple miRNAs can converge on a single pathway to promote a common outcome. Interestingly, while miR-125a is directly regulated by BMP4, miR-125b seems to not be regulated by TGF β signaling. This observation suggests that the BMP4 signaling in the first steps of ESC differentiation undergoes different regulations that are dependent (miR-125a) or independent (miR-125b) by itself. This can probably be due to the relevance that the balance between BMP4 and Nodal/Activin pathways have in the control of the transition from ESC to the epiblast stage, thus indicating that ESCs modulate in different ways such important pathways.

Acknowledgments

The authors thank Danila Scarpa for precious assistance. This work was supported by grants from Associazione Italiana Ricerca sul Cancro and Italian Ministry of Research (PON 01_02782).

Conflict of Interest

The authors declare no conflict of interest.

References

1. Niwa, H.; Ogawa, K.; Shimosato, D.; Adachi, K. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature* **2009**, *460*, 118–122.
2. Parisi, S.; D'Andrea, D.; Lago, C.T.; Adamson, E.D.; Persico, M.G.; Minchiotti, G. Nodal-dependent Cripto signaling promotes cardiomyogenesis and redirects the neural fate of embryonic stem cells. *J. Cell Biol.* **2003**, *163*, 303–314.

3. Seuntjens, E.; Umans, L.; Zwijsen, A.; Sampaolesi, M.; Verfaillie, C.M.; Huylebroeck, D. Transforming growth factor type beta and Smad family signaling in stem cell function. *Cytokine Growth Factor Rev.* **2009**, *20*, 449–458.
4. Watabe, T.; Miyazono, K. Roles of TGF-beta family signaling in stem cell renewal and differentiation. *Cell Res.* **2009**, *19*, 103–115.
5. Zhang, K.; Li, L.; Huang, C.; Shen, C.; Tan, F.; Xia, C.; Liu, P.; Rossant, J.; Jing, N. Distinct functions of BMP4 during different stages of mouse ES cell neural commitment. *Development* **2010**, *137*, 2095–2105.
6. Lee, K.L.; Lim, S.K.; Orlov, Y.L.; Yit le, Y.; Yang, H.; Ang, L.T.; Poellinger, L.; Lim, B. Graded Nodal/Activin signaling titrates conversion of quantitative phospho-Smad2 levels into qualitative embryonic stem cell fate decisions. *PLoS Genet.* **2011**, doi:10.1371/journal.pgen.1002130.
7. Galvin, K.E.; Travis, E.D.; Yee, D.; Magnuson, T.; Vivian, J.L. Nodal signaling regulates the bone morphogenic protein pluripotency pathway in mouse embryonic stem cells. *J. Biol. Chem.* **2010**, *285*, 19747–19756.
8. Parisi, S.; Battista, M.; Musto, A.; Navarra, A.; Tarantino, C.; Russo, T. A regulatory loop involving *Dies1* and miR-125a controls BMP4 signaling in mouse embryonic stem cells. *FASEB J.* **2012**, *26*, 3957–3968.
9. Chambers, I.; Tomlinson, S.R. The transcriptional foundation of pluripotency. *Development* **2009**, *136*, 2311–2322.
10. Parisi, S.; Russo, T. Regulatory role of *Klf5* in early mouse development and in embryonic stem cells. *Vitam Horm.* **2011**, *87*, 381–397.
11. Ivey, K.N.; Srivastava, D. MicroRNAs as regulators of differentiation and cell fate decisions. *Cell Stem Cell* **2010**, *7*, 36–41.
12. Martinez, N.J.; Gregory, R.I. MicroRNA gene regulatory pathways in the establishment and maintenance of ESC identity. *Cell Stem Cell* **2010**, *7*, 31–35.
13. Houbaviy, H.B.; Murray, M.F.; Sharp, P.A. Embryonic stem cell specific microRNAs. *Dev. Cell* **2003**, *5*, 351–358.
14. Viswanathan, S.R.; Daley, G.Q.; Gregory, R.I. Selective blockade of microRNA processing by *Lin28*. *Science* **2008**, *320*, 97–100.
15. Sinkkonen, L.; Hugenschmidt, T.; Berninger, P.; Gaidatzis, D.; Mohn, F.; Artus-Revel, C.G.; Zavolan, M.; Svoboda, P.; Filipowicz, W. MicroRNAs control de novo DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.* **2008**, *15*, 259–267.
16. Zovoilis, A.; Smorag, L.; Pantazi, A.; Engel, W. Members of the miR-290 cluster modulate in vitro differentiation of mouse embryonic stem cells. *Differentiation* **2009**, *78*, 69–78.
17. Wang, Y.; Baskerville, S.; Shenoy, A.; Babiarz, J.E.; Baehner, L.; Blelloch, R. Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. *Nat. Genet.* **2008**, *40*, 1478–1483.
18. 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.

19. 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.
20. Delaloy, C.; Liu, L.; Lee, J.A.; Su, H.; Shen, F.; Yang, G.Y.; Young, W.L.; Ivey, K.N.; Gao, F.B. MicroRNA-9 coordinates proliferation and migration of human embryonic stem cell-derived neural progenitors. *Cell Stem Cell* **2010**, *6*, 323–335.
21. Krichevsky, A.M.; Sonntag, K.C.; Isacson, O.; Kosik, K.S. Specific microRNAs modulate embryonic stem cell-derived neurogenesis. *Stem Cells* **2006**, *24*, 857–864.
22. Tarantino, C.; Paoletta, G.; Cozzuto, L.; Minopoli, G.; Pastore, L.; Parisi, S.; Russo, T. miRNA-34a, 100, and 137 modulate differentiation of mouse embryonic stem cells. *FASEB J.* **2010**, *24*, 3255–3263.
23. Wang, J.; Cao, N.; Yuan, M.; Cui, H.; Tang, Y.; Qin, L.; Huang, X.; Shen, N.; Yang, H.T. MicroRNA-125b/Lin28 pathway contributes to the mesendodermal fate decision of embryonic stem cells. *Stem Cells Dev.* **2012**, *21*, 1524–1537.
24. Ge, Y.; Sun, Y.; Chen, J. IGF-II is regulated by microRNA-125b in skeletal myogenesis. *J. Cell Biol.* **2011**, *192*, 69–81.
25. Dahlqvist, C.; Blokzijl, A.; Chapman, G.; Falk, A.; Dannaeus, K.; Ibáñez, C.F.; Lendahl, U. Functional Notch signaling is required for BMP4-induced inhibition of myogenic differentiation. *Development* **2003**, *130*, 6089–6099.
26. *Sequence Detection Systems (SDA)*, version 1.4; software for detecting nucleic acid sequence Available online: <http://www.lifetechnologies.com/global/en/home/technical-resources/software-downloads/applied-biosystems-7500-fast-real-time-pcr-system.html> (accessed on 31 January, 2008).
27. *LSM 510 Meta software*; Available online: http://microscopy.zeiss.com/microscopy/n_us/downloads/lsm-5-series.html (accessed on 27 February 2007).
28. Parisi, S.; Cozzuto, L.; Tarantino, C.; Passaro, F.; Ciriello, S.; Aloia, L.; Antonini, D.; de Simone, V.; Pastore, L.; Russo, T. Direct targets of Klf5 transcription factor contribute to the maintenance of mouse embryonic stem cell undifferentiated state. *BMC Biol.* **2010**, doi:10.1186/1741-7007-8-128.
29. Aloia, L.; Parisi, S.; Fusco, L.; Pastore, L.; Russo, T. Differentiation of embryonic stem cells 1 (Dies1) is a component of bone morphogenetic protein 4 (BMP4) signaling pathway required for proper differentiation of mouse embryonic stem cell. *J. Biol. Chem.* **2010**, *285*, 7776–7783.

Reprinted from *IJMS*. Cite as: Ruan, H.; Liu, S.; Li, F.; Li, X.; Fan, C. Prevention of Tendon Adhesions by ERK2 Small Interfering RNAs. *Int. J. Mol. Sci.* **2013**, *14*, 4361-4371.

Article

Prevention of Tendon Adhesions by ERK2 Small Interfering RNAs

Hongjiang Ruan [†], Shen Liu [†], Fengfeng Li, Xujun Li and Cunyi Fan ^{*}

Department of Orthopaedics, Shanghai Sixth People's Hospital, Shanghai Jiaotong University, School of Medicine, 600 Yishan Road, Shanghai 200233, China;
E-Mails: ruanhongjiang@126.com (H.R.); liushensjtu@sjtu.edu.cn (S.L.); fengmale@yahoo.com.cn (F.L.); lixujun1983@126.com (X.L.)

[†] These authors contributed equally to this work.

^{*} Author to whom correspondence should be addressed; E-Mail: fancunyi888@hotmail.com; Tel.: +86-21-6436-9181; Fax: +86-21-6406-3802.

Received: 21 November 2012; in revised form: 6 January 2013 / Accepted: 7 January 2013 / Published: 21 February 2013

Abstract: Tendon adhesions are one of the most concerning complications after surgical repair of flexor tendon injury. Extracellular signal-regulated kinase (ERK) 2 plays crucial roles in fibroblast proliferation and collagen expression which contributes to the formation of tendon adhesions after flexor tendon surgery. Using a chicken model, we have examined the effects of a small interfering RNA (siRNA) targeting ERK2 delivered by a lentiviral system on tendon adhesion formation with an adhesion scoring system, histological assessment, and biomechanical evaluation. It was found that ERK2 siRNA effectively suppressed the increase of fibroblasts and the formation of tendon adhesions ($p < 0.05$ compared with the control group). Moreover, no statistically significant reduction in breaking force was detected between the ERK2 siRNA group and the control group. These results show that the lentiviral-mediated siRNA system is effective in preventing tendon adhesion formation but not to tendon healing, and may be used for tendon repair after confirmation and improvement by future detailed studies.

Keywords: tendon repair; adhesion; siRNA; extracellular signal-regulated kinase (ERK) 2; fibroblast proliferation; lentivirus

1. Introduction

Tendon injuries are the second most common hand injuries in trauma and orthopedic patients [1]. In most cases, surgical repair or transplantation is required [2]. However, as a result of an inflammatory response at the surgical site, and the loss of physical separation, local tendon adhesions, as one of the most concerning complications in tendon repair, may form between the tendons and the surrounding sheath [3]. Usually, tendon adhesions bind the flexor tendons to each other or to their sheath, which restricts normal tendon gliding and consequently leads to poor functional recovery. A number of biological or synthetic materials, such as amniotic membrane, Seprafilm and polytetrafluoroethylene membrane, have been evaluated as mechanical barriers to prevent tendon adhesions [4–6]. In addition, the inhibitory effects of pharmacologic agents, such as 5-fluorouracil and hyaluronic acid [3,7,8], on adhesion formation have been investigated as well. Although these studies have obtained some improvements in materials used in surgical repair, due to their complexity and restriction, no reliable therapy has been established. Therefore, it is necessary to develop novel practical approaches to clinically prevent adhesion formation.

RNA interference (RNAi) is an evolutionarily conserved process in which cells employ small interfering RNA (siRNA) duplexes to destroy target messenger mRNAs, so as to silence the activity of corresponding genes [9,10]. The process exists in a variety of organisms and is used to regulate many diverse cellular processes. As a novel and revolutionary approach, RNAi has been rapidly and extensively used in basic biological research and the research and development of drugs and therapies, and has demonstrated great value, owing to its high specificity and potency [11–13]. Nevertheless, whether RNAi can be applied in tendon repair to suppress, or block, adhesion formation remains unclear.

Previous studies have shown that fibroblast proliferation and collagen expression play important roles in the formation of tendon adhesions [14,15]. In our previous work, we demonstrated that extracellular signal-regulated kinase (ERK) 2 is involved in the regulation of collagen expression and fibroblast proliferation induced by transforming growth factor (TGF)- β 1 and fibroblast growth factor (FGF)-2 [16]. Furthermore, the effect of the intra-articular administration of a siRNA, targeting ERK2 on joint adhesion formation, has been investigated in a rat model through lentiviral-mediated RNA interference, and the results indicated that local delivery of this siRNA effectively diminishes joint adhesion formation [17]. All these lines of evidence suggest that lentiviral-mediated ERK2 siRNA may be applied in tendon repair, to provide hints for prevention of tendon adhesion formation. Therefore, in this study, the effect of ERK2 siRNA on adhesion formation was investigated in a chicken model of tendon repair.

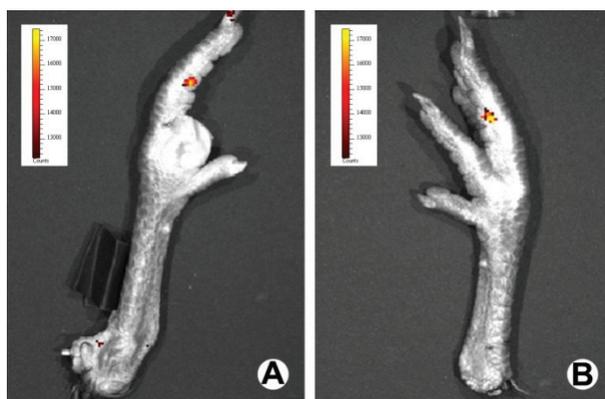
2. Results and Discussion

The lentiviral vector was successfully constructed (data not shown) and the high-titer lentiviral supernatants (for ERK2 siRNA and MS (a mis-sense) siRNA) were obtained for follow-up experiments in a chicken model.

2.1. Lentiviral-Mediated ERK2 siRNA Expression

Fourteen days after surgical repair, bioluminescent imaging demonstrated that the administration of the lentiviral-mediated siRNA system resulted in a localized expression of luciferase around the repaired flexor digitorum profundus (FDP), which also represented the distribution of ERK2 siRNA (Figure 1A). In addition, similar bioluminescent images were obtained in the luciferase fluorescence measurement carried out after 28 days (Figure 1B).

Figure 1. Bioluminescent imaging of a representative chicken foot of the ERK2 siRNA group at 14 (A) and 28 (B) days after surgical repair.



2.2. Effective Silencing of ERK2 and pERK2 by Lentiviral-Mediated RNA Interference

In the control group, the background level of ERK2 and p-ERK2 in normal tendon tissue was relatively low (Figure 2). Western blot analyses revealed that ERK2 siRNA but not MS-siRNA treatment evidently inhibited ERK2 expression in adhesion tissue of the tendon adhesion model (Figure 2A,B). The phosphorylation of ERK levels were increased significantly in adhesion tissue of the tendon adhesion model in a time-dependent manner (Figure 2). ERK2 siRNA significantly reversed increased p-ERK2 levels in the adhesion tissues, similar to its effect on ERK2 expression (Figure 2A,C,D). On the contrary, the MS siRNA showed no obvious effect on the ERK or p-ERK2 levels (Figure 2B,C,D).

2.3. Effective Inhibition of Peritendinous Adhesions by Lentiviral-Mediated ERK2 siRNA

At 28 days after the surgery, the peritendinous adhesions at the repaired tendons were assessed by visual examination based on a scoring system. Obvious fibrous adhesions were observed between the repaired chicken tendons and the peritendinous tissues in the control and MS siRNA groups (Figure 3A,B). For the tendons treated with lentiviral-mediated ERK2 siRNA, fewer and weak adhesions were observed at the repaired sites (Figure 3C), which could also be separated easily. The scoring results suggest that ERK2 siRNA significantly inhibited the adhesion formation, compared with the control and MS siRNA groups (Figure 4).

Figure 2. Effect of ERK2 siRNA on ERK2 and p-ERK levels. ERK2 and phosphorylated ERK1/2 in normal tendon tissues and adhesion tissues from the control group, ERK2 siRNA group (A) and MS siRNA group (B) were detected by Western blotting using specific antibodies at the indicated weeks after surgical repair. Quantification results of the p-ERK2 (C) and ERK2 (D) bands are shown. Results are means \pm SD from three independent experiments. # means $p < 0.05$, ## means $p < 0.01$ vs. the T group, ** means $p < 0.01$ vs. the Control group.

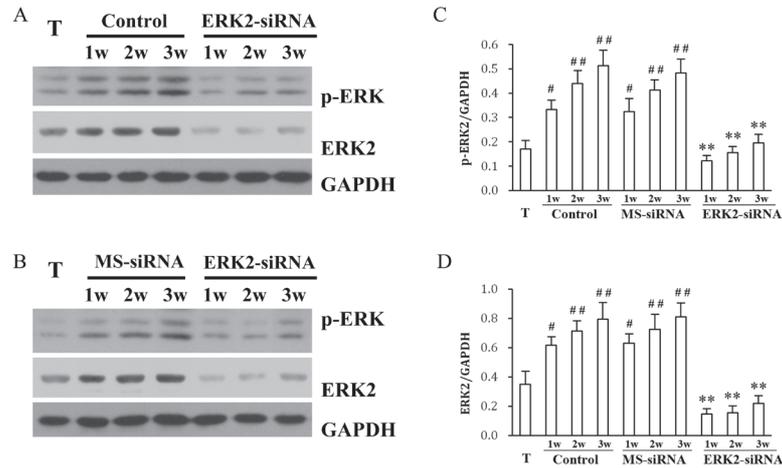
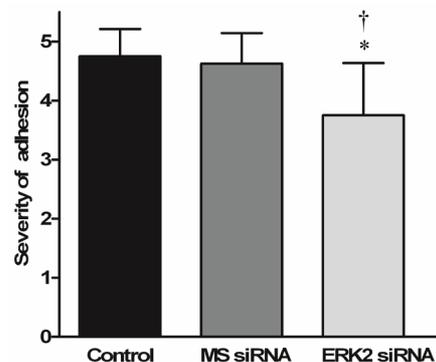


Figure 3. Effect of ERK2 siRNA on adhesion formation. Representative photographs of tendon adhesions in the control group (A), MS siRNA group (B) and ERK2 siRNA group (C) are shown.



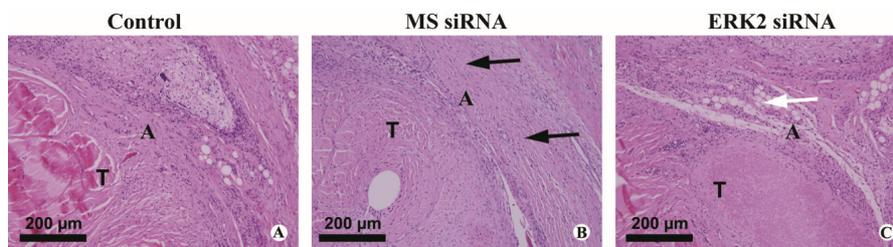
Figure 4. Effect of ERK2 siRNA on adhesion scores. Scores of tendon adhesions in the control group, MS siRNA group and ERK2 siRNA group are presented as mean \pm SD ($n = 8$). (* $p < 0.05$ vs. Control, $\dagger p < 0.05$ vs. MS siRNA).



Histological results of the hematoxylin and eosin (HE) stained sections of the repaired tendons are shown in Figure 5. Thick fibrous adhesion tissues developed at the repair sites in the control and MS

siRNA groups (Figure 5A and B). There was no clear peritendinous adhesion but a little scattered weak fiber formation around the repaired tendons in the ERK2 siRNA group (Figure 5C). The higher concentration of nuclei at the repair sites in the control and MS siRNA groups indicated evident proliferation of fibroblasts, but lower cell density was observed in the ERK2 siRNA treated chickens. Histological assessment of adhesion formation in all treatment groups are shown in Figure 6A. In comparison with the control group and MS siRNA group, the adhesions in the ERK2 siRNA group were significantly lower ($p < 0.05$) (Figure 6A). However, ERK2 siRNA treatment demonstrated no significant effect on the average scores of histological quality of tendon healing (Figure 6B).

Figure 5. Histological observation of repaired tendon sections at 28 days after tendon repair surgery. Representative microscopic photos of HE stained sections of tendon adhesions in the control group (A), MS siRNA group (B) and ERK2 siRNA group (C) are shown. A: Adhesion tissue; T: Tendon. White arrow indicates scattered weak fiber formation surrounding the tendon (T) while black arrows indicate the dense adhesion tissue.



2.4. Effect of ERK2 siRNA on Biomechanical Properties of Repaired Tendons

Compared to the control and MS siRNA group, a significant decrease in the ratio of work of flexion was observed in the ERK2 siRNA group, indicating reduced peritendinous adhesions (Figure 6C). By contrast, there was no significant difference between the breaking forces in the ERK2 siRNA or MS siRNA group and the control group (Figure 6D).

Figure 6. Histological evaluation and biomechanical analysis of repaired tendon sections at 28 days after tendon repair surgery. Scores of histological assessments of adhesions (A), histological quality of tendon healing (B), ratio of work of flexion (C) and the breaking forces of tendons (D) are presented as mean \pm SD ($n = 8$). Asterisk indicates * means $p < 0.05$ vs. Control and \dagger means $p < 0.05$ vs. MS siRNA

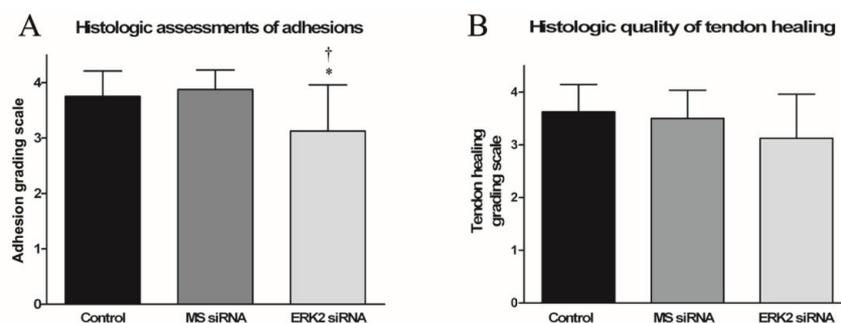
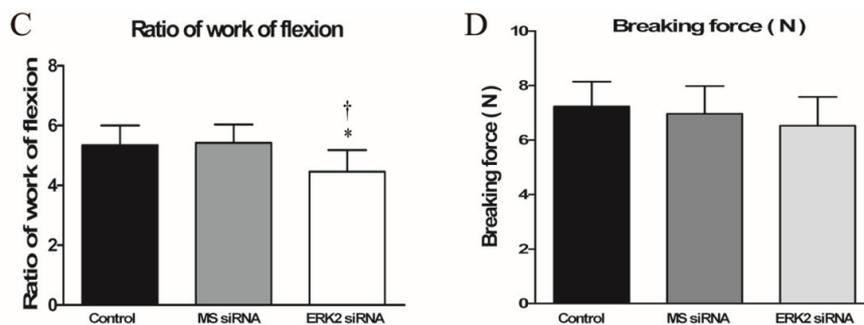


Figure 6. Cont.



Similar to the joint adhesions and abdominal adhesions, the formation of tendon adhesion involves fibroblast migration and proliferation and collagen expression that are mediated by TGF- β 1, IGF-1 and other growth factors [18,19]. The adhesion formation is closely related to the healing of the tendon. It is well established that both intrinsic healing and extrinsic healing, simultaneously, play roles in the healing process after tendon injury. The extrinsic healing is characterized by an evident inflammatory response followed by specialized fibroblast recruitment and proliferation. To inhibit the problematic tendon adhesion formation, and improve the healing quality of the tendon repaired, it is important to restrain the extrinsic healing and promote the intrinsic healing [20]. The migration of fibroblasts of the paratenon plays an important role in adhesion formation. Increased expression of TGF β 1 induces excessive fibroblast proliferation and reduced functionality. TGF β 1 also mediates collagen expression in fibroblasts through the SMAD and ERK pathways. The ERK pathway has been proven to contribute to SMAD-mediated signaling and Ras-dependent cell signaling in some cells [21,22]. In addition, recent researches have suggested that ERK2, but not ERK1, plays a dominant role in cell proliferation [23–25]. All aforementioned evidence indicates the ERK2 may act as a crucial mediator in fibroblast proliferation and collagen production, and the consequent formation of tendon adhesions. Moreover, our previous studies have confirmed that the siRNA targeting ERK2 mediated by lentivirus could effectively reduce the proliferation and collagen expression of rat joint adhesion tissue fibroblasts and decrease joint adhesion formation effectively [16,17]. Due to the similar characteristics of the joint adhesion and tendon adhesion, in the present study, the effect of this siRNA was investigated in the flexor tendon repair model.

Taking its advantages into consideration, such as extensively diverse target cells, high infection efficiency, a capacity to hold long sequences, stable expression due to gene integration into the host cells, and the weak immunogenicity and toxic response [26–31], the lentiviral-mediated siRNA delivery system was used in this study to evaluate the effects of ERK2 siRNA on the formation of flexor tendon adhesions. *In vivo* bioluminescent results show that the lentiviral system could efficiently express the siRNA targeting ERK2 in the peritendinous tissues of the checks, indicating a successful local delivery of siRNA, which avoids side effects caused by overall down-regulation of the target gene. Furthermore, owing to lentiviral integration, the ERK2 siRNA delivery system can steadily affect peritendinous fibroblasts for a long time, so that repeated administrations are not required. Taken together, it is indicated that the lentiviral system may be a safe and durable delivery system for gene regulation.

To assess the efficiency of this siRNA delivery system, a chicken model of flexor tendon repair was used. In the control group, evident peritendinous adhesions were found, suggesting that the model was

appropriately established. The histological results showed the lentiviral system effectively delivered the ERK2 siRNA to the repair sites and remarkably suppressed adhesion formation. The significantly attenuated formation of tendon adhesions by the administration of ERK2 siRNA was confirmed by biomechanical evaluation as well, in comparison with the control group and MS siRNA group. However, no statistically significant difference in breaking forces was noted between the three groups.

In summary, this is the first study to evaluate the effects of ERK2 siRNA in flexor tendon adhesions and the results demonstrate that ERK2 siRNA effectively inhibited the formation of tendon adhesions. Detailed studies should be carried out to provide more relevant evidence and information for the technique in tendon repair.

3. Experimental Section

3.1. *Lentiviral Vector Construction, Virus Production and Infection*

The pshRNA-H1-Luc lentivector purchased from System Biosciences was used in this study, in order to express target siRNA and luciferase originated from the copepod together. The siRNA which had been used to down-regulate ERK2 in rat in our previous work effectively [16], was used to inhibit ERK2 in chicken, since these ERK2 genes bear the identical sequence 5'-GTGATGAGCCTGTAGCTGA-3'. The MS negative control siRNA (5'-CGTTAGTTAGCAGTGAGCG-3') was also included. The synthesized oligonucleotide templates were annealed and inserted into the linear lentivector. The constructed vectors were transfected into 293TN producer cells with pPACK Packaging Plasmid Mix (System Biosciences, Mountain View, CA, USA) using LipofectamineTM 2000 (Invitrogen, Carlsbad, CA, USA) in accordance with the manufacturer's instructions. 48 hours later, the viral supernatants were collected, and cleared by centrifugation and 0.45 μ m PVDF membrane filter. Gradient dilution was used to determine viral titers.

3.2. *Animal Model*

All procedures and handling of the animals were carried out in accordance with the policies of Shanghai Jiao Tong University, School of Medicine and the National Institutes of Health. Leghorn chickens (1.5–2 kg each) were used for this study. They were anesthetized by intramuscular injection of ketamine hydrochloride (50 mg/kg). Then, sterile skin preparation and an elastic tourniquet were applied. A lateral skin incision was created on the proximal phalanx of the third toe. After incising the flexor tendon sheath, the FDP was isolated, transversely incised and then repaired using a modified Kessler tendon repair with 6-0 prolene suture (Ethicon Ltd., Edinburgh, UK). The animals were randomly assigned to three groups. In groups I and II, MS siRNA or ERK2 siRNA was injected around the repair site of the FDP, while no treatment was performed before wound closure in the control group. After skin closure, the extremity was immobilized in a weight-bearing splint.

3.3. *In Vivo Bioluminescence Assay*

Bioluminescence assays comprise a high-sensitivity and non-invasive technique for monitoring specific cellular and genetic activities in a living organism. At 14 and 28 days after surgical

manipulation, the luciferase expression and distribution in the individual chickens, in the ERK2 siRNA group, were measured using a Xenogen IVIS 50 Bioluminescence System (R&D Systems).

3.4. Western Blotting

The adhesion tissues from the three groups, and the normal tendon tissues were dissected and homogenized in RIPA lysis buffer. Protein concentrations were determined using BCA assay. Equal amounts of proteins were separated by SDS-PAGE, and transferred onto PVDF membranes (Millipore, MA, USA). The membranes were blocked in TBST containing 5% nonfat milk at room temperature for two hours and incubated with primary antibodies against ERK2 (1:400; Santa Cruz, Santa Cruz, CA, USA), p-ERK (1:500; Santa Cruz, Boston, MA, USA) or GAPDH (1:2000; Santa Cruz, Santa Cruz, MA, USA) at 4 °C overnight. Membranes were then incubated with corresponding HRP-conjugated secondary antibodies against mouse (1:4000; Cell Signaling Technology, Boston, MA, USA) or rabbit IgG (1:3000; Cell Signaling Technology, Boston, MA, USA) at room temperature for one hour. The bands were visualized with an enhanced chemiluminescence reagent (Amersham Biosciences, Sunnyvale, NJ, USA). ERK2 and pERK2 levels were quantified and normalized to GAPDH bands by densitometry.

3.5. Macroscopic Evaluation

Before sacrificing the animals, the repair site was visually examined for signs of inflammation or ulceration. The severity of peritendinous adhesion was evaluated by a scoring system [32]. To evaluate the severity of peritendinous adhesions, an adhesion scoring system was used to grade a particular area into grades of 1–5, based on the surgical findings: grade 1, no adhesion; grade 2, adhesion area can be separated by blunt dissection alone; grade 3, adhesion area less than or equal to 50% which required sharp dissection for separation; grade 4, 51%–97.5% adhesion area which required sharp dissection for separation; and grade 5, more than 97.5% of the adhesion area requiring sharp dissection for separation.

3.6. Histological Evaluation of Adhesion Tissues

The third toes were fixed in 4% paraformaldehyde for one day and then decalcified in 10% EDTA for one month at room temperature. Samples were dehydrated through increasing concentrations of ethanol and then paraffin embedded. Sections were cut in 4- μ m sagittal slices and stained with hematoxylin-eosin(HE). Histologic assessments of adhesions and tendon healing were performed [33]. Adhesions were quantified into four grades as follows: grade 4, severe (>66% of the tendon surface); grade 3, moderate (33%–66% of the tendon surface); grade 2, mild (<33% of the tendon surface); or grade 1, no adhesions. Tendon healing was quantified into four grades as follows: grade 4, poor (failed healing or massive overgrowth of granulation tissue); grade 3, fair (irregularly arranged and partly broken intratendinous collagen bundles); grade 2, good (intratendinous collagen bundles exhibited good repair, but the epitenon was interrupted by adhesions); or grade 1, excellent (good tendon continuity and smooth epitenon surface). These histological sections were evaluated under light microscopy (LEICA DM 4000 B) by two independent investigators blinded to the treatment.

3.7. Biomechanical Evaluation

To evaluate peritendinous adhesions and tendon healing, the work of flexion and the breaking force were both measured using a rheometer (Instron 5548, Instron, Norwood, MA, USA). To evaluate the work of flexion, the proximal end of the FDP tendon was fixed to a force gauge and the proximal phalanx of the toe was attached to a home-made device with the proximal interdigital joint fixed by stainless steel rods. The load (Newtons) and the displacement (mm) were measured when the FDP tendon was pulled at 20 mm/min until the angle of the distal interdigital joint was 40°. The work of flexion was then calculated by curve integration. To avoid individual variation, both the repaired and the intact tendons of both sides in each animal were evaluated and the ratio of repaired work of flexion vs. intact work of flexion was used as a parameter to determine the difference among different groups. To evaluate breaking force, the repaired chicken FDP tendons were harvested. The proximal and distal ends of the tendon were fixed to the force gauge of the rheometer. The tendon ends were pulled apart at a speed of 20 mm/min until rupture of the tendon occurred, and breaking force was recorded by the rheometer.

3.8. Statistical Analysis

Results are expressed as mean \pm standard deviation (SD). Statistical software SPSS 10.0 (Chicago, IL, USA) was used to analyze the data by one-way analysis of variance; $p < 0.05$ was considered significant.

4. Conclusions

The present study demonstrated that in a chicken flexor tendon repair model, the lentiviral-mediated siRNA targeting ERK2 substantially inhibited peritendinous adhesion formation. This finding may provide a potential novel therapeutic intervention which can be used to eliminate or reduce the formation of postoperative tendon adhesions in the treatment of hand tendon injury.

From the anti-adhesion results of lentiviral-mediated siRNA targeting ERK2, we can learn that the mitogen-activated protein kinase pathway may be involved in the peritendinous adhesion formation. However, further studies of this mechanism and the effect of inflammatory component are needed.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (81271999) and the Doctoral Innovation Fund of SJTU (BXJ201235). We would like to thank G.W. Liu for their excellent technical assistance.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Angermann, P.; Lohmann, M. Injuries to the hand and wrist. A study of 50,272 injuries. *J. Hand Surg. Br.* **1993**, *18*, 642–644.
2. Schöffl, V.; Heid, A.; Küpper, T. Tendon injuries of the hand. *World J. Orthop.* **2012**, *3*, 62–69.

3. Miller, J.A.; Ferguson, R.L.; Powers, D.L.; Burns J.W.; Shalaby, S.W. Efficacy of hyaluronic acid/nonsteroidal anti-inflammatory drug systems in preventing postsurgical tendon adhesions. *J. Biomed. Mater. Res.* **1997**, *38*, 25–33.
4. Özgenel, G.Y. The effects of a combination of hyaluronic and amniotic membrane on the formation of peritendinous adhesions after flexor tendon surgery in chickens. *J. Bone Joint Surg. Br.* **2004**, *86*, 301–307.
5. Menderes, A.; Mola, F.; Tayfur, V.; Vayvada, H.; Barutcu, A. Prevention of peritendinous adhesions following flexor tendon injury with Seprafilm. *Ann. Plast Surg.* **2004**, *53*, 560–564.
6. Hanff, G.; Hagberg, L. Prevention of restrictive adhesions with expanded polytetrafluoroethylene diffusible membrane following flexor tendon repair, an experimental study in rabbits. *J. Hand Surg. Am.* **1998**, *23*, 658–664.
7. Moran, S.L.; Ryan, C.K.; Orlando, G.S.; Pratt, C.E.; Michalko, K.B. Effects of 5-fluorouracil on flexor tendon repair. *J. Hand Surg. Am.* **2000**, *25*, 242–251.
8. Akasaka, T.; Nishida, J.; Imaeda, T.; Shimamura, T.; Amadio, P.C.; An, K.N. Effect of hyaluronic acid on the excursion resistance of tendon graft, a biomechanical in vitro study in a modified human model. *Clin. Biomech.* **2006**, *21*, 810–815.
9. Elbashir, S.M.; Harborth, J.; Lendeckel, W.; Yalcin, A.; Weber, K.; Tuschl, T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **2001**, *411*, 494–498.
10. Hannon, G.J. RNA interference. *Nature* **2002**, *418*, 244–251.
11. Sledz, C.A.; Williams, B.R.G. RNA interference in biology and disease. *Blood* **2005**, *106*, 787–794.
12. Wang, Z.; Rao, D.D.; Senzer, N.; Nemunaitis, J. RNA interference and cancer therapy. *Pharm. Res.* **2011**, *28*, 2983–2995.
13. Dykxhoorn, D.M.; Lieberman, J. The silent revolution, RNA interference as basic biology, research tool, and therapeutic. *Annu. Rev. Med.* **2005**, *56*, 401–423.
14. Khan, U.; Occleston, N.L.; Khaw, P.T.; McGrouther, D.A. Differences in proliferative rate and collagen lattice contraction between endotenon and synovial fibroblasts. *J. Hand Surg. Am.* **1998**, *23*, 266–273.
15. Nyska, M.; Nyska, A.; Rivlin, E.; Porat, S.; Pines, M.; Shoshan, S.; Nagler, A. Topically applied halofuginone, an inhibitor of collagen type I transcription, reduces peritendinous fibrous adhesions following surgery. *Connect. Tissue Res.* **1996**, *34*, 97–103.
16. Li, F.; Fan, C.; Cheng, T.; Jiang, C.; Zeng, B. Efficient inhibition of fibroblast proliferation and collagen expression by ERK2 siRNAs. *Biochem. Biophys. Res. Commun.* **2009**, *382*, 259–263.
17. Li, F.; Ruan, H.; Fan, C.; Zeng, B.; Wang, C.; Wang, X. Efficient inhibition of the formation of joint adhesions by ERK2 small interfering RNAs. *Biochem. Biophys. Res. Commun.* **2010**, *391*, 795–799.
18. Kashiwagi, K.; Mochizuki, Y.; Yasunaga, Y.; Ishida, O.; Deie, M.; Ochi, M. Effects of transforming growth factor-beta 1 on the early stages of healing of the Achilles tendon in a rat model. *Scand. J. Plast. Reconstr. Surg. Hand Surg.* **2004**, *38*, 193–197.
19. Abrahamsson, S.O.; Lohmander, S. Differential effects of insulin-like growth factor-I on matrix and DNA synthesis in various regions and types of rabbit tendons. *J. Orthop. Res.* **1996**, *14*, 370–376.
20. Wang, E.D. Tendon repair. *J. Hand Ther.* **1998**, *11*, 105–110.

21. Hayashida, T.; Decaestecker, M.; Schnaper, H.W. Cross-talk between ERK MAP kinase and Smad-signaling pathways enhances TGF-beta dependent responses in human mesangial cells. *FASEB J.* **2003**, *17*, 1576–1578.
22. Vantaggiato, C.; Formentini, I.; Bondanza, A.; Bonini, C.; Naldini, L.; Brambilla, R. ERK1 and ERK2 mitogen-activated protein kinases affect Ras-dependent cell signaling differentially. *J. Biol.* **2006**, *5*, doi:10.1186/jbiol38.
23. Frémin, C.; Ezan, F.; Boisselier, P.; Bessard, A.; Pagès, G.; Pouysségur, J.; Baffet, G. ERK2 but not ERK1 plays a key role in hepatocyte replication, an RNAi-mediated ERK2 knockdown approach in wild-type and ERK1 null hepatocytes. *Hepatology* **2007**, *45*, 1035–1045.
24. Lefloch, R.; Pouysségur, J.; Lenormand, P. Single and combined silencing of ERK1 and ERK2 reveals their positive contribution to growth signaling depending on their expression levels. *Mol. Cell Biol.* **2008**, *28*, 511–527.
25. Samuels, I.S.; Karlo, J.C.; Faruzzi, A.N.; Pickering, K.; Herrup, K.; Sweatt, J.D.; Saitta, S.C.; Landreth, G.E. Deletion of ERK2 mitogen-activated protein kinase identifies its key roles in cortical neurogenesis and cognitive function. *J. Neurosci.* **2008**, *28*, 6983–6995.
26. Sioud, M.; Sørensen, D.R. Cationic liposome-mediated delivery of siRNAs in adult mice. *Biochem. Biophys. Res. Commun.* **2003**, *312*, 1220–1225.
27. Houry, M.; Louis-Pence, P.; Escriou, V. Efficient new cationic liposome formulation for systemic delivery of small interfering RNA silencing tumor necrosis factor- α in experimental arthritis. *Arthritis Rheum.* **2006**, *54*, 1867–1877.
28. Lois, C.; Hong, E.J.; Pease, S.; Brown, E.J.; Baltimore, D. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* **2002**, *295*, 868–872.
29. Pfeifer, A.; Ikawa, M.; Dayn, Y.; Verma, I.M. Transgenesis by lentiviral vectors, Lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2140–2145.
30. Lai, Z.; Brady, R.O. Gene transfer into the central nervous system in vivo using a recombinant lentivirus vector. *J. Neurosci. Res.* **2002**, *67*, 363–371.
31. Yu, X.; Zhan, X.; D’Costa, J.; Tanavde, V.M.; Ye, Z.; Peng, T.; Malehorn, M.T.; Yang, X.; Civin, C.I.; Cheng, L. Lentiviral vectors with two independent internal promoters transfer high-level expression of multiple transgenes to human hematopoietic stem-progenitor cells. *Mol. Ther.* **2003**, *7*, 827–838.
32. Ishiyama, N.; Moro, T.; Ishihara, K.; Ohe, T.; Miura, T.; Konno, T.; Ohyama, T.; Kimura, M.; Kyomoto, M.; Nakamura, K.; Kawaguchi, H. The prevention of peritendinous adhesions by a phospholipid polymer hydrogel formed in situ by spontaneous intermolecular interactions. *Biomaterials* **2010**, *31*, 4009–4016.
33. Güdemez, E.; Ekşioğlu, F.; Korkusuz, P.; Aşan, E.; Gürsel, I.; Hasirci, V. Chondroitin sulfate-coated polyhydroxyethyl methacrylate membrane prevents adhesion in full-thickness tendon tears of rabbits. *J. Hand Surg. Am.* **2002**, *27*, 293–306.

Reprinted from *IJMS*. Cite as: Kushibiki, T.; Hirasawa, T.; Okawa, S.; Ishihara, M. Regulation of miRNA Expression by Low-Level Laser Therapy (LLLT) and Photodynamic Therapy (PDT). *Int. J. Mol. Sci.* **2013**, *14*, 13542-13558.

Review

Regulation of miRNA Expression by Low-Level Laser Therapy (LLLT) and Photodynamic Therapy (PDT)

Toshihiro Kushibiki *, Takeshi Hirasawa, Shinpei Okawa and Miya Ishihara

Department of Medical Engineering, National Defense Medical College 3-2 Namiki, Tokorozawa, Saitama 359-8513, Japan

* Author to whom correspondence should be addressed; E-Mail: toshi@ndmc.ac.jp;
Tel.: +81-4-2995-1211 (ext. 2241); Fax: +81-4-2996-5199.

Received: 31 May 2013; in revised form: 19 June 2013 / Accepted: 20 June 2013 /

Published: 27 June 2013

Abstract: Applications of laser therapy, including low-level laser therapy (LLLT), phototherapy and photodynamic therapy (PDT), have been proven to be beneficial and relatively less invasive therapeutic modalities for numerous diseases and disease conditions. Using specific types of laser irradiation, specific cellular activities can be induced. Because multiple cellular signaling cascades are simultaneously activated in cells exposed to lasers, understanding the molecular responses within cells will aid in the development of laser therapies. In order to understand in detail the molecular mechanisms of LLLT and PDT-related responses, it will be useful to characterize the specific expression of miRNAs and proteins. Such analyses will provide an important source for new applications of laser therapy, as well as for the development of individualized treatments. Although several miRNAs should be up- or down-regulated upon stimulation by LLLT, phototherapy and PDT, very few published studies address the effect of laser therapy on miRNA expression. In this review, we focus on LLLT, phototherapy and PDT as representative laser therapies and discuss the effects of these therapies on miRNA expression.

Keywords: low-level laser therapy (LLLT); phototherapy; photodynamic therapy (PDT); miRNA

1. Low-Level Laser Therapy (LLLT) and Its Effects on miRNA Expression

A laser (light amplification by stimulated emission of radiation) is a device that generates electromagnetic radiation that is relatively uniform in wavelength, phase and polarization. This technology was originally described by Maiman in 1960 in the form of a ruby laser [1]. The properties of lasers have allowed for numerous medical applications, including their use in surgery, activation of photodynamic agents and various ablative therapies in cosmetics, all of which are based on heat generated by the laser beam, in some cases, leading to tissue destruction [2–9]. These applications of lasers are considered “high-energy”, because of their intensities, which range from about 1–100 watt (W)/cm².

This paper will address another type of laser application, low-level laser therapy (LLLT), which elicits its effects through non-thermal means. This field was initiated by the work of Mester *et al.*, who in 1967 reported non-thermal effects of lasers on mouse hair growth [10]. In a subsequent study, the same group reported acceleration of wound healing and improvement in the post-wounding regeneration ability of muscle fibers using a 1 J/cm² ruby laser [11]. Since those early days, numerous *in vitro* and *in vivo* studies of LLLT in the context of regenerative medicine have demonstrated a wide variety of therapeutic effects, including reduction of pain, anti-inflammatory effects and wound healing. According to da Silva *et al.* [12], the types of laser most frequently used for wound healing and tissue repair are helium neon (He-Ne) lasers and diode lasers, including gallium-aluminum-arsenic (Ga-Al-As), arsenic-gallium (As-Ga) and indium-gallium-aluminum-phosphide (In-Ga-Al-P) lasers.

One of the most distinctive features of LLLT relative to other modalities is that the effects are mediated not through induction of thermal effects, but rather, through a process, still not clearly defined, called “photobiostimulation”. Because this effect of LLLT apparently does not depend on coherence, it is therefore possible to achieve photobiostimulation using non-laser light-generating devices, such as inexpensive light-emitting diode (LED) technology [13–17]. To date, several mechanisms of biological action have been proposed, although none have been clearly established. These include augmentation of cellular ATP levels [18–20], manipulation of inducible nitric oxide synthase (iNOS) activity [21–25], suppression of inflammatory cytokines, such as TNF-alpha [19,26–28], IL-1beta [28–30], IL-6 [28,31–34] and IL-8 [28,31,32,35], upregulation of growth factors, such as PDGF, IGF-1, NGF and FGF-2 [30,36–38], alteration of mitochondrial membrane potential [39–42], due to chromophores found in the mitochondrial respiratory chain [43–45], stimulation of protein kinase C (PKC) activation [46], manipulation of NF-kappaB activation [47], induction of reactive oxygen species (ROS) [48,49], modification of extracellular matrix components [50], inhibition of apoptosis [39], stimulation of mast cell degranulation [51] and upregulation of heat shock proteins [52]. We have also proposed that LLLT influences cell differentiation following laser stimulation [53–55].

Unfortunately, these effects have been demonstrated using a variety of laser devices in non-comparable models. To add to the confusion, dose-dependency seems to be confined to a very narrow range, and in numerous systems, the therapeutic effects disappear with increased dose. Consequently, only two studies of miRNA expression dynamics following LLLT have been reported to date, by Wang *et al.* [56] and Gu *et al.* [57]. With the exception of those studies, no data are currently available regarding the overall changes in the global expression of many hundreds of miRNAs following LLLT. Wang *et al.* [56] showed that LLLT increases the migration, proliferation and viability of rat

mesenchymal stem cells (MSCs) and, also, activates the expression of various miRNAs. Using a diode laser (wavelength: 635 nm, 0.5 J/cm²), they found that the proliferation rate and expression of cell cycle-associated genes increased in a time-dependent manner following LLLT treatment of MSCs. Microarray assays revealed subsets of miRNAs that were regulated by LLLT: 19 miRNAs were upregulated and 15 miRNAs were downregulated (Table 1); these dynamic changes were confirmed by quantitative real-time PCR.

Table 1. Aberrations in miRNA expression after low-level laser therapy (LLLT) to mesenchymal stem cells by using a diode laser (wavelength: 635 nm, 0.5 J/cm²) [56].

| Upregulation | Downregulation |
|-----------------|----------------|
| miR-30e * | |
| <u>miR-15b</u> | |
| miR-30b-5p | miR-204 * |
| miR-322 | miR-7a |
| miR-215 | miR-423 |
| miR-449a | miR-678 |
| miR-126 | miR-25 * |
| <u>miR-133b</u> | miR-327 |
| miR-21 * | miR-351 |
| miR-455 | <u>miR-23a</u> |
| miR-759 | miR-667 |
| miR-872 * | miR-770 |
| miR-29b | miR-324-3p |
| miR-192 | miR-30c-2 * |
| miR-219-1-3p | miR-758 |
| miR-301a | <u>miR-320</u> |
| miR-551b | miR-466c |
| <u>miR-224</u> | |
| <u>miR-193</u> | |

miRNAs expression confirmed by quantitative real-time PCR are indicated by underlining. The asterisk, * indicates the star-form of miRNA.

The most highly upregulated miRNA was miR-193. Gain- and loss-of-function experiments demonstrated that miR-193 levels regulate the proliferation of MSCs of both humans and rats; in particular, blockade of miR-193 repressed the MSCs proliferation induced by LLLT. However, this miRNA apparently does not affect apoptosis or differentiation. In addition, Wang *et al.* found that miR-193 regulated expression of cyclin-dependent kinase 2 (CDK2). Bioinformatic analyses and luciferase reporter assays revealed that inhibitor of growth family, member 5 (ING5), was the most likely target of miR-193 to functionally regulate proliferation and CDK2 expression; indeed, the mRNA and protein levels of ING5 are regulated by miR-193. Furthermore, inhibition of ING5 by small interfering RNA (siRNA) upregulated both MSC proliferation and the expression of CDK2. Another miRNA, miR-335, has been shown by others to regulate the proliferation and migration of MSCs [58], so it is likely to play an important role in MSC proliferation after LLLT. Moreover,

several studies have shown that LLLT also stimulates cell differentiation [53–55,59–75], and future work should reveal miRNAs specifically involved in mediating this effect.

Although some literature reported that tumor or apoptosis related miRNAs were induced by UV irradiation to cells [76–81], Gu *et al.* reported UV-phototherapy and its effect on miRNA expression [57]. They showed the effect of narrow-band ultraviolet B (NB-UVB) irradiation on miR-21 and -125b expression in psoriatic epidermis. Psoriasis is an inflammatory skin disease in which dysregulation of p63, a member of the p53 family that is crucial for skin development and maintenance, has been demonstrated [82–84]. Involvement of miR-203, miR-21 and miR-125b were implicated in the regulation of p63 or p53 in the pathogenesis of psoriasis. Skin biopsies from 12 psoriasis patients were collected before, during and after NB-UVB therapy. The p63 expression was not significantly affected, whereas NB-UVB phototherapy significantly decreased expression of miR-21 and increased miR-125b levels. Since NB-UVB phototherapy is commonly used in the treatment of psoriasis [85–87], those results indicate a complex mechanism of p63 regulation, which merits further investigation in order to achieve better long-term clinical improvement.

2. Photodynamic Therapy and Its Effects on miRNA Expression

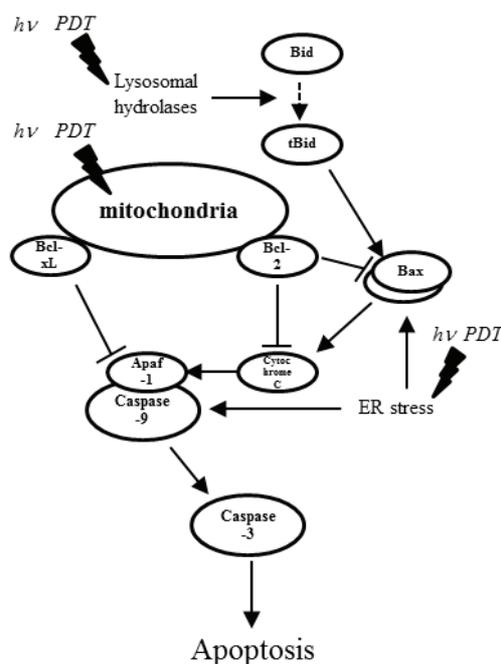
Photodynamic therapy (PDT), a class of laser therapy, is a photochemical modality approved for the treatment of various cancers and diseases in which neovascularization occurs [88,89]. The PDT process consists of injecting a photosensitizer, which selectively accumulates at the lesion site, followed by local irradiation of the tumor with light of an appropriate wavelength to activate a specific drug [90]. Irradiation leads to the generation of singlet oxygen and other reactive oxygen species (ROS) [91]. PDT is being considered not only as palliative therapy, but also as a treatment option for early-stage skin, lung, cervical and esophageal cancers, as well as basal-cell carcinomas. Currently, PDT has been approved for localized diseases and precancerous lesions, such as bladder cancers, pituitary tumors and glioblastomas [92,93]. Furthermore, numerous ongoing clinical studies have been designed to optimize the conditions for PDT; subsequently, PDT has been approved in several countries.

Upon absorption of one or more photons, the excited photosensitizer undergoes one of two possible reactions (type I or/and II) with a neighboring oxygen molecule, yielding ROS [94]. These ROS oxidize various cellular substrates, affecting cellular functions and resulting in cell death. The ROS that are produced during PDT destroy tumors by multiple mechanisms: in contrast to most conventional cytotoxic agents, which usually only trigger apoptotic cell death, PDT can cause cell death by necrosis and/or apoptosis.

The direct destruction of cancer cells (necrosis) by PDT is caused by irreversible damage to the plasma membrane and intracellular organelles, including the mitochondria, lysosomes, Golgi apparatus and endoplasmic reticulum (ER). The mechanisms of PDT-induced apoptosis have been described by many studies. Apoptosis, or programmed cell death, is one mechanism that mediates toxicity in the target tissue following PDT [95]. Apoptosis involves a cascade of molecular events leading to orderly cellular death without an inflammatory response [96–98]. The initiation of apoptosis involves a complex network of signaling pathways, both intrinsic and extrinsic to the individual cell, which are regulated, in part, by pro- and anti-apoptotic factors [96]. The initial damage can involve different molecules, ultimately leading to activation of specific death pathways. Mitochondria-localized

photosensitizers can cause immediate and light-dependent photodamage to mitochondrial components, such as the anti-apoptotic Bcl-2, Bcl-xL and the other apoptosis-related proteins, prompting the release of caspase-activating molecules [99]. Photosensitizers that accumulate in the lysosomes or mitochondria and which were excited by laser light can induce Bax-mediated caspase activation (Figure 1).

Figure 1. Representative signaling pathways of apoptosis induced by photodynamic therapy (PDT). Depending on the nature of the photosensitizer and its intracellular localization, the initial photodamage can involve different molecules, with the consequent activation of specific death pathways that converge on mitochondria. Mitochondria-localized photosensitizer can cause immediate and light-dependent photodamage to the anti-apoptotic Bcl-2 and Bcl-xL proteins, prompting the release of caspase-activating molecules. Lysosomal hydrolases and ER stress also induce Bax-mediated caspase activation.



Another important cellular factor induced by PDT and released from necrotic tumor cells is heat-shock protein 70 (Hsp70) [100]. Hsp70 is significantly induced after stress; when it remains within the cell, it chaperones unfolded proteins and prevents cell death by inhibiting the aggregation of cellular proteins. Hsp70 directly binds to the caspase-recruitment domain of apoptotic-protease activating factor 1 (Apaf-1), thereby preventing the recruitment of Apaf-1 oligomerization and association of Apaf-1 with procaspase 9. These properties not only enable intracellular Hsp70 to inhibit cancer-cell death by apoptosis, but also promote the formation of stable complexes with cytoplasmic tumor antigens. These antigens can then either be expressed at the cell surface or escape intact from dying necrotic cells to interact with antigen-presenting cells, thereby stimulating an anti-tumor immune response.

The mechanisms of cell death following PDT have been thoroughly summarized in the literature [95,101–104]. A better understanding of the molecular differences between apoptosis and

necrosis and identification of the crosstalk between these programs will certainly be crucial to the development of new PDT modalities aimed at increasing the efficiency of cancer-cell killing.

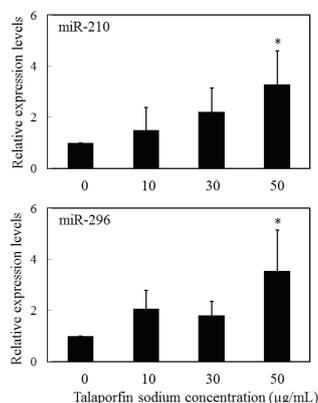
Another inherent consequence of PDT is local hypoxia, which can arise either directly, from oxygen consumption during treatment [105–107], or indirectly, from the destruction of tumor vasculature as a result of effective treatment [108,109]. Hypoxia is a major stimulus for angiogenesis, via its stabilization of the hypoxia-inducible factor-1 α (HIF-1 α) transcription factor [110,111]. HIF-1 is a heterodimeric complex of two helix-loop-helix proteins, HIF-1 α and HIF-1 β (ARNT). ARNT is constitutively expressed, whereas HIF-1 α is rapidly degraded under normoxic conditions. Hypoxia induces the stabilization of the HIF-1 α subunit, which, in turn, allows formation of the transcriptionally active protein complex. A number of HIF-1–responsive genes have been identified, including those encoding vascular endothelial growth factor (VEGF), erythropoietin and glucose transporter-1 [112,113]. Following PDT, increases in VEGF secretion and angiogenic responses stimulated via HIF-1 pathways have been documented *in vivo* [114–117]. VEGF induction can contribute to tumor survival and regrowth and, therefore, may represent one of the factors that prevent PDT from achieving its full tumoricidal potential. PDT has been considered for both palliative therapy and as an early treatment option for cancer. Numerous ongoing clinical studies have been designed to optimize PDT conditions. However, no standardized biological markers of cell death and PDT efficacy, other than cell viability itself, have been reported.

Human cancer is associated with changes in miRNA expression. The pattern of miRNA expression varies dramatically across tumor types, and miRNA profiles reflect the developmental lineage and differentiation state of a tumor [118]. miRNA is also likely to play critical roles in various aspects of hematopoiesis, including the differentiation of hematopoietic stem/progenitor cells, as well as in events that lead to hematological disorders. Nonetheless, very few miRNA expression patterns of specific diseases are available. Moreover, no profiles of miRNA expression after PDT have been reported. Cheng *et al.* found that inhibition of miR-95, -124, -125, -133, -134, -144, -150, -152, -187, -190, -191, -192, -193, -204, -211, -218, -220, -296 and -299 resulted in a decrease in cell growth, whereas inhibition of miR-21 and miR-24 profoundly increased cell growth in HeLa cells [119]. In addition, they identified miRNAs, whose expression increased levels of apoptosis (miR-7, -148, -204, -210, -216 and -296). Those data suggest that specific miRNAs are involved in the cell-death response. We have shown that a miRNA specific to apoptosis is expressed at increased levels in HeLa cells in response to PDT using talaporfin sodium as a photosensitizer [120]. Our study was the first to characterize miRNA expression levels following PDT. In our experiments, miR-210 and miR-296 expression levels increased significantly 1 h after PDT in cells treated with 50 $\mu\text{g}/\text{mL}$ talaporfin sodium, relative to the control group (*i.e.*, 0 $\mu\text{g}/\text{mL}$ talaporfin sodium), as shown in Figure 2. However, the expression levels of other miRNAs, e.g., miR-7, -148a, -204 and -216, were indistinguishable from those of the control group after PDT.

miR-210 is the miRNA most consistently stimulated under hypoxic conditions [121]. Because hypoxia and stabilization of intracellular HIF are inherent consequences of PDT [92], Giannakakis *et al.* investigated miR-210 expression in the context of its hypoxic effect, and they reported evidence for the involvement of the HIF signaling pathway in miR-210 regulation. To study the biological impacts of a partial or complete loss of miR-210 functions, they also identified the putative mRNA targets of miR-210. According to their report, miR-210 targets important regulators of transcription,

cell metabolism, differentiation and development, *i.e.*, processes that are critically affected by hypoxia [121]. The identification of key regulators of important cellular processes among miR-210 target mRNAs, as well as the high frequency of gene copy-number aberrations in tumors, underscore the involvement of miR-210 in oncogenesis and highlight miR-210 as a potential link between hypoxia and cell-cycle control in cancer cells.

Figure 2. Expression of miR-210 and miR-296 after PDT in HeLa cells. miR-210 and miR-296 expression levels were significantly increased 1 h after PDT (60 mW/cm², 90 s) in cells treated with 50 µg/mL talaporfin sodium relative to levels in the control group (*i.e.*, talaporfin sodium concentration of 0 µg/mL) (1×10^4 cells/well). The asterisk, * indicates $p < 0.05$, a significant difference between the relative expression levels of PDT-treated cells and non-PDT-treated cells. All experiments were performed four times independently. All data are expressed as the means \pm SD of four replicates from four experiments (Adapted from [120]).



Würdinger *et al.* reported a role for miR-296 in promoting angiogenesis in tumors [122], and in particular, they showed that VEGF alone is capable of increasing miR-296 expression levels. Their results revealed a feedback loop, wherein VEGF induces miR-296 expression, which targets the hepatocyte growth factor-regulated tyrosine kinase substrate (HGS), which, in turn, results in increased levels of VEGF receptor 2 and platelet-derived growth factor (PDGF) receptor β protein and, ultimately, in an increased response to VEGF. Because increased VEGF sensitivity of cancer cells is one of the inherent consequences of PDT [115], our results suggest that inhibition of miR-296 expression should improve PDT efficacy [120]. Our study also suggested that hypoxia induced by PDT induces miR-210 expression, followed by an increased expression of both VEGF and miR-296 [120]. Hence, we reported that miR-210 and miR-296 expression levels represent markers for the efficacy of talaporfin sodium-mediated PDT in cancer cells.

Furthermore, a recently published paper by Bach *et al.* described a comprehensive analysis of changes in miRNA levels following PDT, using polyvinylpyrrolidone hypericin (PVPH) as a photosensitizer, against A431 human epidermoid carcinoma cells [123]. That study was the first comprehensive analysis of changes in miRNA induced by PDT. Using microarray analysis, Bach *et al.* identified eight miRNAs that were significantly differentially expressed 5 hr after treatment, compared with baseline levels, and three miRNAs with more than two-fold differential expression that could be detected in one or two biological replicates. The verification of these results by quantitative real-time

PCR, including a detailed time course, revealed an up to 15-fold transient upregulation of miR-634, -1246 and -1290 relative to their basal levels (Table 2).

Table 2. Aberrations in miRNA expression after PDT to human epidermoid carcinoma cells (A431) by using polyvinylpyrrolidone hypericin (PVPH) [123].

| Upregulation | Downregulation |
|--------------|----------------|
| miR-1290 | miR-1260b |
| miR-634 | miR-720 |
| miR-1246 | miR-1260 |
| | miR-1280 |

In silico prediction of the targets of these miRNAs yielded numerous mRNAs encoding proteins, including the apoptotic protease activating factor-1 interacting protein and the BMI1 polycomb ring finger oncogene in the apoptosis/cell death category, cyclin-dependent kinase 20 and the cell division cycle 25 homolog C in the proliferation/cell cycle category, frizzled family receptor 3 and bone morphogenetic protein 4 in the cell signaling/adhesion category and the DNA excision repair protein ERCC-8 and peroxiredoxin-6 in the cell stress category. Although several studies have investigated the PDT-induced changes in the transcriptome and proteome, no comprehensive data are currently available regarding the effect of PDT on the miRNA transcriptome. Using a comprehensive microarray platform covering 1223 mature human miRNAs, Bach *et al.* did not observe up- or down-regulation by PDT of the miRNAs reported in our study (miR-210 and -296 [120]). This difference is likely attributable to the PDT conditions, such as cell type, photosensitizer and laser dose. Furthermore, the significant increase in the apoptosis-related miRNAs (3–4-fold increase) observed in our study was measured in a mixed population of cells, consisting predominantly of surviving cells [124]. Given these discrepancies, there is a need for additional experiments that might uncover additional miRNAs that are transiently regulated following photodynamic damage. It will be also of paramount interest to study miRNA-related cellular responses under explicitly non-lethal PDT conditions, as this approach could identify possible miRNA targets, whose manipulation might increase cells' sensitivity towards PDT.

Interestingly, Bach *et al.* also found that the incubation with the photosensitizer induced a slight to moderate increase in the expression of several miRNAs (*i.e.*, miR-1260b, -1260, -1280, -3182, -1290 and -1246), particularly at later time points [123]. Conversely, several miRNAs were transiently up-regulated by light-only treatment, especially at earlier time points (miR-1260b, -1260, -1280, -3182 and -1290). They concluded that the detailed functions of the increased expression of these miRNAs following apoptosis induced by PDT remain to be elucidated [123].

3. Conclusions

In this review, we focused on miRNA expression after LLLT and PDT. As mentioned above, only a few papers have been published regarding miRNA expression in this context, and those few reports discuss only a small number of laser therapy conditions. The ability of LLLT to induce growth-factor production, inhibition of inflammation, stimulation of angiogenesis, pain reduction and direct effects on stem cells suggests that there is an urgent need to combine this modality with regenerative medicine. PDT has been employed in the treatment of many tumor types, and its effectiveness as a

curative and palliative treatment is well documented, especially in the context of skin cancer. A detailed understanding of LLLT-, phototherapy- and PDT-related molecular mechanisms, including the specific effects on miRNA and protein expression, will provide an important source for new applications of laser therapy and for the development of individualized treatments.

Acknowledgments

This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI, Grant Number 25713009.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Maiman, T.H. Stimulated optical radiation in ruby. *Nature* **1960**, *187*, 493–494.
2. Khatri, K.A.; Mahoney, D.L.; McCartney, M.J. Laser scar revision: A review. *J. Cosmet. Laser Ther.* **2011**, *13*, 54–62.
3. Chung, S.H.; Mazur, E. Surgical applications of femtosecond lasers. *J. Biophotonics* **2009**, *2*, 557–572.
4. Zhao, Z.; Wu, F. Minimally-invasive thermal ablation of early-stage breast cancer: A systemic review. *Eur. J. Surg. Oncol.* **2010**, *36*, 1149–1155.
5. Siribumrungwong, B.; Noorit, P.; Wilasrusmee, C.; Attia, J.; Thakkinstian, A. A systematic review and meta-analysis of randomised controlled trials comparing endovenous ablation and surgical intervention in patients with varicose vein. *Eur. J. Vasc. Endovasc. Surg.* **2012**, *44*, 214–223.
6. Vuylsteke, M.E.; Mordon, S.R. Endovenous laser ablation: A review of mechanisms of action. *Ann. Vasc. Surg.* **2012**, *26*, 424–433.
7. Vogel, A.; Venugopalan, V. Mechanisms of pulsed laser ablation of biological tissues. *Chem. Rev.* **2003**, *103*, 577–644.
8. Casas, A.; di Venosa, G.; Hasan, T.; Al, B. Mechanisms of resistance to photodynamic therapy. *Curr. Med. Chem.* **2011**, *18*, 2486–2515.
9. Anand, S.; Ortel, B.J.; Pereira, S.P.; Hasan, T.; Maytin, E.V. Biomodulatory approaches to photodynamic therapy for solid tumors. *Cancer Lett.* **2012**, *326*, 8–16.
10. Mester, E.; Szende, B.; Gartner, P. The effect of laser beams on the growth of hair in mice. *Radiobiol. Radiother.* **1968**, *9*, 621–626.
11. Mester, E.; Spiry, T.; Szende, B.; Tota, J.G. Effect of laser rays on wound healing. *Am. J. Surg.* **1971**, *122*, 532–535.
12. Da Silva, J.P.; da Silva, M.A.; Almeida, A.P.; Lombardi Junior, I.; Matos, A.P. Laser therapy in the tissue repair process: A literature review. *Photomed. Laser Surg.* **2010**, *28*, 17–21.
13. Buravlev, E.A.; Zhidkova, T.V.; Vladimirov, Y.A.; Osipov, A.N. Effects of laser and led radiation on mitochondrial respiration in experimental endotoxic shock. *Lasers Med. Sci.* **2012**, *28*, 785–790.

14. De Sousa, A.P.; Santos, J.N.; dos Reis, J.A., Jr.; Ramos, T.A.; de Souza, J.; Cangussu, M.C.; Pinheiro, A.L. Effect of led phototherapy of three distinct wavelengths on fibroblasts on wound healing: A histological study in a rodent model. *Photomed. Laser Surg.* **2010**, *28*, 547–552.
15. Lev-Tov, H.; Brody, N.; Siegel, D.; Jagdeo, J. Inhibition of fibroblast proliferation *in vitro* using low-level infrared light-emitting diodes. *Dermatol. Surg.* **2012**, *39*, 422–425.
16. Nishioka, M.A.; Pinfildi, C.E.; Sheliga, T.R.; Arias, V.E.; Gomes, H.C.; Ferreira, L.M. Led (660 nm) and laser (670 nm) use on skin flap viability: Angiogenesis and mast cells on transition line. *Lasers Med. Sci.* **2012**, *27*, 1045–1050.
17. Pinheiro, A.L.; Soares, L.G.; Cangussu, M.C.; Santos, N.R.; Barbosa, A.F.; Silveira Junior, L. Effects of led phototherapy on bone defects grafted with mta, bone morphogenetic proteins and guided bone regeneration: A raman spectroscopic study. *Lasers Med. Sci.* **2012**, *27*, 903–916.
18. AlGhamdi, K.M.; Kumar, A.; Moussa, N.A. Low-level laser therapy: A useful technique for enhancing the proliferation of various cultured cells. *Lasers Med. Sci.* **2012**, *27*, 237–249.
19. Gao, X.; Xing, D. Molecular mechanisms of cell proliferation induced by low power laser irradiation. *J. Biomed. Sci.* **2009**, *16*, 4.
20. Tafur, J.; van Wijk, E.P.; van Wijk, R.; Mills, P.J. Biophoton detection and low-intensity light therapy: A potential clinical partnership. *Photomed. Laser Surg.* **2010**, *28*, 23–30.
21. Gavish, L.; Perez, L.S.; Reissman, P.; Gertz, S.D. Irradiation with 780 nm diode laser attenuates inflammatory cytokines but upregulates nitric oxide in lipopolysaccharide-stimulated macrophages: Implications for the prevention of aneurysm progression. *Lasers Surg. Med.* **2008**, *40*, 371–378.
22. Lindgard, A.; Hulten, L.M.; Svensson, L.; Soussi, B. Irradiation at 634 nm releases nitric oxide from human monocytes. *Lasers Med. Sci.* **2007**, *22*, 30–36.
23. Moriyama, Y.; Moriyama, E.H.; Blackmore, K.; Akens, M.K.; Lilge, L. *In vivo* study of the inflammatory modulating effects of low-level laser therapy on inos expression using bioluminescence imaging. *Photochem. Photobiol.* **2005**, *81*, 1351–1355.
24. Moriyama, Y.; Nguyen, J.; Akens, M.; Moriyama, E.H.; Lilge, L. *In vivo* effects of low level laser therapy on inducible nitric oxide synthase. *Lasers Surg. Med.* **2009**, *41*, 227–231.
25. Tuby, H.; Maltz, L.; Oron, U. Modulations of vegf and inos in the rat heart by low level laser therapy are associated with cardioprotection and enhanced angiogenesis. *Lasers Surg. Med.* **2006**, *38*, 682–688.
26. Fukuda, T.Y.; Tanji, M.M.; Silva, S.R.; Sato, M.N.; Plapler, H. Infrared low-level diode laser on inflammatory process modulation in mice: Pro- and anti-inflammatory cytokines. *Lasers Med. Sci.* **2012**, doi:10.1007/s10103-012-1231-z.
27. Oliveira, R.G.; Ferreira, A.P.; Cortes, A.J.; Aarestrup, B.J.; Andrade, L.C.; Aarestrup, F.M. Low-level laser reduces the production of tnf-alpha, ifn-gamma, and il-10 induced by ova. *Lasers Med. Sci.* **2013**, doi:10.1007/s10103-012-1262-5.
28. Yamaura, M.; Yao, M.; Yaroslavsky, I.; Cohen, R.; Smotrich, M.; Kochevar, I.E. Low level light effects on inflammatory cytokine production by rheumatoid arthritis synoviocytes. *Lasers Surg. Med.* **2009**, *41*, 282–290.

29. Aimbire, F.; Ligeiro de Oliveira, A.P.; Albertini, R.; Correa, J.C.; Ladeira de Campos, C.B.; Lyon, J.P.; Silva, J.A., Jr.; Costa, M.S. Low level laser therapy (lllt) decreases pulmonary microvascular leakage, neutrophil influx and il-1beta levels in airway and lung from rat subjected to lps-induced inflammation. *Inflammation* **2008**, *31*, 189–197.
30. Safavi, S.M.; Kazemi, B.; Esmaceli, M.; Fallah, A.; Modarresi, A.; Mir, M. Effects of low-level he-ne laser irradiation on the gene expression of il-1beta, tnf-alpha, ifn-gamma, tgf-beta, bfgf, and pdgf in rat's gingiva. *Lasers Med. Sci.* **2008**, *23*, 331–335.
31. Boschi, E.S.; Leite, C.E.; Saciura, V.C.; Caberlon, E.; Lunardelli, A.; Bitencourt, S.; Melo, D.A.; Oliveira, J.R. Anti-inflammatory effects of low-level laser therapy (660 nm) in the early phase in carrageenan-induced pleurisy in rat. *Lasers Surg. Med.* **2008**, *40*, 500–508.
32. Shiba, H.; Tsuda, H.; Kajiyama, M.; Fujita, T.; Takeda, K.; Hino, T.; Kawaguchi, H.; Kurihara, H. Neodymium-doped yttrium-aluminium-garnet laser irradiation abolishes the increase in interleukin-6 levels caused by peptidoglycan through the p38 mitogen-activated protein kinase pathway in human pulp cells. *J. Endod.* **2009**, *35*, 373–376.
33. Houreld, N.N.; Sekhejane, P.R.; Abrahamse, H. Irradiation at 830 nm stimulates nitric oxide production and inhibits pro-inflammatory cytokines in diabetic wounded fibroblast cells. *Lasers Surg. Med.* **2010**, *42*, 494–502.
34. Simunovic-Soskic, M.; Pezelj-Ribaric, S.; Brumini, G.; Glazar, I.; Grzic, R.; Miletic, I. Salivary levels of tnf-alpha and il-6 in patients with denture stomatitis before and after laser phototherapy. *Photomed. Laser Surg.* **2010**, *28*, 189–193.
35. Fushimi, T.; Inui, S.; Nakajima, T.; Ogasawara, M.; Hosokawa, K.; Itami, S. Green light emitting diodes accelerate wound healing: Characterization of the effect and its molecular basis *in vitro* and *in vivo*. *Wound Repair Regen.* **2012**, *20*, 226–235.
36. Saygun, I.; Karacay, S.; Serdar, M.; Ural, A.U.; Sencimen, M.; Kurtis, B. Effects of laser irradiation on the release of basic fibroblast growth factor (bfgf), insulin like growth factor-1 (igf-1), and receptor of igf-1 (igfbp3) from gingival fibroblasts. *Lasers Med. Sci.* **2008**, *23*, 211–215.
37. Schwartz, F.; Brodie, C.; Appel, E.; Kazimirsky, G.; Shainberg, A. Effect of helium/neon laser irradiation on nerve growth factor synthesis and secretion in skeletal muscle cultures. *J. Photochem. Photobiol. B* **2002**, *66*, 195–200.
38. Yu, W.; Naim, J.O.; Lanzafame, R.J. The effect of laser irradiation on the release of bfgf from 3t3 fibroblasts. *Photochem. Photobiol.* **1994**, *59*, 167–170.
39. Hu, W.P.; Wang, J.J.; Yu, C.L.; Lan, C.C.; Chen, G.S.; Yu, H.S. Helium-neon laser irradiation stimulates cell proliferation through photostimulatory effects in mitochondria. *J. Invest. Dermatol.* **2007**, *127*, 2048–2057.
40. Lan, C.C.; Wu, C.S.; Chiou, M.H.; Chiang, T.Y.; Yu, H.S. Low-energy helium-neon laser induces melanocyte proliferation via interaction with type iv collagen: Visible light as a therapeutic option for vitiligo. *Br. J. Dermatol.* **2009**, *161*, 273–280.
41. Wu, S.; Xing, D.; Gao, X.; Chen, W.R. High fluence low-power laser irradiation induces mitochondrial permeability transition mediated by reactive oxygen species. *J. Cell. Physiol.* **2009**, *218*, 603–611.

42. Zungu, I.L.; Hawkins Evans, D.; Abrahamse, H. Mitochondrial responses of normal and injured human skin fibroblasts following low level laser irradiation—An *in vitro* study. *Photochem. Photobiol.* **2009**, *85*, 987–996.
43. Karu, T. Photobiology of low-power laser effects. *Health Phys.* **1989**, *56*, 691–704.
44. Karu, T.I. Mitochondrial signaling in mammalian cells activated by red and near-ir radiation. *Photochem. Photobiol.* **2008**, *84*, 1091–1099.
45. Tiphlova, O.; Karu, T. Role of primary photoacceptors in low-power laser effects: Action of helium laser radiation on bacteriophage t4-escherichia coli interaction. *Lasers Surg. Med.* **1989**, *9*, 67–69.
46. Zhang, L.; Xing, D.; Zhu, D.; Chen, Q. Low-power laser irradiation inhibiting abeta25-35-induced pc12 cell apoptosis via pkc activation. *Cell. Physiol. Biochem.* **2008**, *22*, 215–222.
47. Aimbire, F.; Santos, F.V.; Albertini, R.; Castro-Faria-Neto, H.C.; Mittmann, J.; Pacheco-Soares, C. Low-level laser therapy decreases levels of lung neutrophils anti-apoptotic factors by a nf- κ dependent mechanism. *Int. Immunopharmacol.* **2008**, *8*, 603–605.
48. Kushibiki, T.; Hirasawa, T.; Okawa, S.; Ishihara, M. Blue laser irradiation generates intracellular reactive oxygen species in various types of cells. *Photomed. Laser Surg.* **2013**, *31*, 95–104.
49. Lipovsky, A.; Nitzan, Y.; Lubart, R. A possible mechanism for visible light-induced wound healing. *Lasers Surg. Med.* **2008**, *40*, 509–514.
50. Ignatieva, N.; Zakharkina, O.; Andreeva, I.; Sobol, E.; Kamensky, V.; Lunin, V. Effects of laser irradiation on collagen organization in chemically induced degenerative annulus fibrosus of lumbar intervertebral disc. *Lasers Surg. Med.* **2008**, *40*, 422–432.
51. Silveira, L.B.; Prates, R.A.; Novelli, M.D.; Marigo, H.A.; Garrocho, A.A.; Amorim, J.C.; Sousa, G.R.; Pinotti, M.; Ribeiro, M.S. Investigation of mast cells in human gingiva following low-intensity laser irradiation. *Photomed. Laser Surg.* **2008**, *26*, 315–321.
52. Coombe, A.R.; Ho, C.T.; Darendeliler, M.A.; Hunter, N.; Philips, J.R.; Chapple, C.C.; Yum, L.W. The effects of low level laser irradiation on osteoblastic cells. *Clin. Orthod. Res.* **2001**, *4*, 3–14.
53. Kushibiki, T.; Awazu, K. Controlling osteogenesis and adipogenesis of mesenchymal stromal cells by regulating a circadian clock protein with laser irradiation. *Inter. J. Med. Sci.* **2008**, *5*, 319–326.
54. Kushibiki, T.; Awazu, K. Blue laser irradiation enhances extracellular calcification of primary mesenchymal stem cells. *Photomed. Laser Surg.* **2009**, *27*, 493–498.
55. Kushibiki, T.; Tajiri, T.; Ninomiya, Y.; Awazu, K. Chondrogenic mrna expression in prechondrogenic cells after blue laser irradiation. *J. Photochem. Photobiol. B* **2010**, *98*, 211–215.
56. Wang, J.; Huang, W.; Wu, Y.; Hou, J.; Nie, Y.; Gu, H.; Li, J.; Hu, S.; Zhang, H. Microrna-193 pro-proliferation effects for bone mesenchymal stem cells after low-level laser irradiation treatment through inhibitor of growth family, member 5. *Stem Cells Dev.* **2012**, *21*, 2508–2519.
57. Gu, X.; Nylander, E.; Coates, P.J.; Nylander, K. Effect of narrow-band ultraviolet b phototherapy on p63 and microrna (mir-21 and mir-125b) expression in psoriatic epidermis. *Acta Derm. Venereol.* **2011**, *91*, 392–397.

58. Tome, M.; Lopez-Romero, P.; Albo, C.; Sepulveda, J.C.; Fernandez-Gutierrez, B.; Dopazo, A.; Bernad, A.; Gonzalez, M.A. Mir-335 orchestrates cell proliferation, migration and differentiation in human mesenchymal stem cells. *Cell Death Differ.* **2011**, *18*, 985–995.
59. Bouvet-Gerbettaz, S.; Merigo, E.; Rocca, J.P.; Carle, G.F.; Rochet, N. Effects of low-level laser therapy on proliferation and differentiation of murine bone marrow cells into osteoblasts and osteoclasts. *Lasers Surg. Med.* **2009**, *41*, 291–297.
60. Da Silva, A.P.; Petri, A.D.; Crippa, G.E.; Stuani, A.S.; Stuani, A.S.; Rosa, A.L.; Stuani, M.B. Effect of low-level laser therapy after rapid maxillary expansion on proliferation and differentiation of osteoblastic cells. *Lasers Med. Sci.* **2012**, *27*, 777–783.
61. Ebrahimi, T.; Moslemi, N.; Rokn, A.; Heidari, M.; Nokhbatolfoghahaie, H.; Fekrazad, R. The influence of low-intensity laser therapy on bone healing. *J. Dent.* **2012**, *9*, 238–248.
62. Fujimoto, K.; Kiyosaki, T.; Mitsui, N.; Mayahara, K.; Omasa, S.; Suzuki, N.; Shimizu, N. Low-intensity laser irradiation stimulates mineralization via increased bmps in mc3t3-e1 cells. *Lasers Surg. Med.* **2010**, *42*, 519–526.
63. Hou, J.F.; Zhang, H.; Yuan, X.; Li, J.; Wei, Y.J.; Hu, S.S. *In vitro* effects of low-level laser irradiation for bone marrow mesenchymal stem cells: Proliferation, growth factors secretion and myogenic differentiation. *Lasers Surg. Med.* **2008**, *40*, 726–733.
64. Kim, H.; Choi, K.; Kweon, O.K.; Kim, W.H. Enhanced wound healing effect of canine adipose-derived mesenchymal stem cells with low-level laser therapy in athymic mice. *J. Dermatol. Sci.* **2012**, *68*, 149–156.
65. Lin, F.; Josephs, S.F.; Alexandrescu, D.T.; Ramos, F.; Bogin, V.; Gammill, V.; Dasanu, C.A.; de Necochea-Campion, R.; Patel, A.N.; Carrier, E.; *et al.* Lasers, stem cells, and copd. *J. Transl. Med.* **2010**, *8*, 16.
66. Luo, L.; Sun, Z.; Zhang, L.; Li, X.; Dong, Y.; Liu, T.C. Effects of low-level laser therapy on ros homeostasis and expression of igf-1 and tgf-beta1 in skeletal muscle during the repair process. *Lasers Med. Sci.* **2013**, *28*, 725–734.
67. Medrado, A.P.; Soares, A.P.; Santos, E.T.; Reis, S.R.; Andrade, Z.A. Influence of laser photobiomodulation upon connective tissue remodeling during wound healing. *J. Photochem. Photobiol. B* **2008**, *92*, 144–152.
68. Nogueira, G.T.; Mesquita-Ferrari, R.A.; Souza, N.H.; Artilheiro, P.P.; Albertini, R.; Bussadori, S.K.; Fernandes, K.P. Effect of low-level laser therapy on proliferation, differentiation, and adhesion of steroid-treated osteoblasts. *Lasers Med. Sci.* **2012**, *27*, 1189–1193.
69. Renno, A.C.; McDonnell, P.A.; Parizotto, N.A.; Laakso, E.L. The effects of laser irradiation on osteoblast and osteosarcoma cell proliferation and differentiation *in vitro*. *Photomed. Laser Surg.* **2007**, *25*, 275–280.
70. Rosa, A.P.; de Sousa, L.G.; Regalo, S.C.; Issa, J.P.; Barbosa, A.P.; Pitol, D.L.; de Oliveira, R.H.; de Vasconcelos, P.B.; Dias, F.J.; Chimello, D.T.; *et al.* Effects of the combination of low-level laser irradiation and recombinant human bone morphogenetic protein-2 in bone repair. *Lasers Med. Sci.* **2012**, *27*, 971–977.
71. Saito, K.; Hashimoto, S.; Jung, H.S.; Shimono, M.; Nakagawa, K. Effect of diode laser on proliferation and differentiation of pc12 cells. *Bull. Tokyo Dent. Coll.* **2011**, *52*, 95–102.

72. Soleimani, M.; Abbasnia, E.; Fathi, M.; Sahraei, H.; Fathi, Y.; Kaka, G. The effects of low-level laser irradiation on differentiation and proliferation of human bone marrow mesenchymal stem cells into neurons and osteoblasts—An *in vitro* study. *Lasers Med. Sci.* **2012**, *27*, 423–430.
73. Song, S.; Zhou, F.; Chen, W.R. Low-level laser therapy regulates microglial function through src-mediated signaling pathways: Implications for neurodegenerative diseases. *J. Neuroinflammation* **2012**, *9*, 219.
74. Stein, A.; Benayahu, D.; Maltz, L.; Oron, U. Low-level laser irradiation promotes proliferation and differentiation of human osteoblasts *in vitro*. *Photomed. Laser Surg.* **2005**, *23*, 161–166.
75. Stein, E.; Koehn, J.; Sutter, W.; Wendtlandt, G.; Wanschitz, F.; Thurnher, D.; Baghestanian, M.; Turhani, D. Initial effects of low-level laser therapy on growth and differentiation of human osteoblast-like cells. *Wien. Klin. Wochenschr.* **2008**, *120*, 112–117.
76. Dziunycz, P.; Iotzova-Weiss, G.; Eloranta, J.J.; Lauchli, S.; Hafner, J.; French, L.E.; Hofbauer, G.F. Squamous cell carcinoma of the skin shows a distinct microRNA profile modulated by uv radiation. *J. Invest. Dermatol.* **2010**, *130*, 2686–2689.
77. Glorian, V.; Maillot, G.; Poles, S.; Iacovoni, J.S.; Favre, G.; Vagner, S. Hur-dependent loading of mirna risc to the mrna encoding the ras-related small gtpase rhob controls its translation during uv-induced apoptosis. *Cell Death Differ.* **2011**, *18*, 1692–1701.
78. Guo, L.; Huang, Z.X.; Chen, X.W.; Deng, Q.K.; Yan, W.; Zhou, M.J.; Ou, C.S.; Ding, Z.H. Differential expression profiles of micrnas in nih3t3 cells in response to UVB irradiation. *Photochem. Photobiol.* **2009**, *85*, 765–773.
79. Pothof, J.; Verkaik, N.S.; van IJcken, W.; Wiemer, E.A.; Ta, V.T.; van der Horst, G.T.; Jaspers, N.G.; van Gent, D.C.; Hoeijmakers, J.H.; Persengiev, S.P. MicroRNA-mediated gene silencing modulates the uv-induced DNA-damage response. *EMBO J.* **2009**, *28*, 2090–2099.
80. Tan, G.; Niu, J.; Shi, Y.; Ouyang, H.; Wu, Z.H. Nf-kappab-dependent microRNA-125b up-regulation promotes cell survival by targeting p38alpha upon ultraviolet radiation. *J. Biol. Chem.* **2012**, *287*, 33036–33047.
81. Tan, G.; Shi, Y.; Wu, Z.H. MicroRNA-22 promotes cell survival upon uv radiation by repressing pten. *Biochem. Biophys. Res. Commun.* **2012**, *417*, 546–551.
82. Gu, X.; Lundqvist, E.N.; Coates, P.J.; Thurffjell, N.; Wettersand, E.; Nylander, K. Dysregulation of tap63 mrna and protein levels in psoriasis. *J. Invest. Dermatol.* **2006**, *126*, 137–141.
83. Okuyama, R.; Ogawa, E.; Nagoshi, H.; Yabuki, M.; Kurihara, A.; Terui, T.; Aiba, S.; Obinata, M.; Tagami, H.; Ikawa, S. P53 homologue, p51/p63, maintains the immaturity of keratinocyte stem cells by inhibiting notch1 activity. *Oncogene* **2007**, *26*, 4478–4488.
84. Shen, C.S.; Tsuda, T.; Fushiki, S.; Mizutani, H.; Yamanishi, K. The expression of p63 during epidermal remodeling in psoriasis. *J. Dermatol.* **2005**, *32*, 236–242.
85. Menter, A.; Korman, N.J.; Elmets, C.A.; Feldman, S.R.; Gelfand, J.M.; Gordon, K.B.; Gottlieb, A.; Koo, J.Y.; Lebwohl, M.; Lim, H.W.; *et al.* Guidelines of care for the management of psoriasis and psoriatic arthritis: Section 5. Guidelines of care for the treatment of psoriasis with phototherapy and photochemotherapy. *J. Am. Acad. Dermatol.* **2010**, *62*, 114–135.
86. Ozawa, M.; Ferenczi, K.; Kikuchi, T.; Cardinale, I.; Austin, L.M.; Coven, T.R.; Burack, L.H.; Krueger, J.G. 312-nanometer ultraviolet b light (narrow-band UVB) induces apoptosis of t cells within psoriatic lesions. *J. Exp. Med.* **1999**, *189*, 711–718.

87. Schneider, L.A.; Hinrichs, R.; Scharffetter-Kochanek, K. Phototherapy and photochemotherapy. *Clin. Dermatol.* **2008**, *26*, 464–476.
88. Celli, J.P.; Spring, B.Q.; Rizvi, I.; Evans, C.L.; Samkoe, K.S.; Verma, S.; Pogue, B.W.; Hasan, T. Imaging and photodynamic therapy: Mechanisms, monitoring, and optimization. *Chem. Rev.* **2010**, *110*, 2795–2838.
89. Dolmans, D.E.; Fukumura, D.; Jain, R.K. Photodynamic therapy for cancer. *Nat. Rev. Cancer* **2003**, *3*, 380–387.
90. Verma, S.; Watt, G.M.; Mai, Z.; Hasan, T. Strategies for enhanced photodynamic therapy effects. *Photochem. Photobiol.* **2007**, *83*, 996–1005.
91. Buytaert, E.; Dewaele, M.; Agostinis, P. Molecular effectors of multiple cell death pathways initiated by photodynamic therapy. *Biochim. Biophys. Acta* **2007**, *1776*, 86–107.
92. Brown, S.B.; Brown, E.A.; Walker, I. The present and future role of photodynamic therapy in cancer treatment. *Lancet Oncol.* **2004**, *5*, 497–508.
93. Dougherty, T.J. An update on photodynamic therapy applications. *J. Clin. Laser Med. Surg.* **2002**, *20*, 3–7.
94. Tomioka, Y.; Kushibiki, T.; Awazu, K. Evaluation of oxygen consumption of culture medium and *in vitro* photodynamic effect of talaporfin sodium in lung tumor cells. *Photomed. Laser Surg.* **2010**, *28*, 385–390.
95. Oleinick, N.L.; Morris, R.L.; Belichenko, I. The role of apoptosis in response to photodynamic therapy: What, where, why, and how. *Photochem. Photochem. Photobiol. Sci.* **2002**, *1*, 1–21.
96. Danial, N.N.; Korsmeyer, S.J. Cell death: Critical control points. *Cell* **2004**, *116*, 205–219.
97. Ferri, K.F.; Kroemer, G. Organelle-specific initiation of cell death pathways. *Nat. Cell Biol.* **2001**, *3*, E255–E263.
98. Hengartner, M.O. The biochemistry of apoptosis. *Nature* **2000**, *407*, 770–776.
99. Piette, J.; Volanti, C.; Vantieghem, A.; Matroule, J.Y.; Habraken, Y.; Agostinis, P. Cell death and growth arrest in response to photodynamic therapy with membrane-bound photosensitizers. *Biochem. Pharmacol.* **2003**, *66*, 1651–1659.
100. Helbig, D.; Simon, J.C.; Paasch, U. Photodynamic therapy and the role of heat shock protein 70. *Int. J. Hyperthermia* **2011**, *27*, 802–810.
101. Matroule, J.Y.; Volanti, C.; Piette, J. Nf-kappab in photodynamic therapy: Discrepancies of a master regulator. *Photochem. Photobiol.* **2006**, *82*, 1241–1246.
102. Agostinis, P.; Buytaert, E.; Breysens, H.; Hendrickx, N. Regulatory pathways in photodynamic therapy induced apoptosis. *Photochem. Photobiol. Sci.* **2004**, *3*, 721–729.
103. Dewaele, M.; Verfaillie, T.; Martinet, W.; Agostinis, P. Death and survival signals in photodynamic therapy. *Methods Mol. Biol.* **2010**, *635*, 7–33.
104. Kessel, D.; Oleinick, N.L. Photodynamic therapy and cell death pathways. *Methods Mol. Biol.* **2010**, *635*, 35–46.
105. Chen, Q.; Huang, Z.; Chen, H.; Shapiro, H.; Beckers, J.; Hetzel, F.W. Improvement of tumor response by manipulation of tumor oxygenation during photodynamic therapy. *Photochem. Photobiol.* **2002**, *76*, 197–203.

106. Henderson, B.W.; Busch, T.M.; Vaughan, L.A.; Frawley, N.P.; Babich, D.; Sosa, T.A.; Zollo, J.D.; Dee, A.S.; Cooper, M.T.; Bellnier, D.A.; *et al.* Photofrin photodynamic therapy can significantly deplete or preserve oxygenation in human basal cell carcinomas during treatment, depending on fluence rate. *Cancer Res.* **2000**, *60*, 525–529.
107. Sitnik, T.M.; Hampton, J.A.; Henderson, B.W. Reduction of tumour oxygenation during and after photodynamic therapy *in vivo*: Effects of fluence rate. *Br. J. Cancer* **1998**, *77*, 1386–1394.
108. Engbrecht, B.W.; Menon, C.; Kachur, A.V.; Hahn, S.M.; Fraker, D.L. Photofrin-mediated photodynamic therapy induces vascular occlusion and apoptosis in a human sarcoma xenograft model. *Cancer Res.* **1999**, *59*, 4334–4342.
109. Fingar, V.H.; Kik, P.K.; Haydon, P.S.; Cerrito, P.B.; Tseng, M.; Abang, E.; Wieman, T.J. Analysis of acute vascular damage after photodynamic therapy using benzoporphyrin derivative (bpd). *Br. J. Cancer* **1999**, *79*, 1702–1708.
110. 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.
111. Semenza, G.L. Hypoxia-inducible factors in physiology and medicine. *Cell* **2012**, *148*, 399–408.
112. Forsythe, J.A.; Jiang, B.H.; Iyer, N.V.; Agani, F.; Leung, S.W.; Koos, R.D.; Semenza, G.L. Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1. *Mol. Cell. Biol.* **1996**, *16*, 4604–4613.
113. Takenaga, K. Angiogenic signaling aberrantly induced by tumor hypoxia. *Front. Biosci.* **2011**, *16*, 31–48.
114. Deininger, M.H.; Weinschenk, T.; Morgalla, M.H.; Meyermann, R.; Schluesener, H.J. Release of regulators of angiogenesis following hypocrellin-a and -b photodynamic therapy of human brain tumor cells. *Biochem. Biophys. Res. Commun.* **2002**, *298*, 520–530.
115. Ferrario, A.; von Tiehl, K.F.; Rucker, N.; Schwarz, M.A.; Gill, P.S.; Gomer, C.J. Antiangiogenic treatment enhances photodynamic therapy responsiveness in a mouse mammary carcinoma. *Cancer Res.* **2000**, *60*, 4066–4069.
116. Jiang, F.; Zhang, Z.G.; Katakowski, M.; Robin, A.M.; Faber, M.; Zhang, F.; Chopp, M. Angiogenesis induced by photodynamic therapy in normal rat brains. *Photochem. Photobiol.* **2004**, *79*, 494–498.
117. Schmidt-Erfurth, U.; Schlotzer-Schrehard, U.; Cursiefen, C.; Michels, S.; Beckendorf, A.; Naumann, G.O. Influence of photodynamic therapy on expression of vascular endothelial growth factor (vegf), vegf receptor 3, and pigment epithelium-derived factor. *Invest. Ophthalmol. Vis. Sci.* **2003**, *44*, 4473–4480.
118. 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.
119. Cheng, A.M.; Byrom, M.W.; Shelton, J.; Ford, L.P. Antisense inhibition of human mirnas and indications for an involvement of mirna in cell growth and apoptosis. *Nucleic Acids Res.* **2005**, *33*, 1290–1297.
120. Kushibiki, T. Photodynamic therapy induces microRNA-210 and -296 expression in HeLa cells. *J. Biophotonics* **2010**, *3*, 368–372.

121. Giannakakis, A.; Sandaltzopoulos, R.; Greshock, J.; Liang, S.; Huang, J.; Hasegawa, K.; Li, C.; O'Brien-Jenkins, A.; Katsaros, D.; Weber, B.L.; *et al.* Mir-210 links hypoxia with cell cycle regulation and is deleted in human epithelial ovarian cancer. *Cancer Biol. Ther.* **2008**, *7*, 255–264.
122. Wurdinger, T.; Tannous, B.A.; Saydam, O.; Skog, J.; Grau, S.; Soutschek, J.; Weissleder, R.; Breakefield, X.O.; Krichevsky, A.M. Mir-296 regulates growth factor receptor overexpression in angiogenic endothelial cells. *Cancer Cell* **2008**, *14*, 382–393.
123. 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.
124. Sato, M.; Kubota, N.; Inada, E.; Saitoh, I.; Ohtsuka, M.; Nakamura, S.; Sakurai, T.; Watanabe, S. Hela cells consist of two cell types, as evidenced by cytochemical staining for alkaline phosphatase activity: A possible model for cancer stem cell study. *Adv. Stem. Cell* **2013**, doi:10.5171/2013.208514.

Reprinted from *IJMS*. Cite as: Richter, H.; Randau, L.; Plagens, A. Exploiting CRISPR/Cas: Interference Mechanisms and Applications. *Int. J. Mol. Sci.* **2013**, *14*, 14518-14531.

Review

Exploiting CRISPR/Cas: Interference Mechanisms and Applications

Hagen Richter, Lennart Randau and André Plagens *

Prokaryotic Small RNA Biology, Max Planck Institute for Terrestrial Microbiology,
Karl-von-Frisch-Straße 10, 35043 Marburg, Germany;

E-Mails: hagen.richter@mpi-marburg.mpg.de (H.R.); lennart.randau@mpi-marburg.mpg.de (L.R.)

* Author to whom correspondence should be addressed; E-Mail: andre.plagens@mpi-marburg.mpg.de;
Tel.: +49-0-6421-178-650; Fax: +49-0-6421-178-599.

Received: 31 May 2013; in revised form: 26 June 2013 / Accepted: 1 July 2013 /

Published: 12 July 2013

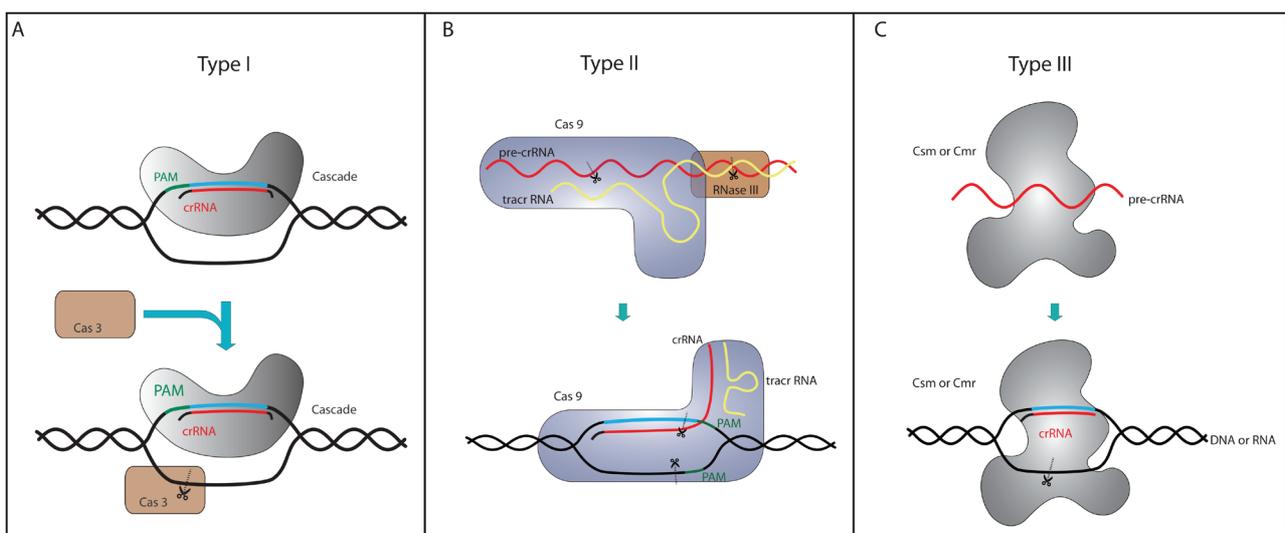
Abstract: The discovery of biological concepts can often provide a framework for the development of novel molecular tools, which can help us to further understand and manipulate life. One recent example is the elucidation of the prokaryotic adaptive immune system, clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) that protects bacteria and archaea against viruses or conjugative plasmids. The immunity is based on small RNA molecules that are incorporated into versatile multi-domain proteins or protein complexes and specifically target viral nucleic acids via base complementarity. CRISPR/Cas interference machines are utilized to develop novel genome editing tools for different organisms. Here, we will review the latest progress in the elucidation and application of prokaryotic CRISPR/Cas systems and discuss possible future approaches to exploit the potential of these interference machineries.

Keywords: CRISPR; crRNA; Cas9; Cascade; interference; genome editing; RGEN; TALEN; ZNF

1. Introduction

The clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems are found in many bacteria and nearly all archaea and constitute an adaptive immune system that recognizes and prevents viral attacks. A breakthrough for the basic understanding of the immunity mechanism was achieved by Barrangou and coworkers, who could show that *Streptococcus thermophilus* can acquire resistance against a bacteriophage by integrating a genome fragment of an infectious virus into its CRISPR locus [1]. A CRISPR cluster is a genomic DNA element that consists of a series of short repeat sequences (typically 24–37 bp) that are separated by unique spacer sequences of similar length [2]. These sequences are often fragments derived from a viral genome, illustrating the genetic memory of previous infections [3,4]. The second part of the CRISPR/Cas machinery is encoded in the Cas genes, and the Cas proteins fulfill essential functions within the immunity mechanism.

Figure 1. The interference step in the three clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) types. Indicated for all three systems are the targeted DNA region (blue), the targeting crRNA (red) and the protospacer adjacent motif (PAM, green). In type I systems (A), the invading DNA is recognized by the Cascade:crRNA complex. The PAM motif promotes the identification of the foreign DNA. Subsequently, the nuclease, Cas3, is recruited and degrades the target DNA. Type II systems (B) require only Cas9 for interference and do not rely on a multi-protein complex. Complex formation of Cas9, tracrRNA and pre-crRNA enables RNase III to mature crRNAs. The resulting complex of Cas9 and the tracrRNA:crRNA duplex recognizes the invading nucleic acid, and the Cas9 nuclease generates blunt-ended cleavage of both DNA strands. In type III systems (C), a multi-protein complex (Csm or Cmr) or Cas6 processes pre-crRNA into mature crRNA. The complex-bound crRNA recognizes invading DNA (Csm) or RNA (Cmr), resulting in target degradation.



The CRISPR/Cas defense response is classified into three main stages, which are conserved throughout all CRISPR systems. In the first stage, termed adaptation, the injected viral DNA is recognized, and a fragment of this DNA inserted as a new spacer into a host CRISPR array. A short

conserved sequence (2 to 5 nt), called the protospacer adjacent motif (PAM), flanks the spacer sequence in the viral genome (termed protospacer) and, thus, determines the targets of most CRISPR/Cas systems [5,6]. A new spacer is always integrated at the AT-rich leader site of a CRISPR, which is proposed to contain specific sequence elements that direct spacer DNA addition [7,8]. The detailed mechanism of spacer acquisition is only partially understood, but the highly conserved proteins, Cas1 and Cas2, were identified as key players in this process [9,10]. The second stage is characterized by the transcription of a CRISPR cluster into a long precursor-crRNA (pre-crRNA). Transcription is controlled by sequences in the leader region that contains promoter elements and possible binding sites for regulatory proteins [11,12]. In many systems, the long pre-crRNA transcript is processed by a Cas endonuclease (Cas6) into short crRNAs [13–17]. A typical crRNA generated by Cas6 cleavage contains the complete spacer sequence flanked by an 8 nt 5'-hydroxyl repeat tag and a 2'-3' cyclic phosphate repeat end [17,18]. In some cases, 3' termini of mature crRNAs are trimmed further, but the mechanism is unknown. In the final stage, the interference reaction, these mature crRNAs are incorporated into a larger Cas protein complex and used to target and degrade the viral DNA during a repeated attack (Figure 1) [19–21]. In the next section, we will focus on the molecular details of different interference mechanisms with diverse sets of Cas proteins and multi-subunit complexes. These crRNA/Cas protein complexes exhibit a large potential for the development of genetic tools.

2. Three Strategies to Cope with Viruses

CRISPR/Cas systems are identified in diverse bacterial and archaeal species, often living in the most extreme ecological niches. Some basic principles of the CRISPR/Cas systems are conserved within all prokaryotes, but the plethora of identified Cas protein families reflects the divergence of mechanistic details during evolution. The continuous co-evolution of viruses and their hosts led to the emergence of anti-CRISPR measures in viruses [22], which might explain the necessity for CRISPR/Cas diversification. Computational studies classified all CRISPR/Cas systems into three types and at least ten subtypes [23–25].

Type I CRISPR/Cas systems are found in both bacteria and archaea and are comprised of six different subtypes (subtypes I-A to I-F). The essential and significantly conserved marker protein in the interference reaction is Cas3, which contains a HD phosphohydrolase domain and a DExH-like helicase domain [24,25]. Both domains are also found to be encoded separately by two discrete genes. These two domains have been shown to unwind dsDNA (helicase domain) and cleave ssDNA (HD nuclease domain), depending on ATP and Mg^{2+} ions [26–28]. Cas3 interacts with a complex of different Cas proteins that bind and deliver the crRNA (Figure 1A). This complex is termed Cascade (CRISPR-associated complex for antiviral defense) and is best studied for the type I-E system from *Escherichia coli* [18,19]. The type I-E Cascade was shown to bind the short crRNAs, which are utilized to recognize DNA target sequences that are complementary to the spacer sequence within the crRNA. Subsequently, Cascade recruits Cas3 to degrade the targeted viral DNA molecule, relying on negatively supercoiled DNA [29,30]. The first 6–12 nt of the crRNA spacer are most important for target binding and are termed the seed sequence [20,21]. The point mutation of single seed sequence nucleotides results in drastic binding defects. However, mismatches in the crRNA spacer sequence

following the seed sequence are tolerated, and the mutated crRNA can still be bound [31]. To avoid that the Cascade:Cas3 complex degrades the host genomic encoded CRISPR cluster, it has to be ensured that the 5' terminal tag of the crRNA and the PAM sequence located upstream of the viral protospacer do not form base pairs. To discriminate between self- and non-self- targets, Cascade screens and specifically binds a PAM sequence, which results in helical destabilization and strand invasion of the matching seed sequence [32,33]. The PAM sequence for type I systems is typically 2–3 bases long and can differ between different subtypes and even organisms. The I-E Cascade complex has a size of 405 kDa and is composed of the five subunits, Cas6e, Cse1, Cse2, Cas7 and Cas5. Cas7 and Cas5 tightly bind and protect the crRNA from degradation [19], whereas Cse1 and Cse2 were shown to be nucleic acid-binding proteins that preferentially interact with the DNA target [34,35].

The targeting complex facilitates base pairing of the crRNA with the complementary DNA strand and additional displacement of the non-complementary strand to produce a so-called R-loop structure [18,36]. Using cryo-electron microscopy, the overall structure of the I-E Cascade complex was resolved and revealed a general outline that is often described as a seahorse-like shape [18,20]. The conservation of the crRNA-binding subunits, Cas7 and Cas5, as well as Cas3, throughout all type I subtypes, suggest structural and functional similarities of Cascade, but further biochemical and structural data are required to compare complexes and the interference reaction in different prokaryotic families in detail. The related I-A Cascade complex shows a conserved Cas7, Cas5 and crRNA assembly platform, which might recruit Cas3 and additional subtype-specific proteins to form the active interference machinery [36,37]. Plasmid-based interference assays were established that could show *in vivo* interference activity for I-A and I-B Cascade modules [38,39].

Type II systems have only been found in bacterial genomes and are characterized by a distinct minimal set of *cas* genes [24,25]. In these systems, the large multifunctional protein, Cas9, is involved in both the maturation of crRNAs and in the subsequent interference reaction [40]. The processing of crRNAs is dependent on a trans-activating crRNA (tracrRNA) encoded in the vicinity of CRISPR loci and containing a 25 nt long stretch that is complementary to the crRNA repeat sequence (Figure 1B) [41]. The comparison of several tracrRNA molecules did not identify any highly conserved sequence or structure elements other than the anti-repeat sequence [42]. Cas9 facilitates the base pairing of tracrRNA and pre-crRNA, which form a RNA duplex that is then targeted by the host endonuclease, RNase III. Cleavage of this duplex by RNase III generates mature crRNAs with 20 nt spacer-derived 5'-tags and 19–22 nt repeat-derived 3'-tags [41,43]. In the interference step, the cleavage of target dsDNA requires not only crRNA and Cas9, but also the presence of tracrRNA. Cas9 cleaves the DNA strand complementary to the crRNA with a McrA/HNH nuclease domain and the non-complementary strand with a RuvC-like (RNase H fold) domain in the presence of Mg²⁺ ions [43]. The interference against the viral DNA requires a conserved 5 nt-long PAM sequence (NGGNG), located immediately downstream of the protospacer [5,44]. The precise DNA cleavage site was identified 3 nt upstream of the PAM for the complementary strand, whereas the non-complementary DNA strand is cleaved at additional sites within three to eight base pairs upstream of the PAM, producing blunt-ended cleavage products [43,45].

The two known type III systems (type III-A and type III-B) are predominantly found in archaeal genomes [24,25], and interestingly, type III-B systems are only found in combination with one or more other CRISPR subtypes. Type III systems encode the CRISPR-specific endoribonuclease, Cas6, and

the subtype-specific Cas10 protein that is very likely involved in target interference (Figure 1C). Similar to Cas3 proteins of type I systems, Cas10 encodes a HD nuclease domain that is proposed to have similar function in target degradation [24,25]. The type III-A system of *Staphylococcus epidermidis* contains five Csm proteins and was shown to target DNA [46]. DNA targeting by this system does not require a specific PAM sequence, but sequences complementary to the 8 nt 5'-tag of the crRNA are not targeted by this system [47]. In the type III-B system of *Pyrococcus furiosus*, Cas6 is not an integral part of the interference complex after the crRNA processing, but the 8-nt 5' repeat tag serves as an anchor for the assembly of a six protein (Cmr1–Cmr6) ribonucleoprotein interference complex. A similar Cmr complex with seven proteins (Cmr1–Cmr7) was identified for *Sulfolobus solfataricus* and shown to endonucleolytically cleave invading RNA at UA dinucleotides [48]. Targeting of RNA was shown to be PAM-independent for both investigated Cmr complexes. The crystal structure of a Cmr2–Cmr3 complex revealed a conserved RNA binding surface of Cmr3, which is reminiscent of the Cas6 RNA interaction surface and illustrates the RNA binding ability of this Cmr subunit [49,50]. Notably, these two interference complexes differ from all other investigated subtypes, as they specifically target RNA and not DNA [11,51]. However, recently, it could be demonstrated *in vivo* that Cmr proteins can target also plasmid DNA in a PAM-independent manner [52].

3. CRISPR Systems as Genome Editing Tools

The application of the diverse CRISPR/Cas systems as genetic tools offers great potential. Initially, before the discovery of Cas protein functions, the diversity of CRISPR sequences was mainly utilized in a powerful method to rapidly identify closely related bacterial strains (e.g., *Mycobacterium tuberculosis*). This genotyping method is known as spacer oligonucleotide typing (spoligotyping) [53,54]. The investigation of different Cas protein activities introduced them as diverse components for genetic tool development. One example is the pre-crRNA processing enzyme, Cas6f (previously termed Csy4), which was shown to be useful for predictable gene expression. In these approaches, the Cas6f cleavage site sequence of a CRISPR repeat can be fused to a gene of interest. Processing of this site by Cas6f can then be used to physically separate genetic elements (e.g., untranslated regions, regulatory elements, ribosome binding sites) at the mRNA level. Therefore, transcript impedances (e.g., secondary structures) can be reduced, which results in a more predictable gene expression pattern [55]. Additionally, the high substrate affinity of Cas6f was used to create a specific RNA-binding bait protein. This construct can be used in high-throughput RNA affinity purification protocols to isolate RNA molecules with a 16 nt hairpin sequence derived from the 3'-terminal crRNA repeat tag [56].

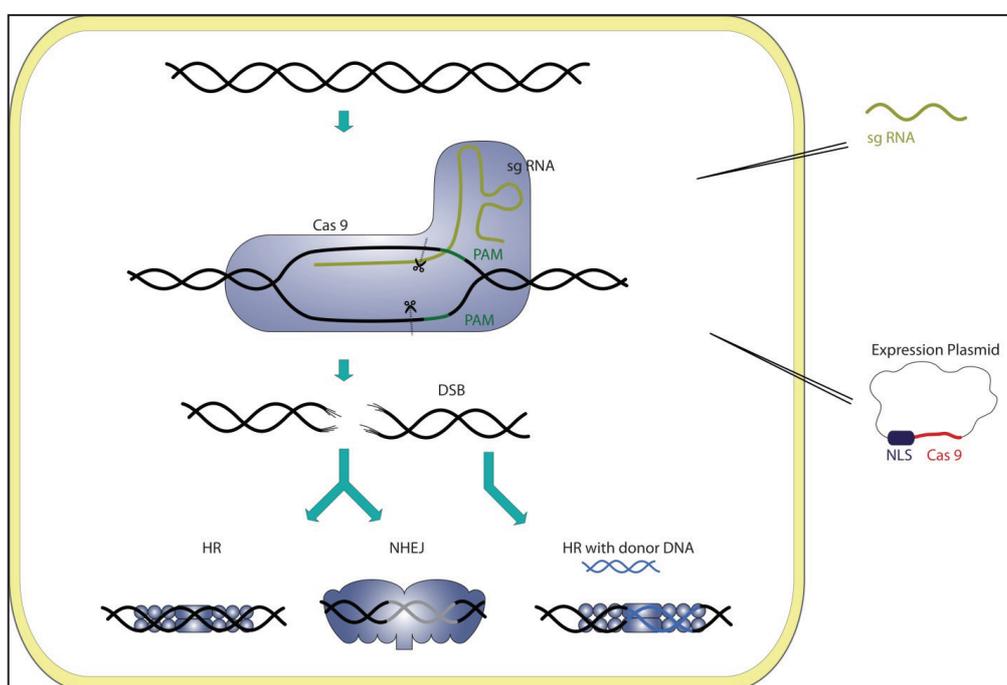
The recent analysis of Cas protein interference complexes immediately revealed their great potential for the development of genetic tools that are required to provide specific DNA or RNA targeting. One key player in this development is the large type II protein, Cas9. Qi and colleagues could show that a nuclease inactive mutant of Cas9 in combination with a sequence specific crRNA can be utilized for targeted DNA recognition to interfere with transcriptional elongation, RNA polymerase or transcription factor binding. This gene silencing activity was termed CRISPRi for CRISPR interference in reference to RNAi [57]. Subsequently, other groups reported the utilization of Cas9/crRNA complexes for genome editing in different organisms, e.g., human cell lines, zebrafish, mice, drosophila, yeast and bacteria. Cas9 interference experiments indicated that the fusion product of crRNA and tracrRNA has similar

efficiency as the RNase III-processed crRNA:tracrRNA duplex [43]. Therefore, these genome editing approaches use Cas9 together with a specific small guide RNA (sgRNA), which was designed to resemble the fused crRNA/tracrRNA sequence. The resulting ribonucleoproteins are termed RNA guided endonucleases (RGENs) and were shown to target single genes or even multiple genes, allowing efficient and site-specific editing of the target sequence [58–67]. Specificity of the targeting reaction is determined by the sgRNA sequence, which immediately displays the advantage of this method over other established genome editing methods, like zinc-finger nucleases (ZNF) or transcription activator-like effector nucleases (TALEN) [68]. A single Cas9 protein can be retargeted using different sgRNA sequences and without the time-consuming protein engineering steps required for target changes of ZNF and TALEN constructs. Moreover, the utilization of several sgRNAs in one reaction was shown to allow for multiplex editing of five genes [66].

All reported CRISPR genome editing approaches facilitate sgRNA:Cas9 systems to generate double strand breaks (DSB) of the target sequence, which can be either repaired by homologous recombination (HR) or non-homologous end joining (NHEJ) (Figure 2). In the case of HR, a full repair of the DSB in eukaryotes is facilitated, as the wild-type allele serves as a donor template. In contrast, NHEJ is an error-prone repair mechanism, which leads to the formation of insertions and deletions (indels) that can result in mutations of the particular sequence. The aim of genome editing using RGEN, ZNF or TALEN is to introduce such indels to mutate the gene of interest. However, HR is the preferred DSB repair mechanism *in vivo*, which leads to a lower efficiency of the editing process. To increase the overall yield of edited sequences, a modified donor DNA can be used, which serves as a template during HR (Figure 2).

The basic principles of established RGENs are similar for different cells. In all cases, the sequence of Cas9 was codon optimized for the particular expression system, and a eukaryotic nuclear localization signal (NLS) was fused to the protein. The plasmid encoded Cas9 production in human, mouse and yeast cells was realized by the *in vivo* production of the protein [58–66]. In zebrafish embryo cells, the mRNA of codon-optimized Cas9 fused to an NLS was microinjected before expression [60,62] (Figure 2). A variety of approaches were chosen to import the targeting sgRNA. One of the employed strategies was to transfect plasmids of a codon-optimized Cas9 together with RNase III, pre-crRNA and tracrRNA [58]. Here, it could be shown that host RNases are sufficient for sgRNA production and that no additional RNase III enzyme is needed. In other experiments, plasmids coding for a single fusion sgRNA were successfully employed [61,64,65]. Additionally, the direct microinjection of a synthesized sgRNA into zebrafish cells or the transfection of *in vitro*-produced sgRNAs in human cells resulted in similar editing efficiencies [59,62]. For the bacterial systems of *E. coli* and *Streptococcus pneumoniae*, HR was exploited with donor template DNA. Here, genomic DNA containing a truncated type II CRISPR system combined with a modified donor DNA template was transformed to yield an edited genome [63,66]. The verification of a successful editing event in mammalian cells was based on the SURVEYOR assay in which a Cel-1 nuclease specifically cleaves mismatches in hybrids of wild-type and mutant PCR products of the targeted sequence region [58,59,62,64,65,69]. Alternatively, successful editing was followed by antibiotic selection in yeast and bacteria cells [61,63].

Figure 2. Targeted genome editing using a small guide RNA (sgRNA):Cas9 complex. A plasmid encoding the codon-optimized Cas9 (red) with a nuclear localization signal (NLS) and an sgRNA (yellow), including the desired targeting sequence, are transferred into the target cell. A functional sgRNA:Cas9 interference complex is assembled in the cell. A double strand breaks (DSB) at the targeted DNA sequence upstream of a PAM (green) is introduced by the sgRNA:Cas9 complex, which can be repaired by the host DNA repair mechanisms, homologous recombination (HR) and non-homologous end joining (NHEJ). While HR restores the wild-type sequence by using the template allele, the error-prone NHEJ mechanism leads to insertions and deletions (indels) at the target site (grey). Increased editing efficiency can be achieved by co-transferring a synthetic donor DNA template for a triggered HR (blue).



4. Type I and Type III Genome Editing Tools

Currently, only the type II CRISPR/Cas systems are established for genome editing or gene silencing [58–66]. What are the barriers that prevent the development of type I or type III editing tools? One difference between the three major CRISPR/Cas types is that the interference reaction of both, type I and III systems, relies on multi-protein complexes. This complicates the transfer of these systems to other organisms, and protein-engineering of a single Cas9 protein is more straightforward than optimization of Cascade or Cmr/Csm complexes would be [11,18]. However, the fact that type III systems do not need PAM sequences for interference should be advantageous for more versatile editing events [32]. One restriction of RGEN genome editing tools is the PAM sequence, which is needed for interference. Therefore, this methodology enables the editing of sequences that occur on average every 8 bp, which leads to approximately 40% of the exons in the human genome being applicable for targeting [58,65,70]. Without the restricting PAM sequences for the interference in type III systems, genome editing could, in principle, be accomplished for any given target sequence.

However, it should be noted that less sequence restriction needs to be in balance with specificity, as targeting of additional unwanted sequences has to be avoided. Similar to established bacterial editing systems, where truncated type II CRISPR/Cas systems were transformed, an approach using minimal type I (e.g., type I-F) or type III interference cassettes in combination with an engineered CRISPR cluster is conceivable. Nevertheless, the interference of a type I and, especially, type III CRISPR/Cas system and their minimal Cas protein assemblies are still not fully understood. Future research will address their potential role as genome editing tools.

5. Genome Editing Tools: RGEN, TALEN, ZFN

The advantages of RGENs over other genome editing tools rest upon the fast, simple and economical design of the targeting crRNAs and the usage of multiple small RNAs for multiplex engineering of several genes [58,62,63,65,66,70]. The ZiFiT Targeter program was designed as a tool to easily find potential sgRNA:Cas9 targeting sites in a given genome [62]. RGEN editing works at a similar efficiency compared to ZFN or TALEN, and so far, no off-site targeting was reported [59]. A further advantage of RGENs is that the CRISPR interference results in a selection for positive clones, as wild-type sequences are constantly targeted until the mutated donor DNA is integrated by HR and inhibits the interference process [61]. Other features of the RGEN approach reveal their restrictions and disadvantages in genome editing. One problem of RGENs is their dependence on a small PAM sequence, which restricts the choice of potential target sequences. Furthermore, the length of the crRNA and its seed sequence limit the range of target sequences and, thus, an optimal design of the sgRNA is required for efficient editing [59,62,64]. ZFN and TALEN systems require protein-engineering steps, but both systems are highly tunable by exchanging the particular domains for sequence specific recognition of the target. Furthermore, libraries of different proteins already exist [71], which reduces the time and costs for the generation of new functional proteins [59,72]. The modular structure of the two systems (TALEN and ZNF) yields a highly variable tool in which the effector domain can be fused to a variety of different proteins (e.g., transposases, nucleases, transcriptional regulators) [73–75]. ZFN- and TALEN-based gene therapy approaches are tested for clinical use. Further improvements, e.g., considering recent advances in re-organizing chromosomes using RGENs, will extend the number of possible applications in gene therapy [72,76].

6. Outlook

The potential of CRISPR/Cas systems to function in genetic tools has been discussed since the identification of the system as a prokaryotic immune system. Six years after its discovery, the first steps have been made, as sgRNA:Cas9 complexes are used for efficient genome editing [58–66] and the establishment of gene silencing [57]. Development of RGEN-based genome editing systems for further model organisms (e.g., plants, insects, Archaea) could simplify their future genetic manipulation. It is required to further increase the efficiency and elucidate the occurrence of off-site targeting of RGENs to turn this system into a potential tool for disease treatment, as it has been, e.g., achieved with ZFN approaches for the treatment of HIV [72]. The establishment of RGEN genome editing tools is only one aspect, in which CRISPR/Cas immunity shows its potential for application in genetic and biotechnological systems. Other approaches include the development of phage-resistant

bacterial strains used in industrial processes or the identification of the interaction of pathogenic bacteria in a clinical setting with either phages or conjugative plasmid that can transfer antibiotic resistance. Future research on basic mechanistic details of the different CRISPR/Cas systems will reveal a more complete picture of the extensive applicability of these immune complexes.

Acknowledgments

The authors would like to thank Norman Ebel for assistance in figure preparation. This work was funded by the Max Planck Society and funds from the Deutsche Forschungsgemeinschaft (DFG FOR1680).

Conflict of Interest

The authors declare no conflict of interest.

References

1. Barrangou, R.; Fremaux, C.; Deveau, H.; Richards, M.; Boyaval, P.; Moineau, S.; Romero, D.A.; Horvath, P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **2007**, *315*, 1709–1712.
2. Grissa, I.; Vergnaud, G.; Pourcel, C. CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* **2007**, *35*, W52–W57.
3. Mojica, F.J.; Diez-Villasenor, C.; Garcia-Martinez, J.; Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **2005**, *60*, 174–182.
4. Pourcel, C.; Salvignol, G.; Vergnaud, G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **2005**, *151*, 653–663.
5. Mojica, F.J.; Diez-Villasenor, C.; Garcia-Martinez, J.; Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **2009**, *155*, 733–740.
6. Swarts, D.C.; Mosterd, C.; van Passel, M.W.; Brouns, S.J. CRISPR interference directs strand specific spacer acquisition. *PLoS One* **2012**, *7*, e35888.
7. Pul, U.; Wurm, R.; Arslan, Z.; Geissen, R.; Hofmann, N.; Wagner, R. Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol. Microbiol.* **2010**, *75*, 1495–1512.
8. Yosef, I.; Goren, M.G.; Qimron, U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* **2012**, *40*, 5569–5576.
9. Babu, M.; Beloglazova, N.; Flick, R.; Graham, C.; Skarina, T.; Nocek, B.; Gagarinova, A.; Pogoutse, O.; Brown, G.; Binkowski, A.; *et al.* A dual function of the CRISPR-Cas system in bacterial antiviral immunity and DNA repair. *Mol. Microbiol.* **2011**, *79*, 484–502.
10. Diez-Villasenor, C.; Guzman, N.M.; Almendros, C.; Garcia-Martinez, J.; Mojica, F.J. CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of *Escherichia coli*. *RNA Biol.* **2013**, *10*, doi:10.4161/rna.24023.

11. Hale, C.R.; Majumdar, S.; Elmore, J.; Pfister, N.; Compton, M.; Olson, S.; Resch, A.M.; Glover, C.V., III; Graveley, B.R.; Terns, R.M.; Terns, M.P. Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. *Mol. Cell* **2012**, *45*, 292–302.
12. Lillestol, R.K.; Shah, S.A.; Brugger, K.; Redder, P.; Phan, H.; Christiansen, J.; Garrett, R.A. CRISPR families of the crenarchaeal genus *Sulfolobus*: Bidirectional transcription and dynamic properties. *Mol. Microbiol.* **2009**, *72*, 259–272.
13. Carte, J.; Wang, R.; Li, H.; Terns, R.M.; Terns, M.P. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **2008**, *22*, 3489–3496.
14. Gesner, E.M.; Schellenberg, M.J.; Garside, E.L.; George, M.M.; Macmillan, A.M. Recognition and maturation of effector RNAs in a CRISPR interference pathway. *Nat. Struct. Mol. Biol.* **2011**, *18*, 688–692.
15. Haurwitz, R.E.; Jinek, M.; Wiedenheft, B.; Zhou, K.; Doudna, J.A. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **2010**, *329*, 1355–1358.
16. Richter, H.; Zoepfel, J.; Schermuly, J.; Maticzka, D.; Backofen, R.; Randau, L. Characterization of CRISPR RNA processing in *Clostridium thermocellum* and *Methanococcus maripaludis*. *Nucleic Acids Res.* **2012**, *40*, 9887–9896.
17. 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.
18. Jore, M.M.; Lundgren, M.; van Duijn, E.; Bultema, J.B.; Westra, E.R.; Waghmare, S.P.; Wiedenheft, B.; Pul, U.; Wurm, R.; Wagner, R.; *et al.* Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.* **2011**, *18*, 529–536.
19. Brouns, S.J.; Jore, M.M.; Lundgren, M.; Westra, E.R.; Slijkhuis, R.J.; Snijders, A.P.; Dickman, M.J.; Makarova, K.S.; Koonin, E.V.; van der Oost, J. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **2008**, *321*, 960–964.
20. Wiedenheft, B.; Lander, G.C.; Zhou, K.; Jore, M.M.; Brouns, S.J.; van der Oost, J.; Doudna, J.A.; Nogales, E. Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* **2011**, *477*, 486–489.
21. 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.; Doudna, J.A. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10092–10097.
22. Bondy-Denomy, J.; Pawluk, A.; Maxwell, K.L.; Davidson, A.R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **2013**, *493*, 429–432.
23. Haft, D.H.; Selengut, J.; Mongodin, E.F.; Nelson, K.E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **2005**, *1*, e60.
24. Makarova, K.S.; Aravind, L.; Wolf, Y.I.; Koonin, E.V. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol. Direct* **2011**, *6*, doi:10.1186/1745-6150-6-38.

25. Makarova, K.S.; Haft, D.H.; Barrangou, R.; Brouns, S.J.; Charpentier, E.; Horvath, P.; Moineau, S.; Mojica, F.J.; Wolf, Y.I.; Yakunin, A.F.; *et al.* Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* **2011**, *9*, 467–477.
26. Beloglazova, N.; Petit, P.; Flick, R.; Brown, G.; Savchenko, A.; Yakunin, A.F. Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. *EMBO J.* **2011**, *30*, 4616–4627.
27. Mulepati, S.; Bailey, S. Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). *J. Biol. Chem.* **2011**, *286*, 31896–31903.
28. Sinkunas, T.; Gasiunas, G.; Fremaux, C.; Barrangou, R.; Horvath, P.; Siksnys, V. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* **2011**, *30*, 1335–1342.
29. Westra, E.R.; Nilges, B.; van Erp, P.B.; van der Oost, J.; Dame, R.T.; Brouns, S.J. Cascade-mediated binding and bending of negatively supercoiled DNA. *RNA Biol.* **2012**, *9*, 1134–1138.
30. Sinkunas, T.; Gasiunas, G.; Waghmare, S.P.; Dickman, M.J.; Barrangou, R.; Horvath, P.; Siksnys, V. *In vitro* reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. *EMBO J.* **2013**, *32*, 385–394.
31. Semenova, E.; Jore, M.M.; Datsenko, K.A.; Semenova, A.; Westra, E.R.; Wanner, B.; van der Oost, J.; Brouns, S.J.; Severinov, K. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10098–10103.
32. Marraffini, L.A.; Sontheimer, E.J. Self *versus* non-self discrimination during CRISPR RNA-directed immunity. *Nature* **2010**, *463*, 568–571.
33. Sashital, D.G.; Wiedenheft, B.; Doudna, J.A. Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Mol. Cell* **2012**, *46*, 606–615.
34. Nam, K.H.; Huang, Q.; Ke, A. Nucleic acid binding surface and dimer interface revealed by CRISPR-associated CasB protein structures. *FEBS Lett.* **2012**, *586*, 3956–3961.
35. Mulepati, S.; Orr, A.; Bailey, S. Crystal structure of the largest subunit of a bacterial RNA-guided immune complex and its role in DNA target binding. *J. Biol. Chem.* **2012**, *287*, 22445–22449.
36. Lintner, N.G.; Kerou, M.; Brumfield, S.K.; Graham, S.; Liu, H.; Naismith, J.H.; Sdano, M.; Peng, N.; She, Q.; Copie, V.; *et al.* Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for antiviral defense (CASCADE). *J. Biol. Chem.* **2011**, *286*, 21643–21656.
37. Plagens, A.; Tjaden, B.; Hagemann, A.; Randau, L.; Hensel, R. Characterization of the CRISPR/Cas subtype I-A system of the hyperthermophilic crenarchaeon *Thermoproteus tenax*. *J. Bacteriol.* **2012**, *194*, 2491–2500.
38. Fischer, S.; Maier, L.K.; Stoll, B.; Brendel, J.; Fischer, E.; Pfeiffer, F.; Dyall-Smith, M.; Marchfelder, A. An archaeal immune system can detect multiple protospacer adjacent motifs (PAMs) to target invader DNA. *J. Biol. Chem.* **2012**, *287*, 33351–33363.

39. Peng, W.; Li, H.; Hallstrom, S.; Peng, N.; Liang, Y.X.; She, Q. Genetic determinants of PAM-dependent DNA targeting and pre-crRNA processing in *Sulfolobus islandicus*. *RNA Biol.* **2013**, *10*, doi:10.4161/rna.23798.
40. Sapranaukas, R.; Gasiunas, G.; Fremaux, C.; Barrangou, R.; Horvath, P.; Siksnys, V. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* **2011**, *39*, 9275–9582.
41. Deltcheva, E.; Chylinski, K.; Sharma, C.M.; Gonzales, K.; Chao, Y.; Pirzada, Z.A.; Eckert, M.R.; Vogel, J.; Charpentier, E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **2011**, *471*, 602–607.
42. Chylinski, K.; Le Rhun, A.; Charpentier, E. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol.* **2013**, *10*, doi:10.4161/rna.24321.
43. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012**, *337*, 816–821.
44. Gasiunas, G.; Barrangou, R.; Horvath, P.; Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E2579–E2586.
45. Garneau, J.E.; Dupuis, M.E.; Villion, M.; Romero, D.A.; Barrangou, R.; Boyaval, P.; Fremaux, C.; Horvath, P.; Magadan, A.H.; Moineau, S. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **2010**, *468*, 67–71.
46. Marraffini, L.A.; Sontheimer, E.J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **2008**, *322*, 1843–1845.
47. Marraffini, L.A.; Sontheimer, E.J. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genetics* **2010**, *11*, 181–190.
48. Zhang, J.; Rouillon, C.; Kerou, M.; Reeks, J.; Brugger, K.; Graham, S.; Reimann, J.; Cannone, G.; Liu, H.; Albers, S.V.; *et al.* Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Mol. Cell* **2012**, *45*, 303–313.
49. Shao, Y.; Coczaki, A.I.; Ramia, N.F.; Terns, R.M.; Terns, M.P.; Li, H. Structure of the Cmr2–Cmr3 subcomplex of the Cmr RNA silencing complex. *Structure* **2013**, *21*, 376–384.
50. Osawa, T.; Inanaga, H.; Numata, T. Crystal structure of the Cmr2–Cmr3 subcomplex in the CRISPR-Cas RNA silencing effector complex. *J. Mol. Biol.* **2013**, in press.
51. Hale, C.R.; Zhao, P.; Olson, S.; Duff, M.O.; Graveley, B.R.; Wells, L.; Terns, R.M.; Terns, M.P. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **2009**, *139*, 945–956.
52. Deng, L.; Garrett, R.A.; Shah, S.A.; Peng, X.; She, Q. A novel interference mechanism by a type IIIB CRISPR-Cmr module in *Sulfolobus*. *Mol. Microbiol.* **2013**, *87*, 1088–1099.
53. Vergnaud, G.; Li, Y.; Gorge, O.; Cui, Y.; Song, Y.; Zhou, D.; Grissa, I.; Dentovskaya, S.V.; Platonov, M.E.; Rakin, A.; *et al.* Analysis of the three *Yersinia pestis* CRISPR loci provides new tools for phylogenetic studies and possibly for the investigation of ancient DNA. *Adv. Exp. Med. Biol.* **2007**, *603*, 327–338.
54. Dale, J.W.; Brittain, D.; Cataldi, A.A.; Cousins, D.; Crawford, J.T.; Driscoll, J.; Heersma, H.; Lillebaek, T.; Quitugua, T.; Rastogi, N.; *et al.* Spacer oligonucleotide typing of bacteria of the *Mycobacterium tuberculosis* complex: Recommendations for standardised nomenclature. *Int. J. Tuberc. Lung Dis.* **2001**, *5*, 216–219.

55. Qi, L.; Haurwitz, R.E.; Shao, W.; Doudna, J.A.; Arkin, A.P. RNA processing enables predictable programming of gene expression. *Nat. Biotechnol.* **2012**, *30*, 1002–1006.
56. Lee, H.Y.; Haurwitz, R.E.; Apffel, A.; Zhou, K.; Smart, B.; Wenger, C.D.; Laderman, S.; Bruhn, L.; Doudna, J.A. RNA-protein analysis using a conditional CRISPR nuclease. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5416–5421.
57. Qi, L.S.; Larson, M.H.; Gilbert, L.A.; Doudna, J.A.; Weissman, J.S.; Arkin, A.P.; Lim, W.A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **2013**, *152*, 1173–1183.
58. Cong, L.; Ran, F.A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P.D.; Wu, X.; Jiang, W.; Marraffini, L.A.; Zhang, F. Multiplex genome engineering using CRISPR/Cas systems. *Science* **2013**, *339*, 819–823.
59. Cho, S.W.; Kim, S.; Kim, J.M.; Kim, J.S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **2013**, *31*, 230–232.
60. Chang, N.; Sun, C.; Gao, L.; Zhu, D.; Xu, X.; Zhu, X.; Xiong, J.W.; Xi, J.J. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res.* **2013**, *23*, 465–472.
61. Dicarlo, J.E.; Norville, J.E.; Mali, P.; Rios, X.; Aach, J.; Church, G.M. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* **2013**, *41*, 4336–4343.
62. Hwang, W.Y.; Fu, Y.; Reyon, D.; Maeder, M.L.; Tsai, S.Q.; Sander, J.D.; Peterson, R.T.; Yeh, J.R.; Joung, J.K. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **2013**, *31*, 227–229.
63. Jiang, W.; Bikard, D.; Cox, D.; Zhang, F.; Marraffini, L.A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **2013**, *31*, 233–239.
64. Jinek, M.; East, A.; Cheng, A.; Lin, S.; Ma, E.; Doudna, J. RNA-programmed genome editing in human cells. *ELife* **2013**, *2*, e00471.
65. Mali, P.; Yang, L.; Esvelt, K.M.; Aach, J.; Guell, M.; DiCarlo, J.E.; Norville, J.E.; Church, G.M. RNA-guided human genome engineering via Cas9. *Science* **2013**, *339*, 823–826.
66. Wang, H.; Yang, H.; Shivalila, C.S.; Dawlaty, M.M.; Cheng, A.W.; Zhang, F.; Jaenisch, R. One-step generation of mice carrying mutations in multiple genes by CRISPR/cas-mediated genome engineering. *Cell* **2013**, *153*, 910–918.
67. Gratz, S.J.; Cummings, A.M.; Nguyen, J.N.; Hamm, D.C.; Donohue, L.K.; Harrison, M.M.; Wildonger, J.; O'Connor-Giles, K.M. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **2013**, doi:10.1534/genetics.113.152710.
68. Segal, D.J.; Meckler, J.F. Genome engineering at the dawn of the golden age. *Ann. Rev. Genomics Hum. Genet.* **2013**, doi:10.1146/annurev-genom-091212-153435.
69. Qiu, P.; Shandilya, H.; D'Alessio, J.M.; O'Connor, K.; Durocher, J.; Gerard, G.F. Mutation detection using Surveyor nuclease. *BioTechniques* **2004**, *36*, 702–707.
70. Horvath, P.; Barrangou, R. RNA-guided genome editing a la carte. *Cell Res.* **2013**, doi:10.1038/cr.2013.39.
71. Kim, Y.; Kweon, J.; Kim, A.; Chon, J.K.; Yoo, J.Y.; Kim, H.J.; Kim, S.; Lee, C.; Jeong, E.; Chung, E.; *et al.* A library of TAL effector nucleases spanning the human genome. *Nat. Biotechnol.* **2013**, *31*, 251–258.

72. Ramalingam, S.; Annaluru, N.; Chandrasegaran, S. A CRISPR way to engineer the human genome. *Genome Biol.* **2013**, *14*, doi:10.1186/gb-2013-14-2-107.
73. Mercer, A.C.; Gaj, T.; Fuller, R.P.; Barbas, C.F., III. Chimeric TALE recombinases with programmable DNA sequence specificity. *Nucleic Acids Res.* **2012**, *40*, 11163–11172.
74. Gaj, T.; Mercer, A.C.; Gersbach, C.A.; Gordley, R.M.; Barbas, C.F., III. Structure-guided reprogramming of serine recombinase DNA sequence specificity. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 498–503.
75. Yant, S.R.; Huang, Y.; Akache, B.; Kay, M.A. Site-directed transposon integration in human cells. *Nucleic Acids Res.* **2007**, *35*, e50.
76. Xiao, A.; Wang, Z.; Hu, Y.; Wu, Y.; Luo, Z.; Yang, Z.; Zu, Y.; Li, W.; Huang, P.; Tong, X.; *et al.* Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.* **2013**, in press.

Reprinted from *IJMS*. Cite as: Pooggin, M.M. How Can Plant DNA Viruses Evade siRNA-Directed DNA Methylation and Silencing?. *Int. J. Mol. Sci.* **2013**, *14*, 17347-17377.

Review

How Can Plant DNA Viruses Evade siRNA-Directed DNA Methylation and Silencing?

Mikhail M. Pooggin

University of Basel, Department of Environmental Sciences, Botany, Schönbeinstrasse 6, Basel 4056, Switzerland; E-Mail: mikhail.pooggin@unibas.ch; Tel.: +41-61-267-2977; Fax: +41-61-267-3504

Received: 13 June 2013; in revised form: 1 July 2013 / Accepted: 1 July 2013 /

Published: 24 July 2013

Abstract: Plants infected with DNA viruses produce massive quantities of virus-derived, 24-nucleotide short interfering RNAs (siRNAs), which can potentially direct viral DNA methylation and transcriptional silencing. However, growing evidence indicates that the circular double-stranded DNA accumulating in the nucleus for Pol II-mediated transcription of viral genes is not methylated. Hence, DNA viruses most likely evade or suppress RNA-directed DNA methylation. This review describes the specialized mechanisms of replication and silencing evasion evolved by geminiviruses and pararetroviruses, which rescue viral DNA from repressive methylation and interfere with transcriptional and post-transcriptional silencing of viral genes.

Keywords: plant virus; DNA virus; geminivirus; pararetrovirus; silencing; siRNA; RNA-directed DNA methylation; cytosine methylation; silencing evasion; suppressor protein

1. Introduction

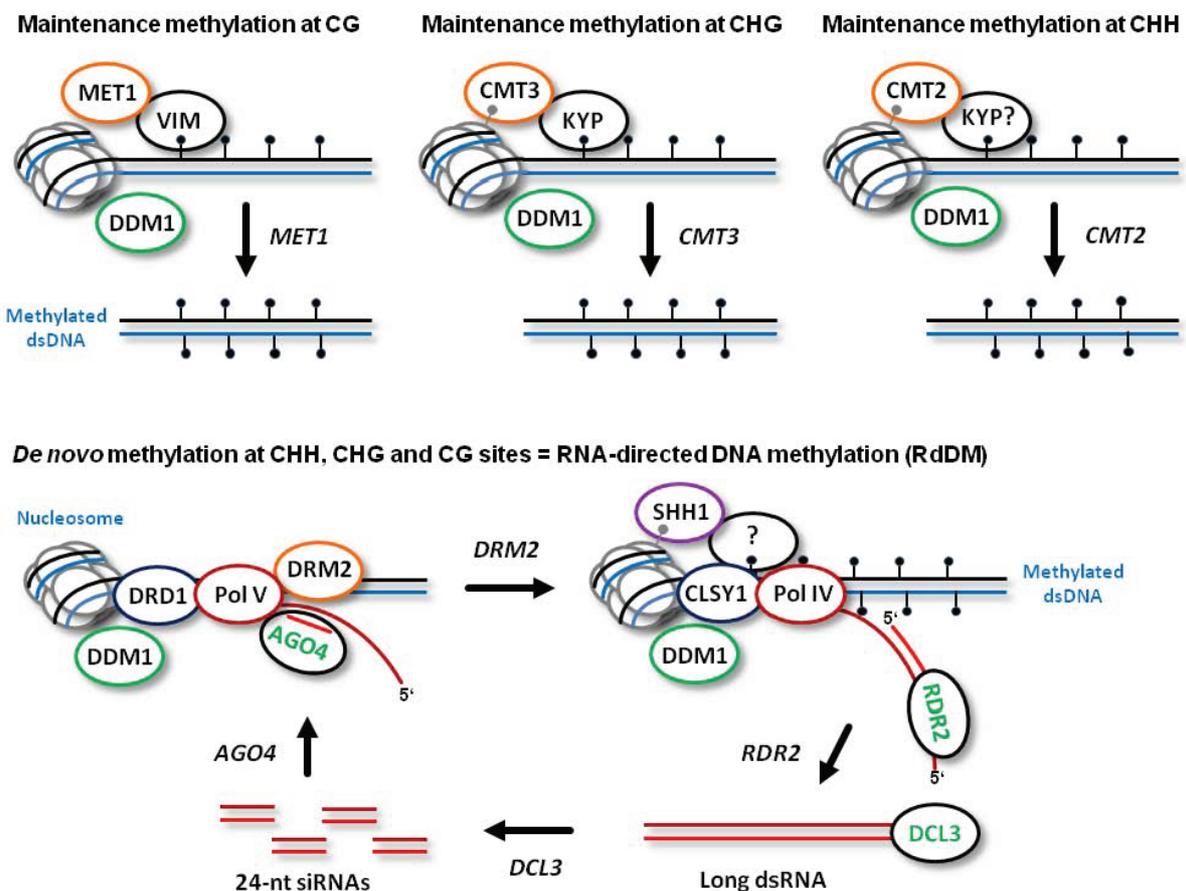
DNA viruses accumulate in the nuclei of infected plant cells as multiple circular minichromosomes, which resemble the host plant chromosomes in that the viral DNA is packaged into nucleosomes forming chromatin. Furthermore, viral minichromosomes are transcribed by the host Polymerase II (Pol II), which generates capped and polyadenylated viral RNAs, similar to mRNAs generated by Pol II from most plant protein-coding genes. Thus, viral minichromosomes must encounter the nuclear pathways that regulate host gene expression and chromatin states. However, DNA viruses have evolved specialized mechanisms of replication that differ from those replicating the plant chromosomes. These replication mechanisms can potentially rescue viral minichromosomes from repressive chromatin marks that silence certain plant genes and repetitive DNA elements in

transcriptionally-inactive heterochromatic regions. Some of the repressive chromatin marks are established by the RNA-directed DNA methylation (RdDM) pathway. RdDM is a nuclear branch of the plant RNA silencing machinery that regulates gene expression and defends against invasive nucleic acids such as transposons, transgenes and viruses. The plant RNA silencing machinery generates 21, 22 and 24 nt small RNAs which are broadly classified into miRNAs and short interfering RNAs (siRNAs). These small RNAs serve as guide molecules for the silencing complexes that repress genes post-transcriptionally and/or transcriptionally in a sequence-specific manner. The transcriptional silencing through *de novo* DNA methylation is directed by 24-nt siRNAs, the most diverse and abundant class of plant small RNAs. Likewise, plant DNA viruses spawn massive quantities of viral 24-nt siRNAs which can potentially silence viral DNA. In this review, I will focus mainly on the nuclear events in life cycles of plant DNA viruses and describe the strategies of silencing evasion evolved by *Geminiviridae* (geminiviruses) and *Caulimoviridae* (pararetroviruses), the two major families of plant DNA viruses. The third DNA virus family, *Nanoviridae*, is discussed, because little is known about interactions of nanoviruses with the plant silencing system. Since they resemble geminiviruses in DNA replication mechanisms [1], the findings for geminiviruses could be extrapolated to nanoviruses. The post-transcriptional RNA silencing mechanisms which contribute to plant defenses against both RNA and DNA viruses, and the biogenesis and function of the three major classes viral siRNAs including 21-nt and 22-nt classes have been reviewed comprehensively [2–5]. Various silencing suppressor proteins encoded by plant viruses have also been reviewed [6,7], and I will focus only on those encoded by DNA viruses and describe emerging evidence that viral suppressor proteins may have effector functions in suppressing plant innate immunity [8].

2. Plant DNA Methylation

DNA methylation at cytosine nucleotides (5meC) is a reversible epigenetic mark that plays a key role in regulation of gene expression and chromatin states in most eukaryotes. Plants and mammals require cytosine methylation for proper development and genome defense against transposons [9,10]. In mammals, methylation occurs predominantly at symmetric CG sites and, following DNA replication, can be maintained by DNA METHYLTRANSFERASE 1 (DNMT1). DNMT1 recognizes hemimethylated double-stranded DNA (dsDNA) with the help of methyl binding domain proteins and catalyzes methylation of symmetric cytosines on the newly-synthesized strand. Establishment of cytosine methylation on unmethylated dsDNA is catalyzed by *de novo* methyltransferases DNMT3a and DNMT3b. Furthermore, methylated dsDNA can be actively demethylated, which ensures dynamic regulation of chromatin states during development and in response to environmental cues. Generally, methylated DNA is repressed transcriptionally, because it is packed into heterochromatin inaccessible to RNA polymerases, whereas unmethylated DNA is present in open actively-transcribed euchromatin.

Figure 1. Models for maintenance methylation and RdDM at the plant genome loci. (Based mostly on the findings using the model plant *Arabidopsis*). The plant dsDNA associated with nucleosomes is depicted as solid lines and the methylated cytosines at one or both strands indicated with black lollypops. Following DNA replication, cytosine methylation at CG, CHG and CHH sites of the newly-synthesized strand (blue) is catalyzed by the maintenance methyltransferases MET1, CMT3, and CMT2, respectively, with the help of co-factors VIM and KYP that recognize hemimethylated dsDNA; CMT3 and CMT2 also bind the repressive histone methylation mark H3K9me2 indicated as grey lollypops. The RNA-directed DNA methylation (RdDM) pathway establishing methylation of dsDNA *de novo* is catalyzed by DRM2 that interacts with the DRD1-Pol V complex generating a scaffold transcript. The nascent scaffold transcript is targeted by the 24-nt siRNA-AGO4 complex. Following DRM2-catalyzed *de novo* methylation of both DNA strands, Pol IV with the help of SHH1 (binding H3K9me2) and CLSY1 initiates siRNA biogenesis. The Pol IV transcript is converted by RDR2 to dsRNA. The resulting dsRNA is processed by DCL3 into 24-nt siRNA duplexes. The duplexes are handed over to AGO4 to form the silencing complexes with a single-stranded siRNA guide. This completes an siRNA amplification loop that reinforces RdDM-mediated transcriptional silencing. The chromatin remodeler DDM1 facilitates the access of all the methyltransferases to dsDNA.



In flowering plants, cytosines in all possible sequence contexts can be methylated, including symmetric (CG and CHG, where H is A, C, or T) and asymmetric (CHH). *De novo* establishment of methylation at CG, CHG and CHH sites is catalyzed by DOMAINS REARRANGED

METHYLTRANSFERASE 2 (DRM2), the plant homolog of mammalian DNMT3a and DNMT3b, which requires 24-nt siRNA guide molecules and other components of the RdDM pathway (Figure 1; see below for more details). Following DNA replication, symmetric CG methylation is maintained by DNA METHYLTRANSFERASE 1 (MET1), the plant homolog of mammalian DNMT1, which recognizes hemimethylated dsDNA with the help of CG-specific methyl binding proteins VARIANT IN METHYLATION 1 (VIM1), VIM2 and VIM3 [11] (Figure 1). Symmetric CHG methylation is maintained by CHROMOMETHYLASE 3 (CMT3), a plant-specific methyltransferase that recognizes dimethylated histone 3 tails at lysine 9 (H3K9m2) on the nucleosomes (Figure 1). In this process, CHG methylation at the template strand is recognized by the H3K9m2 methyltransferase KRYPTONITE (KYP), which can bind methylated cytosines in both CHG and CHH context [12]. Thus, CHG methylation is maintained through a reinforcing loop of DNA and histone (H3K9) methylation. Recently, a homolog of CMT3, CMT2, has been implicated in maintenance methylation at CHH sites [13]. Like CMT3, CMT2 is recruited through direct recognition of the methylated histone H3K9me2 and does not require siRNA guides or other components of RdDM (previously thought to be the only pathway maintaining CHH methylation). Furthermore, indirect recognition of the hemimethylated DNA by CMT2 may also require KYP that binds methylated CHH sites (Figure 1).

Both maintenance methylation and RdDM are facilitated by a chromatin remodeler DEFICIENT IN DNA METHYLATION 1 (DDM1). Indeed, 70% of CG, CHG and CHH methylation is lost in *ddm1* mutant plants. It is believed that maintenance methylation does not take place on naked dsDNA immediately following passage of the DNA replication fork, and that cytosine methylation occurs in a nucleosomal context involving both core and linker histones [14]. DDM1 remodels heterochromatin by removing the repressive linker histone H1 [13]. Obviously, all the DNA methyltransferases need the access to DNA, which can be facilitated by DDM1 (Figure 1). Together, DDM1 and RdDM synergize to maintain all the cytosine methylation in the plant genome [13].

Other factors required for normal DNA methylation include those that have direct or indirect impact on the levels of *S*-adenosyl-L-methionine (SAM), the donor of methyl groups.

3. Mechanism of RNA-Directed DNA Methylation (RdDM)

RdDM is mediated by two plant-specific DNA-dependent RNA polymerases, Pol IV and Pol V: Pol IV functions to initiate siRNA biogenesis, while Pol V generates scaffold transcripts that recruit downstream RdDM factors [15]. Both Pol IV and Pol V are plant-specific enzymes that have evolved from Pol II and share several core Pol II subunits. However, little is known about promoters and other regulatory elements driving transcription at the RdDM loci; the transcripts generated by Pol V and Pol IV were not precisely mapped.

The model depicted in Figure 1 (based mostly on the findings using the model plant *Arabidopsis*) states that Pol V scaffold transcripts are produced at DNA loci to be methylated *de novo*. The nascent scaffold transcript is targeted by an ARGONAUTE 4 (AGO4) protein complex containing a 24-nt siRNA guide molecule via complementary interaction of the siRNA and the scaffold RNA. AGO4 belongs to a family comprising ten members, most of which possess catalytic activity required for sequence-specific cleavage of their target RNAs and subsequent gene silencing at both transcriptional and post-transcriptional levels [16,17]. Besides catalyzing cleavage of the nascent Pol V transcript,

AGO4 interacts with Pol V itself. Together, these interactions are required for recruitment of the methyltransferase DRM2 (or its homolog DRM1) and for subsequent *de novo* methylation of both DNA strands (Figure 1). Other factors that facilitate Pol V transcription and DRM2 recruitment include DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1), DEFECTIVE IN MERISTEM SILENCING 3 (DMS3) and RNA-DIRECTED DNA METHYLATION 1 (RDM1). These proteins form a complex proposed to unwind dsDNA in front of Pol V (via a putative DNA translocase/ATPase activity of DRD1) and to mediate recruitment of DRM2 to the AGO4-bound scaffold transcript. Following AGO4-catalyzed cleavage of the scaffold transcript, the released siRNA-AGO4 complex may bind the complementary DNA and thereby define the region to be methylated by DRM2 [15]. Other members of the nuclear AGO clade, AGO6 and AGO9, which display tissue specific expression, might also function in RdDM together with, or in place of AGO4 [18].

The biogenesis of 24-nt siRNAs at the RdDM loci is initiated by Pol IV transcription. Pol IV transcripts are then converted to double-stranded RNA (dsRNA) by RNA-DEPENDENT RNA-POLYMERASE 2 (RDR2) (Figure 1). RDR2 belongs to a family with at least three functional enzymes involved in the biogenesis of distinct classes of endogenous plant siRNAs and viral secondary siRNAs. Thus, RDR6 generates dsRNA precursors of plant trans-acting siRNAs (tasiRNAs), which have been precisely mapped [19,20], while RDR1 and RDR6 together are involved in the biogenesis of secondary siRNAs derived from RNA viruses [21–23]. RDR2-dependent dsRNA precursors of 24-nt siRNAs have not been mapped, and it is presumed that RDR2 converts to dsRNA a complete Pol IV transcript, or generates Okazaki-like fragments on the nascent Pol IV transcript [15]. Notably, RDR2 and Pol IV form a complex, and RDR2 has no activity in the absence of Pol IV [24]. Together, Pol IV and RDR2 are required for the biogenesis of virtually all endogenous plant 24-nt siRNAs.

Pol IV is localized at the target loci through interaction with SAWADEE HOMEODOMAIN HOMOLOG 1 (SHH1) that recognizes H3K9me2 [25]. Furthermore, Pol IV occupancy at actively-transcribed, siRNA-generating loci may also require methyl binding protein activity, because Pol IV is believed to transcribe methylated DNA following *de novo* methylation (Figure 1). Pol IV transcription of methylated DNA at the RdDM loci would amplify 24-nt siRNAs to reinforce silencing *in cis*, maintain methylation following replication, and enable *de novo* methylation of homologous DNA loci *in trans*.

De novo methylation might also occur at some Pol II loci via targeting of nascent Pol II transcripts by 24-nt siRNAs [26–28]. In fact, such events might trigger *de novo* methylation and transcriptional silencing of active long-terminal repeat (LTR) retrotransposons (whose genomic RNA is generated by Pol II) following their transposition at new loci.

RdDM and Pol IV activity require CLASSY 1 (CLSY1), a putative ATP-dependent nucleic acid translocase predicted to evict nucleosomes and unwind dsDNA (Figure 1). As discussed above, the chromatin remodeler DDM1 might also facilitate RdDM by removing the repressive histone H1. Establishment of other repressive histone modifications at RdDM loci is catalyzed by a Jumonji domain protein JMJ14 and HISTONE DEACETYLASE 6 (HDA6). JMJ14 demethylates histone H3 lysine 4, thus removing the mark associated with active chromatin. Likewise, HDA6 removes acetyl groups from histone lysines (*i.e.*, active chromatin marks), which is a prerequisite for their subsequent methylation creating the repressive marks such as H3K9me2 [15].

The final step in the biogenesis of endogenous 24-nt siRNAs is accomplished by DICER-LIKE 3 (DCL3), an RNase III-like enzyme that belongs to a family of four prototype members [29]. DCL3 catalyzes processing of RDR2-dependent dsRNA into 24-nt siRNA duplexes (Figure 1). These duplexes are then methylated at the 3'-terminal nucleotides' hydroxyls by HUA ENHANCER 1 (HEN1) and sorted by AGO4, AGO6, or AGO9 to form the silencing complexes containing a single-stranded 24-nt siRNA guide molecule [18]. Either strand of the siRNA duplex can get incorporated into the AGO complex, which enables targeting of both sense and antisense transcripts, potentially generated at the RdDM loci.

4. DNA Demethylation

DNA demethylation can occur passively through several rounds of DNA replication in the absence of efficient maintenance methylation, or actively through enzymatic activities. In plants, DNA glycosylases have been implicated in active removal of 5meC from DNA [9,30]. These include DEMETER (DME) which controls imprinting in reproductive tissues, REPRESSOR OF SILENCING 1 (ROS1) initially identified as suppressor of transcriptional silencing of a plant promoter-driven transgene, and two DEMETER-LIKE enzymes (DML2 and DML3) which, together with ROS1, counteract excessive methylation at several hundred loci across the genome [31–33]. The DNA glycosylases can remove repressive cytosine methylation marks in all the sequence contexts without the need for DNA replication and thereby release transcriptional silencing. However, it is not clear what provides sequence specificity for these enzymes. Animals apparently lack 5meC DNA glycosylases and demethylation involves excision of de-aminated and/or oxidized derivatives of 5meC [34].

A crosstalk between demethylation and *de novo* methylation pathways has been recently illustrated by the finding that ROS1 expression is controlled by the RdDM pathway and mutations in Pol IV and Pol V cause transcriptional silencing at the ROS1 target loci [35].

5. Replication Modes of Geminiviruses

The family *Geminiviridae* comprises circular single-stranded DNA (ssDNA) viruses with 2.5–3.2 kb genomes [36]. The *Begomovirus* genus contains monopartite or bipartite geminiviruses with an additional circular ssDNA component of similar size (DNA-B). Viral ssDNA is encapsidated by viral coat protein in twinned (gemininate) virions. The life cycle of geminiviruses, their replication and gene expression strategies have been comprehensively reviewed [37,38]. According to the current model (Figure 2), following insect injection into a plant cell, the viral particle is targeted via a coat protein-based nuclear localization signal to the nucleus, where viral ssDNA is released into nucleoplasm. The circular ssDNA is then converted to circular dsDNA by the host DNA polymerase and other components of the DNA repair machinery. In genus *Begomovirus*, the complementary strand synthesis is primed by an RNA primer [39]. By contrast, in genus *Mastrevirus*, a nested set of complementary strand DNA primers with major species ranging from 78 to 88 nts were found to be associated with virion-derived ssDNA [40].

Figure 2. Models for RCR and RDR modes of geminivirus DNA replication. **(a)** RCR. The viral circular ssDNA is released from the virion (yellow) into the nucleus. The host DNA polymerase synthesizes the complementary strand, yielding circular covalently-closed dsDNA. This dsDNA serves as a template for bidirectional transcription of the early leftward (Rep) and the late rightward (coat protein) genes. Viral mRNAs are transported to the cytoplasm. Following translation, Rep moves to the nucleus to initiate replication of the viral dsDNA by a rolling circle replication (RCR) mechanism. Rep (in yellow) nicks the virion strand in the origin of replication and recruits the host DNA polymerase to extend 3'-end of the cleaved virion strand on the complementary strand template. As the extension progresses, the polymerase complex, associated with Rep covalently linked to the 5'-end of the virion strand, displaces the virion strand. After one or more rounds of replication on the circular complementary strand template, Rep nicks and religates the displaced virion strand extended by one or more copies of the newly-synthesized virion strand and thereby releases one or more copies of circular ssDNA. The resulting circles can re-enter the replication cycle or get packaged into virions; **(b)** The circular covalently-closed dsDNA is invaded by a short viral DNA primer. The primer is extended by the host DNA polymerase on the circular viral template strand. After (or during) one or more rounds of replication, the newly-synthesized linear ssDNA gets fully or partially converted to linear dsDNA by the same (or another) DNA polymerase complex. Thus, RDR generates a heterogeneous population of linear dsDNAs. The long linear dsDNAs that harbor two or more origins of replication are transcribed by Pol II in both orientations to generate viral mRNAs. Following translation, Rep initiates replication of the long linear dsDNA with two or more origins of replication. The replicational release of ssDNA from the multimeric linear dsDNA generates circular ssDNA that can re-enter the replication cycle or get packaged.

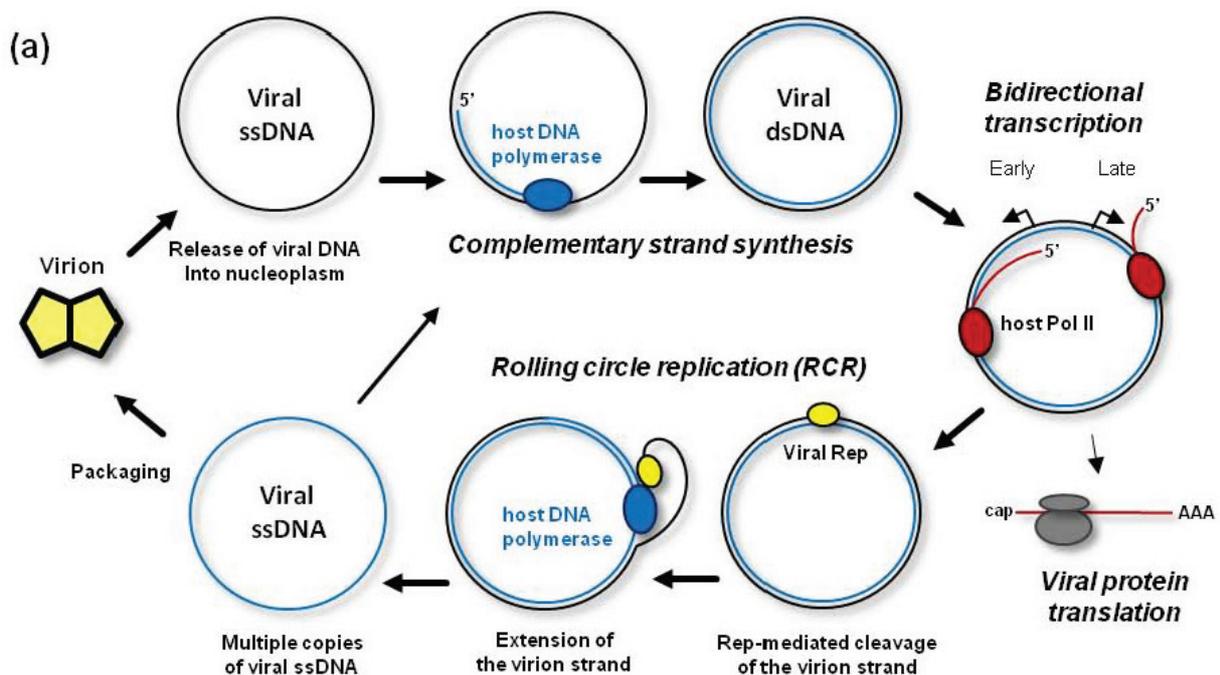
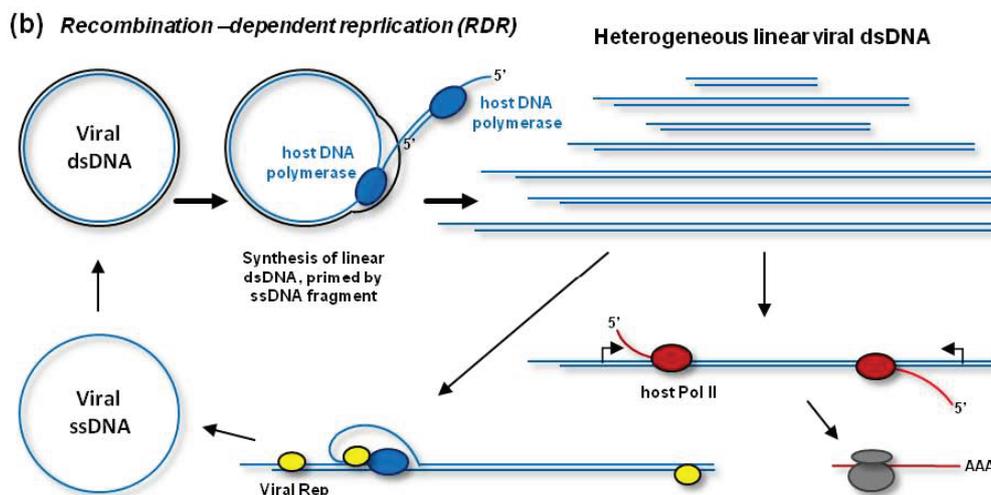


Figure 2. *Cont.*

Following complementary strand synthesis, the resulting covalently-closed circular dsDNA gets associated with nucleosomes [41,42] and transcribed by the host Pol II (Figure 2A). Pol II transcribes the viral minichromosome in the leftward orientation to generate mRNAs for viral replication-initiator protein (Rep) and other proteins assisting replication and transcription. At a later stage, the minichromosome is transcribed by Pol II in the rightward orientation to generate mRNA for coat protein [37,43]. In begomovirus-infected plants, the number of nucleosomes per viral minichromosome is varying between 11 and 12, presumably representing transcriptionally active states, and 13, representing inactive state [42].

After production by the cytoplasmic ribosomes, the viral Rep protein moves to the nucleus to initiate rolling circle replication (RCR) of the viral dsDNA that had given rise to the Rep mRNA. Rep is the only viral protein essential for the RCR mechanism generating multiple copies of circular ssDNA. Rep initiates RCR by nicking the virion strand of dsDNA in a conserved nonanucleotide sequence of the replication origin and by recruiting the host DNA polymerase complex. The polymerase uses the circular complementary strand as a template to extend 3'-end of the cleaved virion strand. During this process, the virion strand with Rep covalently linked to its 5'-end is displaced from the template strand (Figure 2A). Rep helicase activity has also been implicated in a post-initiation phase of RCR [44]. After one or more rounds of RCR, Rep (being associated with the polymerase complex) nicks and ligates the displaced virion strand extended with one or more copies of the newly-synthesized virion strand, and thereby releases circular ssDNA from the complex. Thus, multiple circles of viral ssDNA are synthesized on one complementary ssDNA circle (Figure 2A). These circles can re-enter the replication cycle, or get packaged into virions at later stages of infection, when viral coat protein is accumulated. As a result of RCR, multiple copies of viral minichromosomes accumulate in the initially-infected nucleus and eventually in the nuclei of other cells that are infected by cell-to-cell and long-distance movement of viral particles.

In addition to RCR, geminiviruses can replicate their dsDNA by a recombination-dependent replication (RDR) mechanism [38,45–47]. According to a model shown in Figure 2B, RDR is initiated by a viral ssDNA fragment that invades a homologous region of the circular dsDNA with the help of the host recombination enzymes. Then the host DNA polymerase extends the invaded ssDNA on a

template strand. During (or after) one or more rounds of the extension on the circular template, the resulting linear ssDNA is converted to dsDNA by the DNA polymerase complex primed by a short complementary fragment of viral DNA (or RNA). Thus, RDR generates a heterogeneous population of linear dsDNAs, which accumulate at high levels during viral infection and become targeted for cytosine methylation [48] (see below). RDR priming does not require Rep activity [38]. However, Rep may release circular ssDNA from the heterogeneous linear dsDNA, which contains two or more origins of replication [49] (Figure 2B). In fact, such mechanism is responsible for the release of circular ssDNA from partial dimer clones of geminiviruses widely used for experimental inoculations.

The efficient mechanism of RDR evolved by geminiviruses explains why recombination is a major driving force for their evolution and a frequent cause of epidemics. Indeed, if two geminiviruses enter the same nucleus, RDR will very likely produce a wide variety of chimeric genomes. It should be stressed, however, that Rep-mediated RCR is essential for systemic infection of the plant and for formation of the virions with circular ssDNA, which are transmitted by insects from plant to plant. Thus, both modes of replication are required for robust infection and spread of geminiviruses. The following section describes how RCR and RDR help geminiviruses evade repressive cytosine methylation and transcriptional silencing.

6. Evasion of Maintenance Methylation and RdDM by Geminiviruses

It has been proposed that cytosine methylation is one of the major host defense mechanisms against geminiviruses and therefore these viruses have evolved different suppressor proteins to interfere with repressive methylation and transcriptional silencing of viral DNA [50]. Here I argue that geminiviruses can evade repressive methylation simply via efficient Rep-dependent replication as has been suggested earlier [48,51].

Experimental evidence based on bisulfite treatment of total DNA from geminivirus-infected plants, followed by PCR amplification and sequencing of the virion strand, shows that 50% to 99% (depending on the virus or the host used) of viral DNA is not methylated [52,53]. Note that technical biases of the bisulfite sequencing method may have prevented correct evaluation of the percentage of 5meC in viral DNA, as discussed by Paprotka *et al.* [48]. Interestingly, methylated cytosines were not randomly distributed between the viral molecules, but concentrated in a small fraction of densely methylated molecules [53]. Hence, a large fraction of viral DNA is not methylated at all. Since the most abundant form of viral DNA is circular ssDNA that gets encapsidated into virions, the above findings imply that this form is not methylated and therefore maintenance methylation does not occur during Rep-mediated RCR. As discussed above, maintenance methylation likely occurs in a nucleosomal context (Figure 1). During the first round of RCR the nucleosomes are removed from the replicating viral DNA and their formation is prevented by continuous rounds of replication displacing newly-synthesized ssDNA (Figure 2A). Moreover, the latter ssDNA is only transiently associated with the template strand, thus preventing an access of the hemimethylated dsDNA-binding proteins required for recruitment of methyltransferases (Figure 1). Likewise, the RDR mechanism generating heterogeneous linear dsDNA on a circular dsDNA template (Figure 2B) is not compatible with maintenance methylation. The variable levels of cytosine methylation detected by bisulfite sequencing in all the sequence contexts [52,53] likely reflect the amounts of *de novo* methylated viral dsDNA in

circular or linear forms, both containing the virion strand. Using the bisulfite sequencing approach to evaluate a methylation status of the complementary strand (*i.e.*, RCR template) revealed 36% to 45% of methylation in all the sequence contexts [51]. Hence, a large fraction of viral dsDNA is also not methylated. Taking the above findings and considerations together, detectable methylation of geminiviral DNA is established *de novo*, possibly through RdDM using viral 24-nt siRNA guides. Potential targets of RdDM could be both circular dsDNA and heterogeneous linear dsDNA, which undergo transcription (Figure 2), because RdDM requires on-going transcription at the endogenous target loci (Figure 1).

The lack of maintenance methylation during geminivirus replication is further supported by the findings that geminiviral clones methylated *in vitro* gave rise to unmethylated dsDNA progeny in plant protoplasts, although viral DNA replication was inhibited compared to unmethylated controls [54,55]. These findings illustrate the repressive nature of cytosine methylation, likely inhibiting initial transcription of viral genes. However, more importantly, they demonstrate the ability of geminiviruses to resurrect viral DNA from repressive methylation by evading maintenance methylation during replication.

The methylation status of viral dsDNA in the above-described protoplast studies was evaluated by treatment of total DNA with methylation sensitive enzymes, followed by Southern blot hybridization with virus-specific probes. This approach did not reveal any substantial methylation of viral circular dsDNA in plants infected with different geminiviruses [48,54,56]. The conflicting results obtained with two different methods can be explained by the inability of PCR-based bisulfite sequencing to discriminate between different forms of viral DNA. To resolve this problem, more advanced methods have been applied, using treatment of total DNA with methylation-dependent enzyme McrBC, followed by 1-D or 2-D gel separation and Southern blot analysis or detection with 5mC-specific antibodies [48]. This study has confirmed that circular dsDNA, the template for both replication and transcription, is not methylated. The only viral DNA form that possessed detectable cytosine methylation is heterogeneous linear dsDNA, the product of RDR. Therefore, the extremely variable levels of DNA methylation detected by bisulfite sequencing, ranging for wild-type geminiviruses from 1.25% to 50%–60% [52,53], may reflect the amounts of heterogeneous linear dsDNA accumulated in the respective virus-host systems. The highest methylation level (88%) was reported for an intergenic region of the curtovirus *Beet curly top virus* (BCTV) mutant lacking an L2 gene [52]. *Arabidopsis* plants recover from this mutant virus infection and accumulate very low levels of highly methylated viral DNA. The residual replication of this defective virus in recovered tissues may proceed mainly by RDR that generates linear dsDNA, the target for methylation. Another explanation is that the BCTV L2 protein acts an active suppressor of cytosine methylation [57] (discussed below).

It has been reported that plants deficient in cytosine methylation exhibit increased sensitivity to geminivirus infection [52]. Thus, enhanced disease symptoms were observed for the begomovirus *Cabbage leaf curl virus* (CaLCuV) and the curtovirus BCTV in *Arabidopsis* mutants lacking core components of maintenance methylation or RdDM, which included DRM1/2, Pol IV/V, DDM1, MET1, CMT3, KYP, DCL3, or AGO4. However, mutant plants lacking RDR2, which is also required for RdDM (Figure 1B), did not display enhanced symptoms. Moreover, the mutant plants displaying enhanced symptoms accumulated the wild-type levels of viral DNA [52]. Hence, viral DNA replication is not “de-repressed” in the absence of core components of RdDM or maintenance methylation pathways and, in wild type plants, repressive cytosine methylation can be effectively evaded, likely by

Rep-mediated replication of viral DNA. Consistent with this notion, another study did not reveal increased titres or enhanced symptoms of CaLCuV in *Arabidopsis* mutants lacking Pol IV, RDR2, DCL3, or AGO4 [58]. Notably, viral 24-nt siRNAs were normally produced in all these mutants, except *dcl3*, indicating that the biogenesis of viral 24-nt siRNAs does not require the RdDM components essential for production of dsRNA precursors of endogenous 24-nt siRNAs (see below).

7. Suppression of Cytosine Methylation and Transcriptional Silencing by Geminiviral Proteins

Geminiviral proteins implicated in suppression of cytosine methylation and transcriptional silencing include AC2/AL2/C2/L2 homologs encoded by *Begomovirus* and *Curtovirus* genera and betaC1 encoded by betasatellites associated with some begomoviruses. Since these proteins are not conserved in all genera of *Geminiviridae*, other viral protein(s) may suppress transcriptional silencing (see below) or, as argued here, all geminiviruses should be able to evade cytosine methylation through Rep-dependent replication.

The begomovirus AC2/AL2/C2 gene encodes a transcriptional activator (TrAP) required for activation of late viral genes in the nucleus [37,43]. This protein was also shown to suppress post-transcriptional silencing and its nuclear localization was required for this activity [59,60]. Notably, the suppressor activity of AC2 from two Old World begomoviruses correlated with upregulation of a common subset of host genes including WERNER-LIKE EXONUCLEASE 1 (WEL1), which may act as negative regulators of RNA silencing [60]. Thus, AC2 appears to suppress post-transcriptional silencing indirectly via transcriptional activation of the host genes. Interestingly, the *WEL1* gene that codes for a putative silencing suppressor [60] seats in a transcriptionally-silent locus containing seven *WEL1* paralogs in head-to-tail orientation [60]. Transcriptional silencing of this repetitive DNA locus, likely associated with repressive chromatin marks, might be reversed by the viral TrAP activity. It should be mentioned that in addition to its antisilencing function, *WEL1* may also function in viral DNA replication, because it encodes a putative 3'–5' exonuclease [60] which resembles the WERNER exonuclease involved in DNA replication, recombination and repair.

It has been demonstrated that the begomovirus CaLCuV AL2 and the curtovirus BCTV L2 can reverse transcriptional silencing at transgenic and some endogenous loci repressed by cytosine methylation [57]. The reversal of silencing correlated with partial reduction of non-CG methylation at the respective loci as well as with genome-wide reduction in CHG methylation. In this process, CaLCuV AL2 did not require the C-terminal transcriptional activation domain [57]. This is in contrast to a homologous AC2 protein from the Old World begomovirus *Mungbean yellow mosaic virus*, which needs this domain to activate the host genes and suppress post-transcriptional silencing [60]. Surprisingly, both AL2 from the New World begomoviruses (CaLCuV and TGMV) and L2 from the curtovirus BCTV reverse transcriptional silencing and cytosine methylation by a mechanism that does not require their nuclear localization. These proteins interact with and inactivate ADENOSINE KINASE (ADK), a cytoplasmic enzyme involved in the methyl cycle producing SAM, the donor of methyl groups [50]. Curiously, AL2/L2-mediated inactivation of ADK was required for suppression of both transcriptional [57] and post-transcriptional [61] silencing.

A different mechanism of silencing suppression was reported for C2 of the curtovirus *Beet severe curly top virus*. This protein interacts with SAM DECARBOXYLASE 1 and thereby interferes with

the methyl cycle and DNA methylation [62]. Given that different mechanisms were reported for the closely homologous proteins such as C2 and L2 from curtoviruses as well as AC2 and AL2 from begomoviruses, further research should clarify their activities in the nucleus and the cytoplasm.

Some monopartite begomoviruses are associated with betasatellites that enhance disease symptoms [63]. Betasatellites code for a single protein, betaC1, reported to act as a suppressor of post-transcriptional silencing [64,65]. Like begomoviral AC2/C2, betaC1 from a betasatellite of *Tomato yellow leaf curl China virus* (TYLCCNV) is a nuclear protein and its nuclear localization is required for silencing suppression [64]. The same betaC1 protein was also reported to reverse transcriptional silencing and CHG methylation through inactivation of S-ADENOSYL HOMOCYSTEINE HYDROLASE (SAHH), a methyl cycle enzyme required for synthesis of the methyl donor SAM [53]. Curiously, betaC1 requires an intact nuclear localization signal for the cytoplasmic interaction with SAHH. In the presence of betasatellite, cytosine methylation of the TYLCCNV virion strand was reduced from 5.4% to 1.25%. Since TYLCCNV infection or expression of the TYLCCNV C2 protein failed to reverse transcriptional silencing or cytosine methylation at endogenous loci, betaC1 was proposed to functionally substitute for a loss-of-function mutation in the TYLCCNV C2 gene [53]. However, very low methylation of TYLCCNV DNA in the absence of betasatellite (5.4%) would argue against the absolute necessity for a geminivirus to possess a suppressor of cytosine methylation.

Enhanced symptoms of geminiviral disease in the presence of betasatellite as well as in the methylation-deficient mutant plants described above could be explained by possible involvement of hypomethylation of the host genome in anti-viral defense responses. The symptom severity could be proportional to the expression levels of host defense genes which are induced through demethylation in response to viral infection. The activities of certain geminiviral viral proteins might be recognized by the immune receptors from NUCLEOTIDE BINDING-LEUCINE RICH REPEAT (NB-LRR) family, which induce expression of defense genes in response to both non-viral and viral pathogens [8]. It remains to be investigated if the immune responses to viral infection require active demethylation of the host genome, triggered by recognition of viral proteins.

It has been reported that begomoviral Rep has the ability to reverse transcriptional silencing and reduce CG methylation at endogenous loci, possibly through Rep-mediated downregulation of *MET1* [66]. However, in addition to *MET1*, the transcript levels of *CMT3* and *ROS1* (but not *DRM2*) were also downregulated by transient expression of Rep or by geminivirus infection. It is unclear how the downregulation of the maintenance methyltransferases and the demethylase together would reduce CG methylation and reverse transcription silencing at the endogenous loci, and whether these effects of Rep are important for viral infection. The ability of geminiviral Rep to modify cell cycle and trigger host DNA reduplication [37,67,68] may explain the reduced levels of cytosine methylation at endogenous loci in the Rep transgenic plants upon induction of Rep expression [66]. As described above, Rep-mediated replication of TYLCCNV failed to suppress transcriptional silencing or cytosine methylation [53]. Moreover, geminivirus infection could induce transcriptional silencing of transgenes containing cognate geminiviral sequences, which correlated with hypermethylation of these sequences at CG, CHG and CHH sites [51,69]. This process of virus-induced transgene silencing did not affect geminivirus symptom development or viral DNA accumulation [51]. Thus, while Rep-mediated replication rescues viral DNA from repressive methylation, Rep activity does not prevent *de novo* methylation and silencing of the target transgenes.

8. Plant Recovery from Geminiviral Infection and RdDM

Recovery from virus disease symptoms correlating with reduced viral titres has been observed for RNA and DNA viruses in certain host plants. For RNA viruses that spawn massive quantities of 21-nt and/or 22-nt siRNAs [2–5], the post-transcriptional RNA silencing pathway appears to mediate plant recovery [70]. For geminiviruses that spawn 24-nt siRNAs in addition to 21- and 22-nt siRNAs [58,71–73], both post-transcriptional and transcriptional silencing pathways have been implicated in recovery [72].

Evidence for a role of transcriptional silencing and cytosine methylation in the recovery process comes from several studies. Thus, plant recovery from infections with cassava mosaic begomoviruses correlated with increased accumulation of viral siRNAs of all the size-classes [73], suggesting that viral 24-nt siRNA may direct transcriptional silencing and thereby contribute to recovery. Recovery of pepper plants from begomovirus infection was associated with much low titres of both viral DNA and siRNAs in the youngest recovered leaves, compared to those in the severely infected old leaves. The levels of viral DNA methylation in all the contexts were elevated from ca. 10% in the old leaves to ca. 20% in the youngest leaves [72]. The inverse correlation between 5meC levels and 24-nt siRNA quantities in the respective leaves raises a question whether cytosine methylation is established through the action of viral 24-nt siRNAs. In another study, the abundance of begomovirus-derived siRNAs was also negatively correlated with plant recovery and positively correlated with viral titre [56].

The involvement of RdDM in plant recovery was deduced from the observation that contrary to wild-type plants, the mutant plants lacking AGO4 could not recover from infection with the BCTV mutant lacking L2. This correlated with lower 5meC levels of viral DNA in the disease-displaying *ago4* mutant plants (20%) than in the recovered wild type plants (80%) [52]. However, no difference in cytosine methylation of the wild type BCTV in *ago4* mutant plants *versus* wild-type plants was observed (18% in both cases). Since an AGO4-siRNA complex is a major effector of RdDM (Figure 1), it remains unclear which mechanism mediates cytosine methylation of geminiviral DNA and whether this mechanism requires viral 24-nt siRNAs.

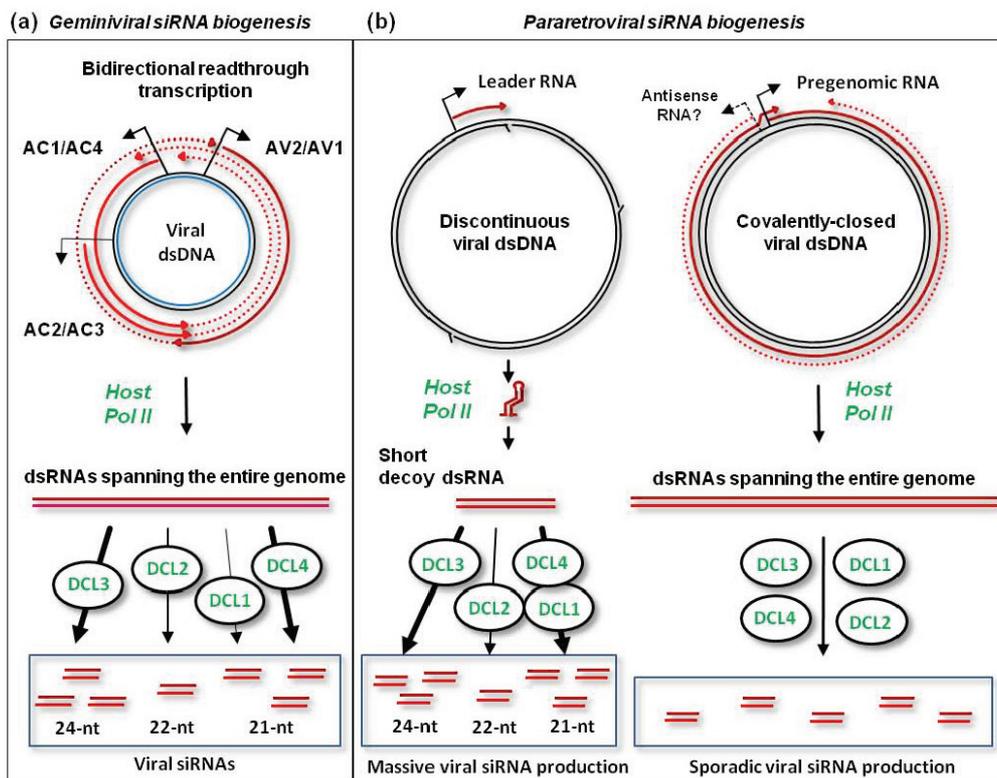
Plant recovery from geminivirus infection can also be triggered by transient or stable expression of inverted-repeat transgenes that generate dsRNA cognate to the geminivirus intergenic region [74,75]. However, it is unclear whether dsRNA-derived 24-nt siRNAs or dsRNA itself contributed to the recovery and whether *de novo* methylation of the viral DNA plays a role in this process. Interestingly, the intergenic region of geminiviruses is a poor source of siRNAs (see below). Thus, targeting this naturally-protected region by artificial dsRNA could help the plant recover from the viral disease.

9. Genetic Requirements for the Biogenesis of Geminiviral siRNAs

Geminivirus-infected plants produce abundant virus-derived 21-, 22- and 24-nt siRNAs. As evaluated by deep sequencing, a sub-population of viral siRNAs can vary from ca. 1%–3% to 30%–50% of the total small RNA population in infected plants [76–78]. Thus, despite a tiny size of the geminivirus genome, the quantity of viral siRNAs in some virus-host systems (e.g., CaLCuV-infected *Arabidopsis* [78]) is comparable to a combined quantity of siRNAs and miRNAs expressed from the plant genome. It should be noted that the percentage of viral siRNAs in a total sRNA population is

lower for those geminiviruses that are strictly limited to phloem tissues. Taking into account the dilution factor, geminivirus-infected phloem cells must produce massive amounts of viral siRNAs.

Figure 3. Models for the biogenesis of geminiviral and pararetroviral siRNAs. **(a)** The biogenesis of geminiviral siRNAs is initiated by bi-directional readthrough transcription beyond the poly(A) signals that normally terminate transcription of the viral leftward genes (in begomoviruses, AC1/Rep, AC4, AC2/TrAP and AC3) and the rightward genes (in begomoviruses, AV2 and AV1/CP). The resulting sense and antisense readthrough transcripts (dotted lines) anneal to the complementary viral mRNAs (solid lines with arrowheads) and to each other (in the intergenic region between the transcription start sites). This creates dsRNAs spanning the entire circular viral genome. Every DCL digests these dsRNAs into siRNAs of different sizes, with DCL3 (24-nt), DCL4 (21-nt) and DCL2 (22-nt) being favored (in that order); **(b)** Pol II transcribes both the discontinuous and the covalently-closed dsDNA forms of pararetroviral dsDNA. Abrupt termination of Pol II transcription at the unrepaired minus-strand DNA gap (Met-tRNA gap), results in production of aberrant 8S RNA lacking poly(A) tail (Leader RNA). This RNA forms a viroid-like secondary structure which can be converted by Pol II to dsRNA. The resulting dsRNA serves as a decoy to engage all the four DCLs in massive production of 21-, 22-, and 24-nt vsRNAs. Pol II-mediated transcription of the covalently-closed circular dsDNA generates pgRNA covering the entire genome as well as antisense transcript(s) (red dotted line). The 35S pgRNA promoter was reported to drive transcription not only in the forward but also in the reverse orientation [79] (indicated with bent lines with arrowheads). The pgRNA and antisense Pol II transcripts form low-abundance dsRNA spanning the entire virus genome. This dsRNA is diced by the four DCLs to generate viral 21, 22 and 24-nt siRNAs.



Although the hotspots of viral siRNA production are not equally distributed along the virus genome, unique (non-redundant) siRNAs of each size-class cover the entire circular viral DNA in both sense and antisense polarities, as demonstrated for CaLCuV DNA-A and DNA-B [78]. Based on genetic evidence combined with small RNA deep sequencing and blot hybridization, the biogenesis of CaLCuV siRNAs is mediated by all the four plant DCLs, but does not require RDR1, RDR2, RDR6, Pol IV, or Pol V [58,78]. The precursors of viral siRNAs are likely produced by Pol II-mediated bi-directional readthrough transcription of viral circular dsDNA far beyond the poly(A) signals (Figure 3A; further discussed in [4,78]). Such readthrough transcripts of sense and antisense polarities can potentially form dsRNA substrates for DCLs. Interestingly, the intergenic region harboring bidirectional promoter elements between the transcription start sites is a poor source of viral siRNAs [77,78]. This implies that the readthrough transcripts of each polarity are preferentially associated with more abundant viral mRNAs to form dsRNA. The resulting dsRNAs covering the leftward and rightward genes as well as less abundant dsRNAs covering the intergenic region are then processed by each of the four DCLs to generate 21-nt (DCL4 and DCL1), 22-nt (DCL2) and 24-nt (DCL3) siRNA duplexes [58,78] (Figure 3A). Both strands of these siRNA duplexes are then methylated at the 3'-terminal nucleotide's hydroxyl by HEN1 [58] and presumably sorted by AGO proteins to form silencing complexes. By using CaLCuV as a vector for virus-induced gene silencing (VIGS) targeting a host gene, it was demonstrated that viral siRNA generated by each DCL has the ability to knock down target mRNA accumulation [58]. Furthermore, CaLCuV-VIGS targeting an enhancer region of 35S promoter-driven transgene could induce transcriptional silencing of the transgene in virus-infected plants [78]. Whether viral 24-nt siRNAs get associated with AGO4 to direct *de novo* methylation and transcriptional silencing remains to be investigated.

The most abundant viral 24-nt siRNAs, which can potentially direct *de novo* methylation of viral DNA, map to the coding regions of the geminivirus genome [77,78], where cytosine methylation may not affect viral transcription. Indeed, substantial methylation is found in the bodies of active *Arabidopsis* genes, *i.e.*, downstream of their promoters. In contrast, inactive, developmentally-regulated and tissue-specific genes tend to have high levels of cytosine methylation in the promoters [9]. As noted above the geminiviral bidirectional promoter region spawns low amounts of 24-nt siRNAs, which may not be sufficient for RdDM and transcriptional silencing.

Taken together, the biogenesis of geminiviral 24-nt siRNAs does not involve the core components of the RdDM pathway such as RDR2, Pol IV, or Pol V. Since these components are required for the biogenesis and function of endogenous 24-nt siRNAs, the RdDM pathway may not be effective in targeting viral dsDNA for cytosine methylation. This is consistent with the findings that circular viral dsDNA is not methylated. It remains to be investigated whether detectable methylation of viral heterogeneous linear dsDNA is established through RdDM. It is feasible that Pol II-mediated transcription of viral linear dsDNA may lead to targeting of the nascent transcript by viral 24-nt siRNA-AGO4 complexes, which would recruit DRM2. However, the resulting methylated DNA may not be able to recruit the Pol IV-RDR2 complex for dsRNA production and siRNA amplification, since this complex does not contribute substantially to production of viral siRNAs [58,78].

Notably, in the absence of three functional RDRs (RDR1, RDR2, RDR6), accumulation of CaLCuV siRNAs of all sizes was elevated, which correlated with increased accumulation of some viral transcripts [78]. Southern blot analysis revealed increased accumulation of viral circular ssDNA,

but not circular dsDNA. This implies that plant RDR activities may repress both Pol II-mediated siRNA production via readthrough transcription and Rep-dependent production of circular ssDNA. Interestingly, a distinct plant siRNA-generating pathway has been implicated in recombination-dependent DNA repair [80]. It is tempting to speculate that viral siRNAs accumulating in the nucleus may facilitate viral DNA replication, e.g., by serving as primers for the host DNA polymerase.

10. Pararetrovirus Replication and Evasion of Transcriptional and Post-Transcriptional Silencing

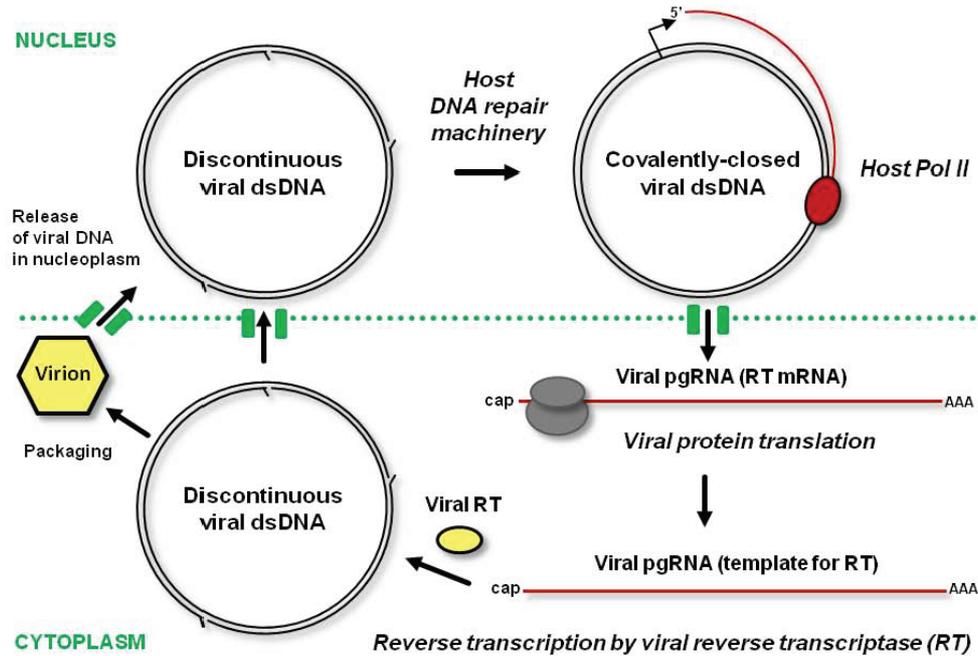
Plants do not host retroviruses, but their genomes are populated by LTR retrotransposons whose transcriptional activity is repressed by RdDM. Only episomal pararetroviruses that do not obligatorily integrate into the host genome can replicate and spread in plants. The family *Caulimoviridae* comprises several genera of pararetroviruses with circular dsDNA genomes of 7.4 to 8 kbp [81,82]. Like retroviruses, the pararetroviruses replicate via reverse transcription. The pararetroviral reverse transcriptase (RT) possesses RNA-dependent and DNA-dependent DNA polymerase activities and RNaseH activity, but lacks an integrase activity [81]. Nonetheless, some plant pararetroviruses have managed to integrate into the host genomes and form complex repetitive integration loci. Some of them, e.g., endogenous *Banana streak virus* and *Petunia vein clearing virus* (PVCV), can be released from the genome upon stress and cause disease [83,84].

The genomic DNA of episomal pararetroviruses is encapsidated in icosahedral or bacilliform virions and transmitted from plant to plant by insect vectors [81]. Like in geminiviruses, a nuclear localization signal of pararetroviral coat protein promotes delivery of viral DNA into the nucleus. The virion-associated circular dsDNA has at least one gap (discontinuity) in each strand, the remnants from reverse transcription of viral pregenomic RNA (pgRNA) in the cytoplasm [85]. These gaps are sealed in the nucleus by the host DNA repair machinery and the resulting covalently-closed circular dsDNA gets associated with nucleosomes to form a viral minichromosome, the template for Pol II transcription (Figure 4). Pol II generates a capped and polyadenylated pgRNA that covers the entire virus genome and has a terminal redundancy, owing to the recognition of the poly(A) signal located at a short distance downstream of the transcription start site only on a second encounter. In some genera, Pol II transcription also generates a subgenomic RNA, the mRNA for P6/TAV protein. This multifunctional protein is involved in formation of dense inclusion bodies in the cytoplasm, translation reinitiation and suppression of plant defenses (see below).

The pgRNA harboring all the viral ORFs serves as an mRNA for polycistronic translation of viral proteins (including coat protein and RT) and as a template for reverse transcription. Following translation in the cytoplasm, the pgRNA is reverse transcribed by viral RT enzymatic activities with the help of coat protein. The resulting open-circular dsDNA with gaps at both strands can be delivered into the nucleus by coat protein, or get incorporated into a mature virion, which can re-infect the same nucleus or move out of the cell (Figure 4). As a result of multiple rounds of replication as well as cell-to-cell and long-distance movement of virions, the infected cells' nuclei accumulate multiple copies of viral minichromosomes.

Figure 4. Model for pararetrovirus replication. Viral circular dsDNA from the virion (in yellow) is released into the nucleus. The gaps at both DNA strands are sealed by the host

DNA repair machinery. The resulting covalently-closed dsDNA serves as a template for Pol II transcription generating viral pregenomic RNA (pgRNA). The capped and polyadenylated pgRNA is transported to the cytoplasm for translation of viral proteins including the reverse transcriptase (RT), and for subsequent reverse transcription catalyzed by RT. The resulting dsDNA with discontinuities at both strands can get packaged into a new virion or targeted to the nucleus for the next round of replication.



The cytoplasmic step of viral replication through pgRNA should effectively protect viral DNA from maintenance methylation and RdDM. However, covalently-closed circular dsDNA, which is transcribed in the nucleus, can potentially be methylated *de novo* by the RdDM machinery charged with viral 24-nt siRNAs. If this is the case, even inefficient transcription of viral minichromosomes with the repressive marks will generate pgRNA, and the next round of pgRNA translation and reverse transcription will produce unmethylated viral dsDNA.

Deep-sequencing analysis of small RNAs from *Arabidopsis* plants infected with *Cauliflower mosaic virus* (CaMV), a type member of genus *Caulimovirus*, has demonstrated that 21-, 22- and 24-nt viral siRNAs accumulate in massive quantities comparable to the entire complement of endogenous plant siRNA and miRNAs [86]. Moreover, massive production of all size-classes of viral siRNAs of both sense and antisense polarities is largely restricted to a 600 bp non-coding region of the CaMV genome, between the pgRNA transcription start site and the reverse transcription primer binding site. Other genomic sequences spawn much less abundant siRNAs of each size-class and polarity. Given that Pol II-mediated transcription of the CaMV genome generating pgRNA and P6 mRNA is mono-directional, the precursors of viral siRNAs covering the entire genome in both polarities are likely generated by antisense transcription driven by cryptic promoter(s) on viral DNA (Figure 3B). Alternatively, host RDR activities may convert viral RNAs into dsRNA. However, genetic evidence combined with siRNA deep sequencing and blot hybridization ruled out this hypothesis. Indeed, the biogenesis of viral siRNAs from both hot and cold regions does not require RDR1, RDR2, or RDR6 [58,86]. Furthermore, Pol V and Pol IV do not contribute to CaMV siRNA

production. Hence, both sense and antisense strands of dsRNA precursors of viral siRNAs are likely generated by Pol II. The resulting dsRNAs are then processed by each of the four Dicers, which generate 21-nt (DCL1 and DCL4), 22-nt (DCL2) and 24-nt (DCL3) siRNAs [58,86]. DCL1, which normally generates plant miRNAs, produces a larger fraction of viral 21-nt siRNAs than DCL4 [58,86]. DCL4 is a primary dicer generating 21-nt siRNAs from RNA viruses [87] but its activity is inhibited by CaMV P6/TAV protein [88,89] (further discussed below).

The 600 bp non-coding region of CaMV genome generating the majority of viral siRNAs was proposed to produce a decoy dsRNA that would engage all the four DCLs and available AGOs in production and sorting of viral siRNAs [86] (Figure 3B). Such decoy strategy would protect other regions from silencing at both transcriptional and post-transcriptional levels. Indeed, the upstream pgRNA promoter elements and the downstream coding sequences spawn only small amounts of viral siRNAs that would have to compete with abundant, decoy dsRNA-derived siRNAs for AGOs to form silencing complexes. Consistent with the decoy model, immuno-precipitation with AGO-specific antibodies revealed that AGO1 is associated with 21-nt siRNAs from the non-coding region but not other regions of CaMV genome [86]. Surprisingly, only a tiny fraction of abundant 24-nt siRNAs from the non-coding region was associated with AGO4. AGO4 complexes in the nucleus are likely saturated with endogenous 24-nt siRNAs and only a small pool of free AGO4 is available. If the non-coding region becomes *de novo* methylated through the action of detectable silencing complexes, transcriptional activity of the upstream promoter will not be affected. At the post-transcriptional level, the 600 nt non-coding leader sequence of pgRNA folds into a stable secondary structure bypassed by ribosomes to initiate translation [90–92], which may not be accessible for 21-nt siRNA-AGO1 complexes. Taken together, the decoy strategy evolved by CaMV [86] and possibly other pararetroviruses with a similar configuration of the non-coding region elements and structures [93] would help the virus evade silencing at both transcriptional and post-transcriptional levels.

Like in the case of geminiviruses, pararetrovirus infection can induce silencing of transgenes sharing homology with the virus. In CaMV-infected plants, the transgenes driven by the CaMV 35S pgRNA promoter were silenced at the transcriptional levels, whereas those with the CaMV 3'UTR sequences at the post-transcriptional level [94,95]. Notably, CaMV replication and viral transcript accumulation were not affected by ongoing silencing of the transgenes [94]. Thus, CaMV can indeed evade both transcriptional and posttranscriptional silencing as argued above. It remains to be investigated if silencing of homologous transgenes is directed by viral siRNAs.

Some host plants can recover from pararetrovirus disease symptoms, but abundant viral dsDNA can still persist in the recovered tissues. The recovery of kohlrabi plants from CaMV infection was preceded by overaccumulation of covalently-closed viral dsDNA in the nucleus, followed by arrest of reverse transcription [96]. Interestingly, overall transcription of viral dsDNA in the nucleus (evaluated by a “nuclear run-on” method) did not change after the transition to recovery, but accumulation of polyadenylated viral transcripts was strongly reduced. This implicates post-transcriptional silencing in the recovery process. Notably, covalently-closed viral dsDNA was not found to be methylated before or after recovery [96]. The mechanisms underlying the overaccumulation of viral minichromosomes before recovery and the posttranscriptional degradation of viral RNAs remain to be further investigated.

Endogenous pararetroviruses integrated in the host genomes are likely repressed by cytosine methylation and histone modifications. These repressive marks can potentially be established *de novo*

by RdDM and efficiently maintained following plant DNA replication. The integrated copies of PVCV in the petunia genome were found to be associated with repressive H3K9me2 marks [84]. In this case, accumulation of 21–24 nt viral siRNAs was barely detectable, and only disease induction could boost viral siRNA production. Hence, the released episomal virus spawns much more abundant siRNAs than the integrated copies. Deep-sequencing of siRNAs combined with cytosine methylation analysis should clarify whether the infectious copies of integrated pararetroviral DNA are densely methylated and whether cytosine methylation is established and maintained by RdDM. In the case of an endogenous tomato pararetrovirus, which cannot be released as episomal virus, the integrated viral sequences were found to be methylated at CHG and CHH sites and virus-derived 21–24 nt siRNAs accumulated at detectable levels [97].

11. Suppression of Plant Defenses by Pararetroviral Proteins

CaMV P6/TAV protein has been implicated in suppression of the plant defenses based on RNA silencing [88,89,98] and innate immunity [8,99]. Since this protein has no homologs in several genera of *Caulimoviridae* [81], plant pararetroviruses have to rely on other strategies to suppress or evade plant defenses. The tungrovirus *Rice tungro bacilliform virus* possesses a P4 gene of unknown function, which is missing in closely related badnaviruses. Like CaMV P6 gene, the P4 gene is located downstream of the RT gene and expressed from a separate mRNA [81]. These similarities suggest that P4 may have been acquired by a badnavirus to cope with plant defenses in a new host.

The mechanism of silencing suppression by CaMV P6/TAV has been extensively investigated [88,89,98]. According to the current model, P6 interferes with amplification of secondary siRNAs by blocking DCL4-mediated processing of RDR6-dependent dsRNAs. Curiously, nuclear import of P6 was required for P6-mediated suppression of endogenous tasiRNA biogenesis, which presumably occurs in the cytoplasm, and for P6 interaction with DOUBLE-STRANDED RNA BINDING 4 (DRB4), a partner of DCL4 [88]. It remains to be demonstrated if these activities of CaMV P6 are also required for suppression of antiviral silencing. Indirect evidence supporting this hypothesis is that only a fraction of CaMV 21-nt siRNAs is produced by DCL4 and the biogenesis of the DCL4-dependent fraction of viral 21-nt siRNAs does not require RDR6 activity [58,86]. However, presumptive RDR6-dependent precursors of viral secondary siRNAs, which should be stabilized by the P6 action, could not be detected in CaMV-infected plants. In contrast the RDR6-dependent dsRNA precursors of plant tasiRNAs are readily detectable in both CaMV-infected and P6 transgenic plants [19,20,58,89]. Thus, CaMV infection or P6 expression does not interfere with RDR6 activity, but viral mRNAs (and their degradation products) appear to be poor substrates for RDR6. Similar findings have been reported for the geminivirus CaLCuV [78]. Thus, DNA viruses have evolved to protect their mRNAs from RDR activity that would amplify and spread antiviral siRNAs. Likewise, most of the plant genes controlled by miRNAs do not spawn RDR6-dependent secondary siRNAs.

The caulimovirus P6/TAV is a multifunctional protein harboring the domains implicated in interactions with ribosomal proteins and translation initiation factors, in binding RNA (single and double-stranded), in formation of inclusion bodies, and in hypersensitive immune responses [100–102]. The domain responsible for suppression of RNA silencing has not been identified yet.

It has been hypothesized that P6/TAV and suppressor proteins of other plant viruses can interfere with the innate immune responses, which restrict growth of non-viral pathogens [8]. In resistant hosts, CaMV P6 triggers hypersensitive responses and its avirulence domain recognized by the immune system has been mapped. Notably, this P6 domain is also required for CaMV virulence in susceptible hosts [101]. By analogy with effector proteins of non-viral pathogens, a primary function for P6 is to suppress basal immune responses. In resistant hosts, P6 effector activity is recognized by the immune receptors of the NB-LRR family, which triggers hypersensitive response and programmed cell death restricting viral infection [8]. This hypothesis is supported by the finding that CaMV P6 expression in transgenic plants promotes growth of a bacterial pathogen [99]. The mechanism of P6 interference with the immune responses remains to be investigated.

12. Concluding Remarks

Unlike animals, land plants do not host “true” dsDNA viruses whose replication mechanisms generate dsDNA genome copies without a ssDNA or RNA intermediate, or “true” retroviruses with a provirus stage of replication that involves viral DNA integration in the host genome. This exclusion is likely because the land plants have evolved the mechanisms of siRNA-directed *de novo* methylation of all cytosines (RdDM) and maintenance methylation at both CG and non-CG sites. These mechanisms establish and maintain cytosine methylation in all sequence contexts of the plant genome and thereby effectively repress unwanted transcription in the nucleus. This repressive methylation system is evaded by ssDNA viruses which can resurrect their dsDNA forms from cytosine methylation by Rep-dependent replication generating unmethylated ssDNA. Likewise, pararetroviruses that omit a host genome integration step can thereby evade the transcriptional silencing reinforced by a concert action of maintenance methylation and RdDM-dependent amplification of siRNAs. Furthermore, episomal pararetroviruses can evade repressive methylation by constant delivery of multiple unmethylated copies of circular dsDNA to the nucleus from the cytoplasm where pgRNA is reverse transcribed. Having these replication strategies, pararetroviruses and ssDNA viruses have not been under a strong pressure in land plants to evolve suppressors of cytosine methylation.

Acknowledgments

I thank Thomas Boller for supporting research of MMP group at the University of Basel and Rajendran Rajeswaran for fruitful collaboration and critical reading of the manuscript. Special thanks to the anonymous reviewer for constructive criticism and ideas. The work of MMP group is being funded by Swiss National Science Foundation (grant 31003A_143882/1), Indo-Swiss Collaboration in Biotechnology, and European Collaboration in Science and Technology (COST grant SER No. C09.0176 and COST Action FA0806).

Conflict of Interest

The author declares no conflict of interest.

References

1. Gronenborn, B. Nanoviruses: Genome organisation and protein function. *Vet. Microbiol.* **2004**, *98*, 103–109.
2. Llave, C. Virus-derived small interfering RNAs at the core of plant-virus interactions. *Trends Plant Sci.* **2010**, *15*, 701–707.
3. Pantaleo, V. Plant RNA silencing in viral defence. *Adv. Exp. Med. Biol.* **2011**, *722*, 39–58.
4. Rajeswaran, R.; Pooggin, M.M. Role of Virus-Derived Small RNAs in Plant Antiviral Defense: Insights from DNA Viruses. In *MicroRNAs in Plant Development and Stress Response*; Sunkar, R., Ed.; Springer-Verlag: Berlin Heidelberg, Germany, 2012; pp. 261–289.
5. Szittyá, G.; Burgyán, J. RNA interference-mediated intrinsic antiviral immunity in plants. *Curr. Top. Microbiol. Immunol.* **2013**, *371*, 153–181.
6. Burgyán, J.; Havelda, Z. Viral suppressors of RNA silencing. *Trends Plant Sci.* **2011**, *16*, 265–272.
7. Incarbone, M.; Dunoyer, P. RNA silencing and its suppression: Novel insights from in planta analyses. *Trends Plant Sci.* **2013**, doi:10.1038/nrg1555.
8. Zvereva, A.S.; Pooggin, M.M. Silencing and innate immunity in plant defense against viral and non-viral pathogens. *Viruses* **2012**, *4*, 2578–2597.
9. Law, J.A.; Jacobsen, S.E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* **2010**, *11*, 204–220.
10. Klose, R.J.; Bird, A.P. Genomic DNA methylation: The mark and its mediators. *Trends Biochem. Sci.* **2006**, *31*, 89–97.
11. Woo, H.R.; Dittmer, T.A.; Richards, E.J. Three SRA-domain methylcytosine-binding proteins cooperate to maintain global CpG methylation and epigenetic silencing in *Arabidopsis*. *PLoS Genet.* **2008**, *4*, e1000156.
12. Johnson, L.M.; Bostick, M.; Zhang, X.; Kraft, E.; Henderson, I.; Callis, J.; Jacobsen, S.E. The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr. Biol.* **2007**, *17*, 379–384.
13. Zemach, A.; Kim, M.Y.; Hsieh, P.H.; Coleman-Derr, D.; Eshed-Williams, L.; Thao, K.; Harmer, S.L.; Zilberman, D. The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **2013**, *153*, 193–205.
14. Pikaard, C.S. Methylating the DNA of the most repressed: Special access required. *Mol. Cell* **2013**, *49*, 1021–1022.
15. Pikaard, C.S.; Haag, J.R.; Pontes, O.M.; Blevins, T.; Cocklin, R. A transcription fork model for Pol IV and Pol V-dependent RNA-Directed DNA methylation. *Cold Spring Harb. Symp. Quant. Biol.* **2013**, doi:10.1101/sqb.2013.77.014803.
16. Mallory, A.; Vaucheret, H. Form, function, and regulation of ARGONAUTE proteins. *Plant Cell* **2010**, *22*, 3879–3889.
17. Carbonell, A.; Fahlgren, N.; Garcia-Ruiz, H.; Gilbert, K.B.; Montgomery, T.A.; Nguyen, T.; Cuperus, J.T.; Carrington, J.C. Functional analysis of three *Arabidopsis* ARGONAUTES using slicer-defective mutants. *Plant Cell* **2012**, *24*, 3613–3629.
18. Havecker, E.R.; Wallbridge, L.M.; Hardcastle, T.J.; Bush, M.S.; Kelly, K.A.; Dunn, R.M.; Schwach, F.; Doonan, J.H.; Baulcombe, D.C. The *Arabidopsis* RNA-directed DNA methylation

- argonautes functionally diverge based on their expression and interaction with target loci. *Plant Cell* **2010**, *22*, 321–334.
19. Rajeswaran, R.; Pooggin, M.M. RDR6-mediated synthesis of complementary RNA is terminated by miRNA stably bound to template RNA. *Nucleic Acids Res.* **2012**, *40*, 594–599.
 20. 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.
 21. Wang, X.B.; Wu, Q.; Ito, T.; Cillo, F.; Li, W.X.; Chen, X.; Yu, J.L.; Ding, S.W. RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 484–489.
 22. Wang, X.B.; Jovel, J.; Udornporn, P.; Wang, Y.; Wu, Q.; Li, W.X.; Gascioli, V.; Vaucheret, H.; Ding, S.W. The 21-Nucleotide, but Not 22-Nucleotide, viral secondary small interfering RNAs direct potent antiviral defense by two cooperative argonautes in *Arabidopsis thaliana*. *Plant Cell* **2011**, *23*, 1625–1638.
 23. Garcia-Ruiz, H.; Takeda, A.; Chapman, E.J.; Sullivan, C.M.; Fahlgren, N.; Brempelis, K.J.; Carrington, J.C. *Arabidopsis* RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. *Plant Cell* **2010**, *22*, 481–496.
 24. Haag, J.R.; Ream, T.S.; Marasco, M.; Nicora, C.D.; Norbeck, A.D.; Pasa-Tolic, L.; Pikaard, C.S. *In vitro* transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. *Mol. Cell* **2012**, *48*, 811–818.
 25. Law, J.A.; Du, J.; Hale, C.J.; Feng, S.; Krajewski, K.; Palanca, A.M.; Strahl, B.D.; Patel, D.J.; Jacobsen, S.E. Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* **2013**, doi:10.1038/nature12178.
 26. Zheng, B.; Wang, Z.; Li, S.; Yu, B.; Liu, J.Y.; Chen, X. Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. *Genes Dev.* **2009**, *23*, 2850–2860.
 27. Stroud, H.; Greenberg, M.V.; Feng, S.; Bernatavichute, Y.V.; Jacobsen, S.E. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. *Cell* **2013**, *152*, 352–364.
 28. You, W.; Lorkovic, Z.J.; Matzke, A.J.; Matzke, M. Interplay among RNA polymerases II, IV and V in RNA-directed DNA methylation at a low copy transgene locus in *Arabidopsis thaliana*. *Plant Mol. Biol.* **2013**, *82*, 85–96.
 29. Schauer, S.E.; Jacobsen, S.E.; Meinke, D.W.; Ray A. DICER-LIKE1: Blind men and elephants in *Arabidopsis* development. *Trends Plant Sci.* **2002**, *7*, 487–491.
 30. Furner, I.J.; Matzke, M. Methylation and demethylation of the *Arabidopsis* genome. *Curr. Opin. Plant Biol.* **2011**, *14*, 137–141.
 31. Penterman, J.; Zilberman, D.; Huh, J.H.; Ballinger, T.; Henikoff, S.; Fischer, R.L. DNA demethylation in the *Arabidopsis* genome. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 6752–6757.
 32. Zhu, J.; Kapoor, A.; Sridhar, V.V.; Agius, F.; Zhu, J.K. The DNA glycosylase/lyase ROS1 functions in pruning DNA methylation patterns in *Arabidopsis*. *Curr. Biol.* **2007**, *17*, 54–59.

33. Ortega-Galisteo, A.P.; Morales-Ruiz, T.; Ariza, R.R.; Roldán-Arjona, T. *Arabidopsis* DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks. *Plant Mol. Biol.* **2008**, *67*, 671–681.
34. Bhutani, N.; Burns, D.M.; Blau, H.M. DNA demethylation dynamics. *Cell* **2011**, *146*, 866–872.
35. Li, X.; Qian, W.; Zhao, Y.; Wang, C.; Shen, J.; Zhu, J.K.; Gong, Z. Antisilencing role of the RNA-directed DNA methylation pathway and a histone acetyltransferase in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11425–11430.
36. Fauquet, C.M.; Briddon, R.W.; Brown, J.K.; Moriones, E.; Stanley, J.; Zerbini, M.; Zhou, X. Geminivirus strain demarcation and nomenclature. *Arch. Virol.* **2008**, *153*, 783–821.
37. Hanley-Bowdoin, L.; Settlage, S.B.; Orozco, B.M.; Nagar, S.; Robertson, D. Geminiviruses: Models for plant DNA replication, transcription, and cell cycle regulation. *Crit. Rev. Plant Sci.* **1999**, *18*, 71–106.
38. Jeske, H. Geminiviruses. *Curr. Top. Microbiol. Immunol.* **2009**, *331*, 185–226.
39. Saunders, K.; Lucy, A.; Stanley, J. RNA-Primed complementary-sense DNA synthesis of the geminivirus African cassava mosaic virus. *Nucleic Acids Res.* **1992**, *20*, 6311–6315.
40. Andersen, M.T.; Richardson, K.A.; Harbison, S.A.; Morris, B.A. Nucleotide sequence of the geminivirus chloris striate mosaic virus. *Virology* **1988**, *164*, 443–449.
41. Pilartz, M.; Jeske, H. Abutilon mosaic geminivirus double-stranded DNA is packed into minichromosomes. *Virology* **1992**, *189*, 800–802.
42. Pilartz, M.; Jeske, H. Mapping of abutilon mosaic geminivirus minichromosomes. *J. Virol.* **2003**, *77*, 10808–10818.
43. Shivaprasad, P.V.; Akbergenov, R.; Trinks, D.; Rajeswaran, R.; Veluthambi, K.; Hohn, T.; Pooggin, M.M. Promoters, transcripts, and regulatory proteins of Mungbean yellow mosaic geminivirus. *J. Virol.* **2005**, *79*, 8149–8163.
44. Singh, D.K.; Malik, P.S.; Choudhury, N.R.; Mukherjee, S.K. MYMIV replication initiator protein (Rep): Roles at the initiation and elongation steps of MYMIV DNA replication. *Virology* **2008**, *380*, 75–83.
45. Jeske, H.; Lütgemeier, M.; Preiss, W. DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. *EMBO J.* **2001**, *20*, 6158–6167.
46. Preiss, W.; Jeske, H. Multitasking in replication is common among geminiviruses. *J. Virol.* **2003**, *77*, 2972–2980.
47. Erdmann, J.B.; Shepherd, D.N.; Martin, D.P.; Varsani, A.; Rybicki, E.P.; Jeske, H. Replicative intermediates of maize streak virus found during leaf development. *J. Gen. Virol.* **2010**, *91*, 1077–1081.
48. Paprotka, T.; Deuschle, K.; Metzler, V.; Jeske, H. Conformation-selective methylation of geminivirus DNA. *J. Virol.* **2011**, *85*, 12001–12012.
49. Stenger, D.C.; Revington, G.N.; Stevenson, M.C.; Bisaro, D.M. Replication release of geminivirus genomes from tandemly repeated copies: Evidence for a rolling circle replication of a plant viral DNA. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 8029–8033.
50. Raja, P.; Wolf, J.N.; Bisaro, D.M. RNA silencing directed against geminiviruses: Post-Transcriptional and epigenetic components. *Biochim. Biophys. Acta* **2010**, *1799*, 337–351.

51. Bian, X.Y.; Rasheed, M.S.; Seemanpillai, M.J.; Rezaian, M.A. Analysis of silencing escape of tomato leaf curl virus: An evaluation of the role of DNA methylation. *Mol. Plant-Microbe Interact.* **2006**, *19*, 614–624.
52. Raja, P.; Sanville, B.C.; Buchmann, R.C.; Bisaro, D.M. Viral genome methylation as an epigenetic defense against geminiviruses. *J. Virol.* **2008**, *82*, 8997–9007.
53. Yang, X.; Xie, Y.; Raja, P.; Li, S.; Wolf, J.N.; Shen, Q.; Bisaro, D.M.; Zhou, X. Suppression of methylation-mediated transcriptional gene silencing by β C1-SAHH protein interaction during geminivirus-betasatellite infection. *PLoS Pathog.* **2011**, *7*, e1002329.
54. Brough, C.L.; Gardiner, W.E.; Inamdar, N.M.; Zhang, X.Y.; Ehrlich, M.; Bisaro, D.M. DNA methylation inhibits propagation of tomato golden mosaic virus DNA in transfected protoplasts. *Plant Mol. Biol.* **1992**, *18*, 703–712.
55. Ermak, G.; Paszkowski, U.; Wohlmuth, M.; Mittelsten Scheid, O.; Paszkowski, J. Cytosine methylation inhibits replication of African cassava mosaic virus by two distinct mechanisms. *Nucleic Acids Res.* **1993**, *21*, 3445–3450.
56. Hagen, C.; Rojas, M.R.; Kon, T.; Gilbertson, R.L. Recovery from Cucurbit leaf crumple virus (family *Geminiviridae*, genus *Begomovirus*) infection is an adaptive antiviral response associated with changes in viral small RNAs. *Phytopathology* **2008**, *98*, 1029–1037.
57. Buchmann, R.C.; Asad, S.; Wolf, J.N.; Mohannath, G.; Bisaro, D.M. Geminivirus AL2 and L2 proteins suppress transcriptional gene silencing and cause genome-wide reductions in cytosine methylation. *J. Virol.* **2009**, *83*, 5005–5013.
58. Blevins, T.; Rajeswaran, R.; Shivaprasad, P.V.; Beknazariants, D.; Si-Ammour, A.; Park, H.S.; Vazquez, F.; Robertson, D.; Meins, F., Jr.; Hohn, T.; *et al.* Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Res.* **2006**, *34*, 6233–6246.
59. Dong, X.; van Wezel, R.; Stanley, J.; Hong, Y. Functional characterization of the nuclear localization signal for a suppressor of posttranscriptional gene silencing. *J. Virol.* **2003**, *77*, 7026–7033.
60. Trinks, D.; Rajeswaran, R.; Shivaprasad, P.V.; Akbergenov, R.; Oakeley, E.J.; Veluthambi, K.; Hohn, T.; Pooggin, M.M. Suppression of RNA silencing by a geminivirus nuclear protein, AC2, correlates with transactivation of host genes. *J. Virol.* **2005**, *79*, 2517–2527.
61. Wang, H.; Buckley, K.J.; Yang, X.; Buchmann, R.C.; Bisaro, D.M. Adenosine kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins. *J. Virol.* **2005**, *79*, 7410–7418.
62. Zhang, Z.; Chen, H.; Huang, X.; Xia, R.; Zhao, Q.; Lai, J.; Teng, K.; Li, Y.; Liang, L.; Du, Q.; *et al.* BSCTV C2 attenuates the degradation of SAMDC1 to suppress DNA methylation-mediated gene silencing in Arabidopsis. *Plant Cell* **2011**, *23*, 273–288.
63. Briddon, R.W.; Stanley, J. Subviral agents associated with plant single-stranded DNA viruses. *Virology* **2006**, *344*, 198–210.
64. Cui, X.; Li, G.; Wang, D.; Hu, D.; Zhou, X. A Begomovirus DNAbeta-encoded protein binds DNA, functions as a suppressor of RNA silencing, and targets the cell nucleus. *J. Virol.* **2005**, *79*, 10764–10775.

65. Eini, O.; Dogra, S.C.; Dry, I.B.; Randles, J.W. Silencing suppressor activity of a begomovirus DNA β encoded protein and its effect on heterologous helper virus replication. *Virus Res.* **2012**, *167*, 97–101.
66. Rodríguez-Negrete, E.; Lozano-Durán, R.; Piedra-Aguilera, A.; Cruzado, L.; Bejarano, E.R.; Castillo, A.G. Geminivirus Rep protein interferes with the plant DNA methylation machinery and suppresses transcriptional gene silencing. *New Phytol.* **2013**, *199*, 464–475.
67. Nagar, S.; Hanley-Bowdoin, L.; Robertson, D. Host DNA replication is induced by geminivirus infection of differentiated plant cells. *Plant Cell* **2002**, *14*, 2995–3007.
68. Kittelmann, K.; Rau, P.; Gronenborn, B.; Jeske, H. Plant geminivirus rep protein induces rereplication in fission yeast. *J. Virol.* **2009**, *83*, 6769–6778.
69. Seemanpillai, M.; Dry, I.; Randles, J.; Rezaian, A. Transcriptional silencing of geminiviral promoter-driven transgenes following homologous virus infection. *Mol. Plant Microbe Interact.* **2003**, *16*, 429–438.
70. Ratcliff, F.G.; MacFarlane, S.A.; Baulcombe, D.C. Gene silencing without DNA. RNA-mediated cross-protection between viruses. *Plant Cell* **1999**, *11*, 1207–1216.
71. Akbergenov, R.; Si-Ammour, A.; Blevins, T.; Amin, I.; Kutter, C.; Vanderschuren, H.; Zhang, P.; Gruissem, W.; Meins, F., Jr.; Hohn, T.; *et al.* Molecular characterization of geminivirus-derived small RNAs in different plant species. *Nucleic Acids Res.* **2006**, *34*, 462–471.
72. Rodríguez-Negrete, E.A.; Carrillo-Tripp, J.; Rivera-Bustamante, R.F. RNA silencing against geminivirus: Complementary action of posttranscriptional gene silencing and transcriptional gene silencing in host recovery. *J. Virol.* **2009**, *83*, 1332–1340.
73. Chellappan, P.; Vanitharani, R.; Pita, J.; Fauquet, C.M. Short interfering RNA accumulation correlates with host recovery in DNA virus-infected hosts, and gene silencing targets specific viral sequences. *J. Virol.* **2004**, *78*, 7465–7477.
74. Pooggin, M.; Shivaprasad, P.V.; Veluthambi, K.; Hohn, T. RNAi targeting of DNA virus in plants. *Nat. Biotechnol.* **2003**, *21*, 131–132.
75. Vanderschuren, H.; Akbergenov, R.; Pooggin, M.M.; Hohn, T.; Gruissem, W.; Zhang, P. Transgenic cassava resistance to African cassava mosaic virus is enhanced by viral DNA-A bidirectional promoter-derived siRNAs. *Plant Mol. Biol.* **2007**, *64*, 549–557.
76. Donaire, L.; Wang, Y.; Gonzalez-Ibeas, D.; Mayer, K.F.; Aranda, M.A.; Llave, C. Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes. *Virology* **2009**, *392*, 203–214.
77. Yang, X.; Wang, Y.; Guo, W.; Xie, Y.; Xie, Q.; Fan, L.; Zhou, X. Characterization of small interfering RNAs derived from the geminivirus/betasatellite complex using deep sequencing. *PLoS One* **2011**, *6*, e16928.
78. Aregger, M.; Borah, B.K.; Seguin, J.; Rajeswaran, R.; Gubaeva, E.G.; Zvereva, A.S.; Windels, D.; Vazquez, F.; Blevins, T.; Farinelli, L.; *et al.* Primary and secondary siRNAs in geminivirus-induced gene silencing. *PLoS Pathog.* **2012**, *8*, e1002941.
79. Xie, M.; He, Y.; Gan, S. Bidirectionalization of polar promoters in plants. *Nat. Biotechnol.* **2001**, *19*, 677–679.

80. Wei, W.; Ba, Z.; Gao, M.; Wu, Y.; Ma, Y.; Amiard, S.; White, C.I.; Rendtlew Danielsen, J.M.; Yang, Y.G.; Qi, Y. A role for small RNAs in DNA double-strand break repair. *Cell* **2012**, *149*, 101–112.
81. Hull, R. Caulimoviridae (Plant Pararetroviruses) In *eLS*; John Wiley & Sons Ltd: Chichester, UK, 2007.
82. Rothnie, H.M.; Chapdelaine, Y.; Hohn, T. Pararetroviruses and retroviruses: A comparative review of viral structure and gene expression strategies. *Adv. Virus Res.* **1994**, *44*, 1–67.
83. Chabannes, M.; Baurens, F.C.; Duroy, P.O.; Bocs, S.; Vernerey, M.S.; Rodier-Goud, M.; Barbe, V.; Gayral, P.; Iskra-Caruana, M.L. Three infectious viral species lying in wait in the banana genome. *J. Virol.* **2013**, doi:10.1128/JVI.00899-13.
84. Noreen, F.; Akbergenov, R.; Hohn, T.; Richert-Pöggeler, K.R. Distinct expression of endogenous *Petunia* vein clearing virus and the DNA transposon dTph1 in two *Petunia hybrida* lines is correlated with differences in histone modification and siRNA production. *Plant J.* **2007**, *50*, 219–229.
85. Pfeiffer, P.; Hohn, T. Involvement of reverse transcription in the replication of cauliflower mosaic virus: A detailed model and test of some aspects. *Cell* **1983**, *33*, 781–789.
86. Blevins, T.; Rajeswaran, R.; Aregger, M.; Borah, B.K.; Schepetilnikov, M.; Baerlocher, L.; Farinelli, L.; Meins, F., Jr.; Hohn, T.; Pooggin, M.M. Massive production of small RNAs from a non-coding region of Cauliflower mosaic virus in plant defense and viral counter-defense. *Nucleic Acids Res.* **2011**, *39*, 5003–5014.
87. Deleris, A.; Gallego-Bartolome, J.; Bao, J.; Kasschau, K.D.; Carrington, J.C.; Voinnet, O. Hierarchical action and inhibition of plant dicer-like proteins in antiviral defense. *Science* **2006**, *313*, 68–71.
88. Haas, G.; Azevedo, J.; Moissiard, G.; Geldreich, A.; Himber, C.; Bureau, M.; Fukuhara, T.; Keller, M.; Voinnet, O. Nuclear import of CaMV P6 is required for infection and suppression of the RNA silencing factor DRB4. *EMBO J.* **2008**, *27*, 2102–2112.
89. Shivaprasad, P.V.; Rajeswaran, R.; Blevins, T.; Schoelz, J.; Meins, F., Jr.; Hohn, T.; Pooggin, M.M. The CaMV transactivator/viroplasm interferes with RDR6-dependent trans-acting and secondary siRNA pathways in Arabidopsis. *Nucleic Acids Res.* **2008**, *36*, 5896–5909.
90. Pooggin, M.M.; Futterer, J.; Skryabin, K.G.; Hohn, T. Ribosome shunt is essential for infectivity of cauliflower mosaic virus. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 886–891.
91. Pooggin, M.M.; Fütterer, J.; Hohn, T. Cross-species functionality of pararetroviral elements driving ribosome shunting. *PLoS One* **2008**, *3*, e1650.
92. Pooggin, M.M.; Rajeswaran, R.; Schepetilnikov, M.V.; Ryabova, L.A. Short ORF-dependent ribosome shunting operates in an RNA picorna-like virus and a DNA pararetrovirus that cause rice tungro disease. *PLoS Pathog.* **2012**, *8*, e1002568.
93. Pooggin, M.M.; Fütterer, J.; Skryabin, K.G.; Hohn, T. A short open reading frame terminating in front of a stable hairpin is the conserved feature in pregenomic RNA leaders of plant pararetroviruses. *J. Gen. Virol.* **1999**, *80*, 2217–2228.
94. Al-Kaff, N.S.; Covey, S.N.; Kreike, M.M.; Page, A.M.; Pinder, R.; Dale, P.J. Transcriptional and posttranscriptional plant gene silencing in response to a pathogen. *Science* **1998**, *279*, 2113–2115.

95. Al-Kaff, N.S.; Kreike, M.M.; Covey, S.N.; Pitcher, R.; Page, A.M.; Dale, P.J. Plants rendered herbicide-susceptible by cauliflower mosaic virus-elicited suppression of a 35S promoter-regulated transgene. *Nat. Biotechnol.* **2000**, *18*, 995–999.
96. Covey, S.N.; Al-Kaff, N.S.; Lángara, A.; Turner, D.S. Plants combat infection by gene silencing. *Nature* **1997**, *385*, 781–782.
97. Staginnus, C.; Gregor, W.; Mette, M.F.; Teo, C.H.; Borroto-Fernández, E.G.; Machado, M.L.; Matzke, M.; Schwarzacher, T. Endogenous pararetroviral sequences in tomato (*Solanum lycopersicum*) and related species. *BMC Plant Biol.* **2007**, *7*, 24.
98. Love, A.J.; Laird, J.; Holt, J.; Hamilton, A.J.; Sadanandom, A.; Milner, J.J. Cauliflower mosaic virus protein P6 is a suppressor of RNA silencing. *J. Gen. Virol.* **2007**, *88*, 3439–3444.
99. Love, A.J.; Geri, C.; Laird, J.; Carr, C.; Yun, B.W.; Loake, G.J.; Tada, Y.; Sadanandom, A.; Milner, J.J. Cauliflower mosaic virus protein P6 inhibits signaling responses to salicylic acid and regulates innate immunity. *PLoS One* **2012**, *7*, e47535.
100. Ryabova, L.A.; Pooggin, M.M.; Hohn, T. Viral strategies of translation initiation: Ribosomal shunt and reinitiation. *Prog. Nucleic Acid Res. Mol. Biol.* **2002**, *72*, 1–39.
101. Kobayashi, K.; Hohn, T. The avirulence domain of Cauliflower mosaic virus transactivator/viroplasmin is a determinant of viral virulence in susceptible hosts. *Mol. Plant Microbe Interact.* **2004**, *17*, 475–483.
102. Schepetilnikov, M.; Kobayashi, K.; Geldreich, A.; Caranta, C.; Robaglia, C.; Keller, M.; Ryabova, L.A. Viral factor TAV recruits TOR/S6K1 signalling to activate reinitiation after long ORF translation. *EMBO J.* **2011**, *30*, 1343–1356.

Reprinted from *IJMS*. Cite as: Gébelin, V.; Leclerc, J.; Hu, S.; Tang, C.; Montoro, P. Regulation of *MIR* Genes in Response to Abiotic Stress in *Hevea brasiliensis*. *Int. J. Mol. Sci.* **2013**, *14*, 19587-19604.

Article

Regulation of *MIR* Genes in Response to Abiotic Stress in *Hevea brasiliensis*

Virginie Gébelin^{1,†}, Julie Leclercq^{1,†}, Songnian Hu², Chaorong Tang³ and Pascal Montoro^{1,*}

¹ Centre de Coopération Internationale en Recherche Agronomique Pour le Développement, Unité Mixte de Recherche Amélioration Génétique et Adaptation des Plantes Méditerranéennes et Tropicales, Avenue Agropolis, Montpellier F-34398, France;

E-Mails: gebelin.v@gmail.com (V.G.); julie.leclercq@cirad.fr (J.L.)

² Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029, China;

E-Mail: husn@big.ac.cn

³ Rubber Research Institute, Chinese Academy of Tropical Agricultural Sciences, Danzhou 531737, China; E-Mail: chaorongtang@126.com

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: pascal.montoro@cirad.fr; Tel.: +33-0-467-61-5682; Fax: +33-0-467-61-5605.

Received: 31 May 2013; in revised form: 6 September 2013 / Accepted: 16 September 2013 /

Published: 27 September 2013

Abstract: Increasing demand for natural rubber (NR) calls for an increase in latex yield and also an extension of rubber plantations in marginal zones. Both harvesting and abiotic stresses lead to tapping panel dryness through the production of reactive oxygen species. Many microRNAs regulated during abiotic stress modulate growth and development. The objective of this paper was to study the regulation of microRNAs in response to different types of abiotic stress and hormone treatments in *Hevea*. Regulation of *MIR* genes differs depending on the tissue and abiotic stress applied. A negative co-regulation between *HbMIR398b* with its *chloroplastic HbCuZnSOD* target messenger is observed in response to salinity. The involvement of *MIR* gene regulation during latex harvesting and tapping panel dryness (TPD) occurrence is further discussed.

Keywords: gene expression; miRNA; *MIR* gene; abiotic stress; rubber tree; tapping panel dryness

1. Introduction

Hevea brasiliensis is the only source of natural rubber (NR), which is produced in latex cells in the phloem. Latex harvesting consists in tapping the soft bark of rubber trees. Ethephon is applied to stimulate production. Increasing demand for NR calls for an increase in latex yield, but also an extension of crops to marginal zones. Such marginal zones are subject to stronger and more detrimental abiotic constraints for latex production (positive cold, frost-prone, drought, *etc.*). Exposure to abiotic stress in addition to latex harvesting stress affects latex production and tree productivity [1]. Laticifer cells are the site of numerous types of stress that lead to oxidative stress. The plant's first response to wounding is to produce jasmonate [2]. The wound and associated jasmonate production also cause oxidative stress in the plant and activate antioxidant systems to overcome it [3]. Oxidative stress, whether produced in response to tapping and/or environmental constraints, may lie behind problems linked to latex flow [4]. It is reflected in an imbalance between ROS (reactive oxygen species) and the detoxification systems existing in laticifers. ROS then attack the membranes of cell organelles, causing the release of coagulant factors existing in the laticifers, which gives rise to *in situ* coagulation of rubber particles. This complex physiological disorder is called TPD (Tapping Panel Dryness). This phenomenon remains reversible up to a certain limit. It amounts to laticifer fatigue, called dry cut. In extreme cases, halted flow is followed by the degeneration and death of the laticifers and, consequently, severe cortical necrosis (Brown Bast) [5]. Annual losses due to TPD are estimated at 10% to 40% [1].

MicroRNAs (miRNAs) form a small class of non-coding RNAs of 21 to 24 nucleotides involved in numerous physiological processes in plants including response to biotic and abiotic stresses through post-transcriptional regulation. miRNA biogenesis requires several stages and several enzymes from *MIR* gene transcription to the generation of mature miRNA [6–10]. The mature miRNA enters the complex called RISC (RNA-induced silencing complex) thereby enabling regulation of the target gene. In the RISC complex, the mature miRNA associates with proteins of the Argonaute family (AGO) to direct the regulation of the target gene [11,12]. In the RISC complex, the mature miRNA sequence links up with the sequence of the messenger RNA through perfect or imperfect sequence homology. This action causes a silencing phenomenon in plants, called PTGS (post-transcriptional gene silencing) [13]. In plants, most target messenger RNAs contain a single site complementary to the miRNA, corresponding to the cleavage site [9], suggesting that plant microRNAs use RNA cleavage rather than preventing translation. Long microRNAs are taken in charge by Dicer-like 3 (DCL3), associate with AGO4 and then guide methylation of the DNA of target genes in rice and mosses [14,15].

Many microRNAs regulated by abiotic stress have been identified in the model species (for review, see [16]). The first involvements of microRNAs in response to stress were described by Jones Rhoades and Bartel who, in *Arabidopsis thaliana*, predicted genes targeted by microRNAs such as superoxide dismutase, the laccases and ATP sulfurylases (APS) [17]. MicroRNA expression was then studied in different contexts of abiotic stress in numerous species (for review, see [16]). Most microRNAs conserved between species target transcription factors, such as miR/target gene pairs *miR156/SBP*, *miR159/319/MYB-TCP*, *miR160/ARF*, *miR166/HDZIPIII*, *miR169/NFY* subunit.

MicroRNAs are also involved in the regulation of antioxidant activities and particularly in the regulation of reactive oxygen species-scavenging enzymes. The production of ROS is linked to abiotic

and biotic stress status in plants [18–20] and the excess ROS can lead to several types of cell damage, such as peroxidation of membrane lipid compounds, degradation of polysaccharides, denaturing of enzymes and lesions in DNA [21]. The miR398 targets the *CSD1* and *CSD2* genes encoding cytosolic and chloroplastic CuZnSOD in *Arabidopsis thaliana* [22]. The level of expression of these microRNAs and of their target is regulated in response to stress, involving a modulation of growth and development during stress. For example, miR398 is under-expressed in response to stress which leads to an accumulation of CuZnSOD enabling detoxification of an excess of reactive oxygen species in *Arabidopsis thaliana* [22].

To date, 28 sequences of *MIR* are accessible in miRBase [23–25] for the rubber tree. They represent 10 conserved families and 14 newly identified families potentially *Hevea*-specific miR [26,27], added to which four additional precursors have been identified recently [28]. Our previous studies on mature trees demonstrated the functionality of miRNA biogenesis in the latex cells. Moreover, we highlighted the reduction of small RNA size from 24 to 21 nucleotides in TPD-affected trees, which could not be explained by a general RNA degradation [28]. From this previous work, we hypothesized that the change in size could also be effective for mature miRNAs and consequently involved to independent miRNA pathway, one leading to targeted transcript cleavage and the other to methylation of targeted genes (for review [29]). Among all pre-miRNAs tested, only the relative *Hbpre-MIR159b* abundance was up-regulated between TPD-affected and healthy trees. Targets of miR159 were predicted and might be involved in various activities such as rubber biosynthesis, antioxidant activity and transcription regulation activity [26,28,30].

The objective of this paper was to study the regulation of microRNAs in response to different types of abiotic stress (cold, saline stress, wounding) or hormone treatments (ethylene and methyl jasmonate). Given some mature miRNAs could be produced by several *MIR* genes, this study focused on the expression analysis pattern using real-time RT-PCR with specific primers for each *MIR* genes. The relative accumulation of 19 premiRs was analyzed by real-time RT-PCR on various tissues from juvenile plant material in order to facilitate the application of treatments. Negative co-regulation of *Hbpre-miR398b* and its target *chloroplastic CuZnSOD* is revealed in response to saline stress.

2. Results and Discussion

2.1. Relative Accumulation of premiR in Response to Abiotic Stress in *in vitro* Plantlets

Relative accumulation of premiR of 20 selected *MIR* genes was analysed in different tissues (leaf, bark and root; Table 1). In control plants, the conserved *MIR* genes were generally very weakly expressed such as *HbMIR156*, *HbMIR319* and *HbMIR398a* and *HbMIR398b* (from 6.78×10^{-4} in roots for *HbMIR398a* to 9.22×10^{-2} for *HbMIR319* in bark). By contrast, the *HbMIR408a* and *HbMIR408b* genes were strongly expressed (from 8.47×10^{-1} to 3.54). Among the *MIR* genes newly identified in *Hevea brasiliensis*, *HbMIR6482*, *HbMIR6483* and *HbMIR6485* were also strongly expressed (from 2.20 for *HbMIR6482* to 7.44×10 for *HbMIR6485* in leaves).

In the same *in vitro* plantlets, statistical analyses revealed a difference in gene expression at tissue level (Table 1). Indeed, three genes *HbMIR166b*, *HbMIR319* and *HbMIR6482* had a relative differential expression level in leaves, bark and roots under untreated conditions. *HbMIR166b* and

HbMIR319 were more strongly expressed in bark while *HbMIR6482* was more strongly expressed in roots. For these three genes, the lowest expression level was found in leaves (Table 1). In addition, seven genes significantly displayed a relative differential expression level, namely *HbMIR159a*, *HbMIR398a*, *HbMIR398b*, *HbMIR408b*, *HbMIR6483*, *HbMIR6484* and *HbMIR6485*. The *MIR398c* transcript was only present in trace form under the majority of conditions.

Table 1. Relative premiR abundance of 13 conserved *MIR* genes and 7 *MIR* genes newly identified in *Hevea brasiliensis* in the leaves, bark and roots of one-year-old *in vitro* plantlets in response to cold and saline (NaCl) treatments compared with non-treated plants (C: control), nc: not calculated. Data with the same letter are not significantly different at the level of 5%.

| Gene | Tissue | Expression value for control | Cold/C | <i>p</i> -value | NaCl/C | <i>p</i> -value |
|---------------------|--------|------------------------------|--------|-----------------|--------|-----------------|
| <i>HbpremiR156</i> | leaf | 1.72×10^{-3} a | 1.79 | 0.52 | 0.17 | 0.08 |
| | bark | 1.80×10^{-3} a | 0.91 | 0.72 | 0.20 | 0.08 |
| | root | 2.63×10^{-3} a | 0.70 | 0.39 | 0.39 | 0.13 |
| <i>HbpremiR159a</i> | leaf | 4.68×10^{-2} b | 3.45 | 0.008 | 0.04 | 0.002 |
| | bark | 7.45×10^{-2} a | 1.25 | 0.22 | 0.17 | <0.001 |
| | root | 5.94×10^{-2} a,b | 0.32 | 0.02 | 0.10 | 0.01 |
| <i>HbpremiR159b</i> | leaf | 1.56×10^{-1} a | 4.06 | 0.004 | 0.05 | 0.003 |
| | bark | 2.21×10^{-1} a | 1.59 | 0.03 | 0.23 | <0.0001 |
| | root | 1.94×10^{-1} a | 0.44 | 0.23 | 0.11 | 0.02 |
| <i>HbpremiR166a</i> | leaf | 1.19×10^{-2} a | 1.38 | 0.38 | 2.08 | 0.08 |
| | bark | 1.78×10^{-2} a | 2.00 | 0.18 | 0.70 | 0.16 |
| | root | 4.16×10^{-2} a | 0.60 | 0.62 | 2.76 | 0.13 |
| <i>HbpremiR166b</i> | leaf | 2.21×10^{-1} c | 1.34 | 0.31 | 0.20 | 0.01 |
| | bark | 5.20×10^0 a | 0.64 | 0.01 | 0.14 | 0.003 |
| | root | 9.67×10^{-1} b | 0.48 | 0.39 | 1.01 | 0.80 |
| <i>HbpremiR319</i> | leaf | 5.73×10^{-3} c | 0.29 | 0.08 | 0.23 | 0.02 |
| | bark | 9.22×10^{-2} a | 1.30 | 0.90 | 0.07 | <0.0001 |
| | root | 1.76×10^{-2} b | nc | – | nc | – |
| <i>HbpremiR396</i> | leaf | 4.73×10^{-1} a | 1.34 | 0.28 | 0.10 | 0.002 |
| | bark | 1.41×10^0 a | 0.93 | 0.61 | 0.17 | <0.0001 |
| | root | 2.33×10^0 a | 0.28 | 0.19 | 0.12 | 0.05 |
| <i>HbpremiR398a</i> | leaf | 6.93×10^{-3} a | 0.34 | 0.12 | 0.21 | 0.02 |
| | bark | 2.27×10^{-3} a,b | 0.23 | 0.31 | 0.49 | 0.64 |
| | root | 6.78×10^{-4} b | 0.47 | 0.39 | 0.10 | 0.20 |
| <i>HbpremiR398b</i> | leaf | 3.27×10^{-2} b | 0.95 | 0.92 | 0.87 | 0.61 |
| | bark | 8.68×10^{-2} a | 0.52 | 0.15 | 0.07 | 0.01 |
| | root | 3.19×10^{-2} b | 0.62 | 0.34 | 0.15 | 0.03 |
| <i>HbpremiR398c</i> | leaf | trace | nc | – | nc | – |
| | bark | trace | nc | – | nc | – |
| | root | trace | nc | – | nc | – |
| <i>HbpremiR408a</i> | leaf | 8.47×10^{-1} a | 0.47 | 0.20 | 0.10 | <0.0001 |
| | bark | 1.50×10^0 a | 0.23 | 0.11 | 0.07 | <0.001 |
| | root | 1.48×10^0 a | 0.22 | 0.02 | 0.06 | 0.01 |

Table 1. Cont.

| Gene | Tissue | Expression value for control | Cold/C | p-value | NaCl/C | p-value |
|---------------------|--------|------------------------------------|--------|---------|--------|---------|
| <i>HbpremiR408b</i> | leaf | 1.60×10^0 ^b | 0.48 | 0.21 | 0.12 | <0.0001 |
| | bark | 3.54×10^0 ^a | 0.22 | 0.11 | 0.08 | 0.01 |
| | root | 2.75×10^0 ^{a,b} | 0.22 | 0.01 | 0.08 | 0.001 |
| <i>HbpremiR476</i> | leaf | 2.71×10^{-2} ^a | 0.66 | 0.39 | 0.16 | 0.01 |
| | bark | 2.84×10^{-2} ^a | 0.36 | 0.01 | 0.37 | 0.002 |
| | root | 8.49×10^{-2} ^a | 0.20 | 0.11 | 0.41 | 0.23 |
| <i>HbpremiR6482</i> | leaf | 2.20×10^0 ^c | 0.09 | 0.001 | 0.30 | 0.06 |
| | bark | 9.86×10^0 ^b | 0.21 | 0.002 | 0.45 | 0.01 |
| | root | 7.06×10^0 ^a | 0.05 | 0.004 | 0.05 | 0.003 |
| <i>HbpremiR6483</i> | leaf | 5.47×10^0 ^a | 0.59 | 0.50 | 0.03 | 0.002 |
| | bark | 1.56×10^0 ^a | 0.70 | 0.43 | 0.18 | 0.01 |
| | root | 5.54×10^{-1} ^b | 1.05 | 0.72 | 0.51 | 0.59 |
| <i>HbpremiR6484</i> | leaf | 7.17×10^{-1} ^a | 0.70 | 0.26 | 0.22 | 0.003 |
| | bark | 7.47×10^{-1} ^a | 0.65 | 0.06 | 0.55 | 0.02 |
| | root | 3.47×10^{-1} ^b | 0.25 | 0.02 | 0.35 | 0.05 |
| <i>HbpremiR6485</i> | leaf | 7.44×10^0 ^a | 0.73 | 0.49 | 0.03 | 0.000 |
| | bark | 3.14×10^0 ^a | 0.82 | 0.77 | 0.22 | 0.02 |
| | root | 6.06×10^0 ^b | 0.73 | 0.14 | 0.27 | 0.01 |
| <i>HbpremiRn11</i> | leaf | trace | nc | – | nc | – |
| | bark | trace | nc | – | nc | – |
| | root | trace | nc | – | nc | – |
| <i>HbpremiRn12</i> | leaf | 7.26×10^{-2} ^a | 1.25 | 0.50 | 6.45 | <0.0001 |
| | bark | 8.80×10^{-2} ^a | 0.75 | 0.29 | 5.93 | <0.001 |
| | root | 1.47×10^{-1} ^a | 0.48 | 0.10 | 4.17 | 0.01 |
| <i>HbpremiRn13</i> | leaf | 3.77×10^2 ^a | 1.17 | 0.60 | 0.34 | 0.11 |
| | bark | 3.41×10^2 ^a | 1.17 | 0.74 | 0.46 | 0.001 |
| | root | 2.65×10^2 ^a | 0.81 | 0.51 | 1.90 | 0.06 |

The regulation of *MIR* genes in response to cold and salinity stresses was shown as the ratio between the mean of three treated *in vitro* plantlets and three untreated *in vitro* plantlets (Table 1). An analysis of the relative expression of the *MIR* genes showed that 8 *MIR* genes were significantly regulated in response to cold (Table 1). In leaves, both *HbMIR159* genes displayed a significant increase in the relative number of transcripts in response to cold while the relative expression of *HbMIR6482* was greatly inhibited. In bark, the relative expression of 3 *MIR* genes (*HbMIR166b*, *HbMIR476* and *HbMIR6482*) was repressed while that of *HbMIR159b* was slightly stimulated. In roots, 5 *MIR* genes (*HbMIR159a*, *HbMIR408a*, *HbMIR408b*, *HbMIR6482* and *HbMIR6484*) displayed a significant reduction in transcripts in response to cold. The drop in temperature led to a drop in transcripts for *HbmiR6482* in the 3 tissues. *HbMIR159a* displayed an expression profile that was antagonistic between leaves and roots. With the exception of *HbMIR159a* and *HbMIR159b*, the significantly regulated *MIR* genes were repressed in response to cold stress.

The *in vitro* plantlets were also subjected to saline stress by watering with a 300 mM NaCl solution rather than water twice a day for 24 h. In response to NaCl, 16 of the 20 genes displayed a significant

repression in their level of relative transcript abundance. Only *HbmiRn12* was highly induced (4 to 6 times) in most tissues.

2.2. Relative Accumulation of preMIR in Response to Ethylene (ET), Methyl Jasmonate (MeJA) and Wounding in Three-Month-Old Epicormic Shoots

Based on previous studies on budded plant material [31], the effect of ET, MeJA and wounding was studied on three-month-old epicormic shoots. In control plants, the relative level of premiR was highly accumulated for *HbMIR6485* and *HbMIRn13* (1.60×10 and 3.92×10^2 respectively) and at low level for *HbMIR156*, *HbMIR319* and *HbMIR398a* (4.42×10^{-4} , 1.15×10^{-2} and 3.61×10^{-4} , respectively; Table 2). This result was similar to the basal expression profile found in the *in vitro* plantlets. 7 MIR genes had a relative expression level that differed significantly between bark and leaves (Table 2). They were *HbMIR159a*, *HbMIR159b*, *HbMIR166b*, *HbMIR319*, *HbMIR396*, *HbMIR398a* and *HbMIR6485*. In comparison, the aforementioned genes were also regulated differently depending on the tissues in the *in vitro* plantlets, apart from *HbMIR396*. Again, the *MIR398c* transcript was only present in trace form under the majority of conditions.

Table 2. Relative premiR abundance of 13 conserved MIR genes and 7 MIR genes newly identified in *Hevea brasiliensis* in the leaves and bark collected on 3-month-old epicormic shoot of budded plants in response to ethylene (ET), methyl jasmonate (MeJA) and wounding (W) compared with non-treated plants (C: control), nc: not calculated. Data with the same letter are not significantly different at the level of 5%.

| Gene | Tissue | Expression value for control | ET/C | p-value | MeJA/C | p-value | W/C | p-value |
|---------------------|--------|------------------------------|-------|---------|--------|---------|------|---------|
| <i>HbpremiR156</i> | leaves | 4.42×10^{-4} a | 1.02 | 0.94 | 6.23 | 0.06 | 2.92 | 0.68 |
| | bark | 6.80×10^{-4} a | 1.63 | 0.26 | 0.66 | 0.62 | 1.68 | 0.35 |
| <i>HbpremiR159a</i> | leaves | 2.34×10^{-1} a | 1.22 | 0.91 | 1.40 | 0.04 | 1.04 | 0.88 |
| | bark | 4.41×10^{-2} b | 0.40 | 0.30 | 0.75 | 0.99 | 0.33 | 0.20 |
| <i>HbpremiR159b</i> | leaves | 6.28×10^{-1} a | 0.81 | 0.38 | 1.64 | 0.02 | 0.76 | 0.14 |
| | bark | 1.96×10^{-1} b | 0.50 | 0.35 | 0.86 | 0.92 | 0.41 | 0.22 |
| <i>HbpremiR166a</i> | leaves | 4.02×10^{-2} a | 0.49 | 0.39 | 2.84 | 0.05 | 0.98 | 0.84 |
| | bark | 1.62×10^{-2} a | 0.35 | 0.17 | 1.00 | 0.99 | 1.06 | 0.72 |
| <i>HbpremiR166b</i> | leaves | 5.98×10^0 a | 1.50 | 0.24 | 2.45 | 0.22 | 1.78 | 0.07 |
| | bark | 1.00×10^{-1} b | 1.08 | 0.67 | 1.81 | 0.15 | 0.32 | 0.03 |
| <i>HbpremiR319</i> | leaves | 2.80×10^{-1} a | 0.73 | 0.63 | 1.46 | 0.33 | 0.87 | 0.83 |
| | bark | 1.15×10^{-2} b | 0.31 | 0.03 | 1.28 | 0.41 | 0.18 | 0.02 |
| <i>HbpremiR396</i> | leaves | 2.87×10^0 a | 0.62 | 0.20 | 3.19 | 0.09 | 0.57 | 0.22 |
| | bark | 4.80×10^{-1} b | 0.27 | 0.02 | 0.50 | 0.13 | 0.43 | 0.07 |
| <i>HbpremiR398a</i> | leaves | 3.61×10^{-4} b | 0.22 | 0.05 | 0.69 | 0.53 | 1.85 | 0.34 |
| | bark | 1.66×10^{-2} a | 3.59 | 0.08 | 3.38 | 0.21 | 2.99 | 0.19 |
| <i>HbpremiR398b</i> | leaves | 1.89×10^{-3} b | 0.52 | 0.67 | 0.09 | 0.41 | 0.05 | 0.27 |
| | bark | 5.01×10^{-2} a,b | 15.19 | 0.09 | 0.61 | 0.46 | 4.66 | 0.60 |
| <i>HbpremiR398c</i> | leaves | trace | 0.91 | 0.82 | nc | – | 0.04 | 0.16 |
| | bark | trace | nc | – | nc | – | nc | – |

Table 2. Cont.

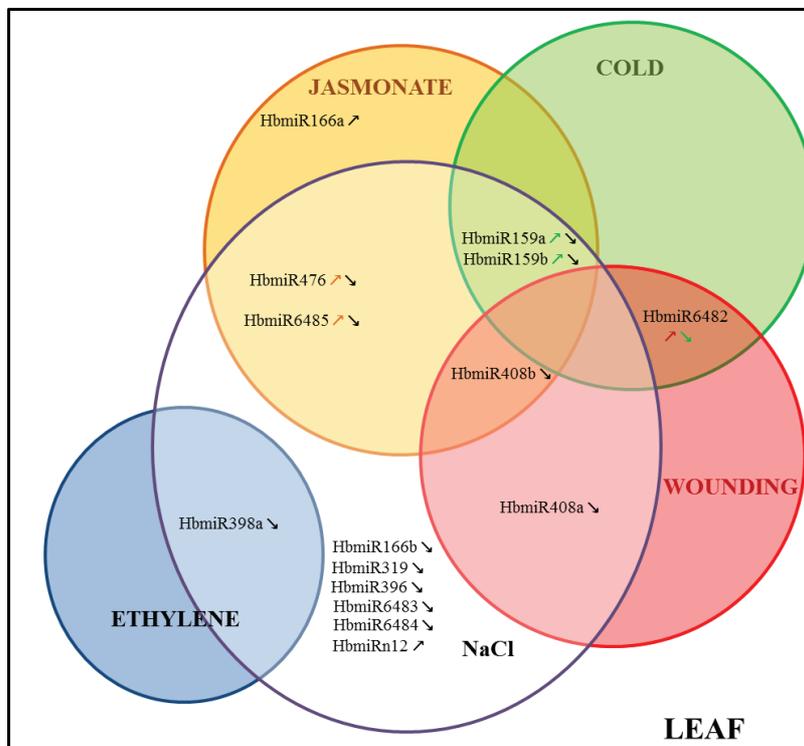
| Gene | Tissue | Expression value for control | ET/C | <i>p</i> -value | MeJA/C | <i>p</i> -value | W/C | <i>p</i> -value |
|---------------------|--------|------------------------------------|------|-----------------|--------|-----------------|------|-----------------|
| <i>HbpremiR408a</i> | leaves | 3.49×10^0 ^a | 0.40 | 0.61 | 0.05 | 0.21 | 0.04 | 0.002 |
| | bark | 1.29×10^0 ^a | 0.50 | 0.64 | 0.002 | 0.52 | 0.05 | 0.62 |
| <i>HbpremiR408b</i> | leaves | 5.40×10^0 ^a | 0.42 | 0.55 | 0.02 | 0.01 | 0.04 | 0.001 |
| | bark | 2.52×10^0 ^a | 0.73 | 0.30 | 0.003 | 0.30 | 0.05 | 0.91 |
| <i>HbpremiR476</i> | leaves | 1.49×10^{-1} ^a | 0.84 | 0.52 | 17.93 | 0.001 | 0.98 | 0.73 |
| | bark | 1.73×10^{-2} ^a | 0.26 | 0.10 | 6.18 | 0.06 | 0.32 | 0.30 |
| <i>HbpremiR6482</i> | leaves | 4.41×10^0 ^a | 1.14 | 0.60 | 1.60 | 0.26 | 7.19 | 0.05 |
| | bark | 1.63×10^0 ^a | 0.62 | 0.48 | 1.32 | 0.53 | 2.68 | 0.14 |
| <i>HbpremiR6483</i> | leaves | 5.56×10^0 ^a | 2.97 | 0.18 | 4.14 | 0.13 | 0.81 | 0.85 |
| | bark | 8.63×10^0 ^a | 1.13 | 0.67 | 1.34 | 0.69 | 0.68 | 0.53 |
| <i>HbpremiR6484</i> | leaves | 6.10×10^{-1} ^a | 0.92 | 0.65 | 1.67 | 0.14 | 0.48 | 0.12 |
| | bark | 8.92×10^{-1} ^a | 0.99 | 1.00 | 1.15 | 0.72 | 0.73 | 0.41 |
| <i>HbpremiR6485</i> | leaves | 1.60×10^b | 1.19 | 0.58 | 2.62 | 0.004 | 0.61 | 0.11 |
| | bark | 3.00×10^a | 1.00 | 0.94 | 0.82 | 0.44 | 0.52 | 0.06 |
| <i>HbpremiRn11</i> | leaves | trace | nc | – | 3.52 | 0.08 | 1.35 | 0.98 |
| | bark | trace | 0.97 | 0.93 | 0.56 | 0.24 | 0.60 | 0.39 |
| <i>HbpremiRn12</i> | leaves | 6.19×10^{-2} ^a | 1.23 | 0.65 | 1.51 | 0.12 | 7.12 | 0.13 |
| | bark | 6.13×10^{-2} ^a | 1.20 | 0.58 | 0.84 | 0.96 | 1.42 | 0.44 |
| <i>HbpremiRn13</i> | leaves | 3.16×10^2 ^a | 1.23 | 0.39 | 1.24 | 0.55 | 1.41 | 0.96 |
| | bark | 3.92×10^2 ^a | 1.85 | 0.16 | 0.91 | 0.55 | 1.85 | 0.63 |

The regulation of *MIR* genes in response to ethylene, MeJA and wounding was shown as the ratio between the mean of three treated *in vitro* plantlets and three untreated *in vitro* plantlets. 3 *MIR* genes were significantly regulated by the ethylene treatment (Table 2). In leaves, only the relative expression of the *HbMIR398a* gene was repressed. In bark, 2 *MIR* genes were negatively regulated by ethylene, and the expression of *HbMIR319* and *HbMIR396* was inhibited (Table 2). Application of methyl jasmonate induced regulation of the relative expression of 6 *MIR* genes in leaves and not in bark (Table 2). The *HbMIR159a*, *HbMIR159b*, *HbMIR166a*, *HbMIR476* and *HbMIR6485* genes displayed a relative increase in transcripts, while *HbMIR408b* displayed a significant large drop in transcripts in response to external methyl jasmonate (Table 2). In response to wounding, 5 genes were significantly regulated (Table 2). In leaves, *HbMIR408a* and *HbMIR408b* displayed a large reduction in their relative expression levels while *HbMIR6482* significantly showed an increase in relative abundance for these transcripts (Table 2). In bark, the *HbMIR166b* and *HbMIR319* genes displayed a significant drop in their relative transcript abundance (Table 2). It should be noted that the *HbMIR319* gene was regulated in the same way in response to both ethylene and wounding.

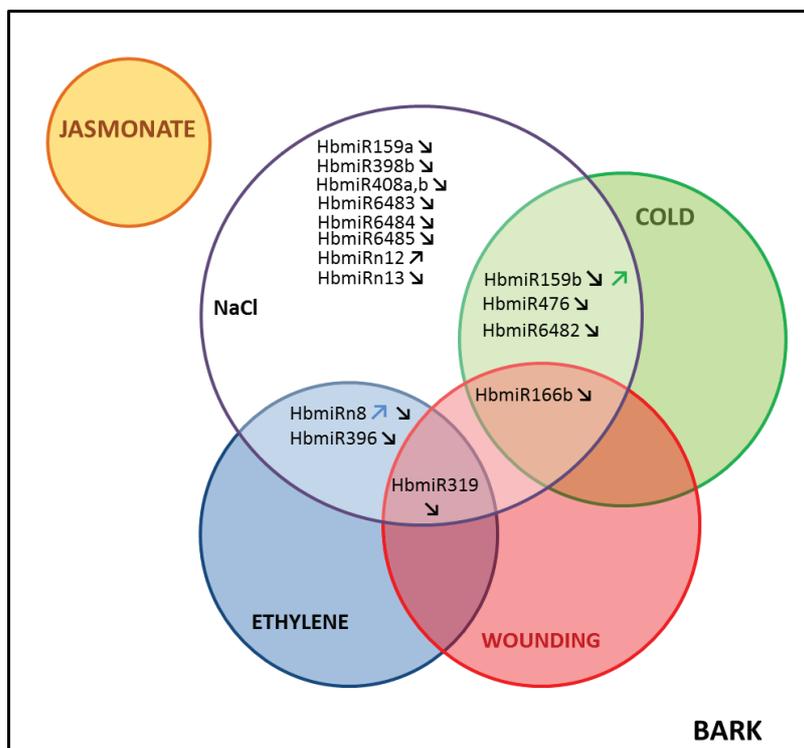
The response to jasmonate in leaves seemed to differ from the other abiotic stresses. In fact, the expression of 5 *MIR* genes (*MIR159a*, *MIR159b*, *MIR166*, *MIR476* and *MIR6485*) was stimulated during the treatment. MiR159 and miR476 are known to target type MYB transcription factors [17,32] and PPRs (pentatricopeptide repeat proteins) [33], respectively. These targets were not identified during the prediction of targets of miR159 and miR476 in *Hevea*. In contrast, miR166 is known to target proteins of the HD-ZIPIII family in *Arabidopsis* and also in *Hevea* [17,26].

The expression of *MIR* genes was differentially regulated depending on the stress and the tissue (bark and leaf), suggesting specific response pathways for each type of stress (Figure 1a,b).

Figure 1. Venn diagram representing the regulation of the *MIR* genes common to all the treatments tested. (a): in leaves; and (b): in bark.



(a)



(b)

In maize, conversely, the comparison of expression pattern of mature miRNAs in plant subjected to saline stress and a viral infection lead the authors to suggest some cross-talk between abiotic and biotic stress [34]. In addition to differential expression depending on the tissue and stress applied, some differences between species are worth noting. In *Hevea brasiliensis*, the response to cold stress differs from that found in *Arabidopsis thaliana* [35]. In *Arabidopsis*, miR166b is accumulated in response to cold stress whereas in *Hevea*, repression of the *HbMIR166b* gene is found. Inhibition of the relative expression of the *HbMIR476* gene in *Hevea* bark shows an expression profile similar to that in poplar [36]. For saline stress, regulation of the expression of microRNAs appears to be specific to each species. In *Hevea*, saline stress led to global repression of the expression of all the *MIR* genes studied, except *HbMIRn12*. In this case, the result suggests that response mechanisms were brought into play by blocking the repression of genes targeted by microRNAs. By contrast, in *Arabidopsis*, accumulation of the mature microRNAs miR156, miR159, miR319 and miR396 and a reduction in the quantity of miR398 are seen [37]. The quantity of miR396 and miR156 is reduced in rice and maize respectively [38]. In poplar, the quantity of miR530a, miR1445, miR1446a-e, miR1447 and miR1711 is reduced while that of miR482.2 and miR1450 is increased [36].

The difference observed between *Hevea* and other plants could be explained by post-transcriptional regulations acting between pri-*MIR* transcript and the release of mature miRNA [39]. Moreover, most of studies reported on total mature miRNAs using Northern-blot hybridization highlighting the final products of several *MIR* genes. In our study, the use of real-time RT-PCR allowed identifying the specific expression of each *MIR* gene isoform. The microRNAs newly identified in *Hevea brasiliensis* have a high level of expression compared to that of conserved microRNAs under normal conditions. Different elements tend to show that the new microRNAs (young miRNAs) have little or no function [40]. Although conserved microRNAs are generally more abundant than the new microRNAs [41], an abundance of the latter does not mean greater regulation of the targets. Regulation of the targets of new microRNAs is not affected in mutant transgenic plants in the miRNA biogenesis pathway. The new microRNAs would seem not to be integrated yet in the regulation networks or may only function in a precise spatio-temporal context, or during response to a specific type of biotic or abiotic stress [40].

2.3. Analysis of the Co-Regulated Expression of the HbMIR398a, HbMIR398b and HbMIR398c Genes and Their Putative Target Gene Chloroplastic HbCuZnSOD in Response to Abiotic Stress and to Hormone Treatments

Three miR/target pairs have been experimentally validated in the rubber tree (miR156/Squamosa promoter binding protein, miR160/ARF and miR398/chloroplastic CuZnSOD) [26]. We chose to monitor expression of the *MIR398/chloroplastic CuZnSOD* pair. HbmiR398 is generated by three *HbMIR398* genes (3 genes, a, b and c; Figure 2).

In order to monitor expression of the *chloroplastic HbCuZnSOD* target, two primer pairs were designed, one flanking the cleavage site and the other at 3' UnTranslated Region (3'UTR) (Table 3).

Figure 2. Stem-loop structure of three precursors of *MIR398* in *Hevea brasiliensis*. The mature microRNA sequence is shown in dark.

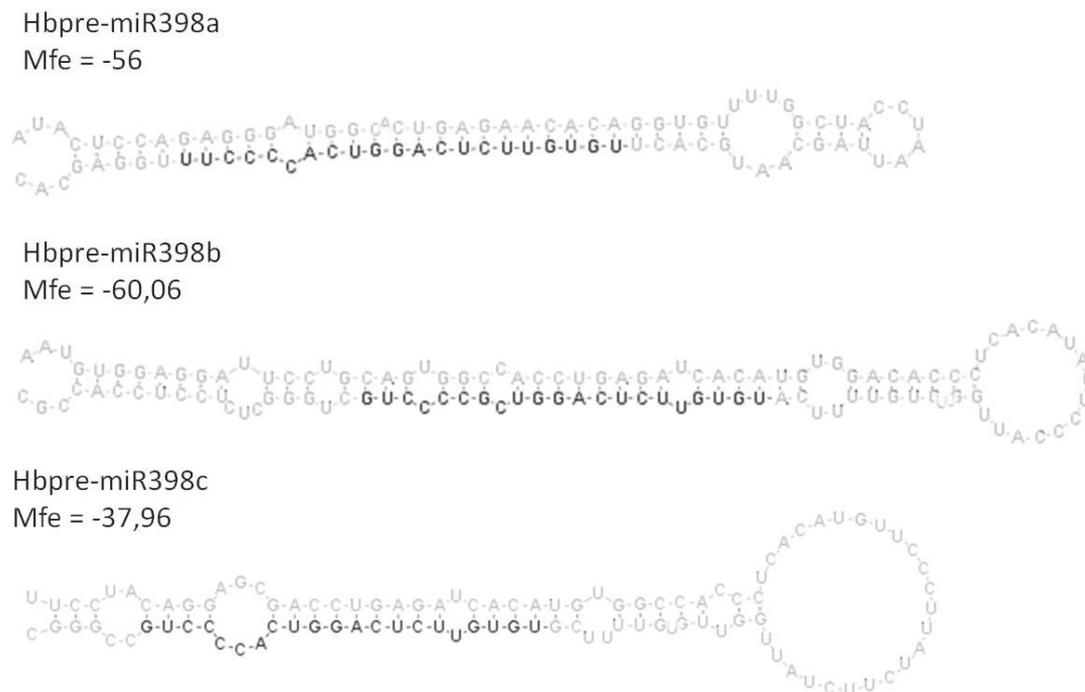


Table 3. List of primers and their sequences used for *MIR* gene expression analysis by RT-PCR. All primers are presented from 5' to 3' end.

| Pre-microRNA | Forward primer | Reverse primer | PCR efficiency |
|----------------------|------------------------------------|--------------------------------|----------------|
| <i>Hbpre-miR156</i> | TGGTGATGTTGTTGACAGAAG ATAGAGAGC | GCACAAAGGAGTGAGATGCA GAGTCC | 1.79 |
| <i>Hbpre-miR159a</i> | GGTTAAGAAGTGGAGCTCCTT GAAGTC | GCTCCCTTCAATCCAAACAAG GATC | 1.958 |
| <i>Hbpre-miR159b</i> | GTGGAGCTCCTTGAAGTCCAA TAGAGG | AGAGCTCCCTTCAATCCAAAC AAGG | 1.881 |
| <i>Hbpre-miR166a</i> | TTCTTTTTGAGGGGAATGTTG TCTGG | GGAATGAAGCCTGGTCCGAG GAG | 1.820 |
| <i>Hbpre-miR166b</i> | GGGGAATGTTGTCTGGTTCGA TG | TCAAATCAAACCCTGTTGGGG G | 1.738 |
| <i>Hbpre-miR319</i> | CCAGTCACGGTGGGCAATGG G | GGAGCTCCCTTCAGTCCAAGT ACAGG | 1.847 |
| <i>Hbpre-miR396</i> | TGACCCTCTTCGTATTCTTCCA CAGC | CCCACAGCTTTATTGAACCGC AAC | 1.782 |
| <i>Hbpre-mir398a</i> | TGAGAACACAGGTGTTTTGGC TACC | GTGCTCAAAGGGGTGACCTG AG | 1.879 |
| <i>Hbpre-mir398b</i> | ACCTGAGATCACATGTGGACA CCC | GCGGTGGAGGAGAGCCCAG | 1.939 |
| <i>Hbpre-mir398c</i> | TGGCCACCCTCACATGTTCCC | CCGGCAGGGGTGACCTGAG | 1.965 |
| <i>Hbpre-miR408a</i> | ACTGGGAACAGGCAGAGCAT GG | GCCACAAGCCAGGGAAGAGG C | 1.723 |

Table 3. Cont.

| Pre-microRNA | Forward primer | Reverse primer | PCR efficiency |
|---|-----------------------------------|------------------------------------|----------------|
| <i>Hbpre-miR408b</i> | GACATACAAAGACTGGGAAC AGGCAG | GCCACAAGCCAGGGGAAGAG GC | 1.792 |
| <i>Hbpre-miR476</i> | GCCTTGATGTTTCATTTAGTA ATCCTTCT | GATAATCCTTCTATGCAAAGT CTTTTATGC | 1.732 |
| <i>Hbpre-miR6482</i> | ACCAGGAACTGGTATCAACCC AGC | TGCTACCAATGAATCGGACCC ACC | 1.837 |
| <i>Hbpre-miR6483</i> | CAGTAAATAGCAGTATCGTGG ATAGGG | GTCCAATCATTGATCCTGAAA ATTTCTAC | 1.828 |
| <i>Hbpre-miR6484</i> | TGGATTGGAGCCCAATACTGT GAC | CTGCTCCATTGATTTTACCATC TATGC | 1.873 |
| <i>Hbpre-miR6485</i> | ACCTAGGATGTAGAAGAGCA TAAC | ACTACATGAGTGGATATATAG GAATCC | 1.787 |
| <i>Hbpre-miRn11</i> | GTATCAACGCAGATGTGCCGC C | CCCCAGCCAAACTCCCCACC | 1.828 |
| <i>Hbpre-miRn12</i> | AGCTTTCACCCAATAACCTTT GCAGT | GCTCTTCCAATTCTATCCAA AGTGGT | 1.78 |
| <i>Hbpre-miRn13</i> | TGTGTTGGCCTTCGGGATCGG | CGAATGCCCCCGACTGTCCC | 1.889 |
| <i>Hb-RH2b</i> | GAGGTGGATTGGCTAACTGAG AAG | GTTGAACATCAAGTCCCCGAG C | 1.68 |
| <i>HbCuZnSOD</i> (flanking miRNA site) | GCTCTATCTCTCGCCGCCGCC TCC | CCGCAATTGTTGCTTCTGCC | 1.785 |
| <i>HbCuZnSOD (3'UTR)</i> | TGGCAGAAGCAACAATTGCGG | GCAGGGAACAATGGCTGCC | 2 |

In response to abiotic stress, the transcripts of the *chloroplastic HbCuZnSOD* were accumulated significantly in roots in response to cold, in bark and roots in response to saline stress, and particularly after 24 h of treatment revealing that these transcripts are not cleaved by miR398 (Table 4, Tables S1 and S2). When the relative abundance of *chloroplastic HbCuZnSOD* transcripts was visualized (cleaved and non-cleaved with the primers at 3'UTR), the relative expression level for *chloroplastic HbCuZnSOD* transcripts decreased in leaves after 24 h of saline stress. However, that relative expression level rose significantly in response to saline stress in roots after 24 h of saline stress, and in wounded leaves (Table 4).

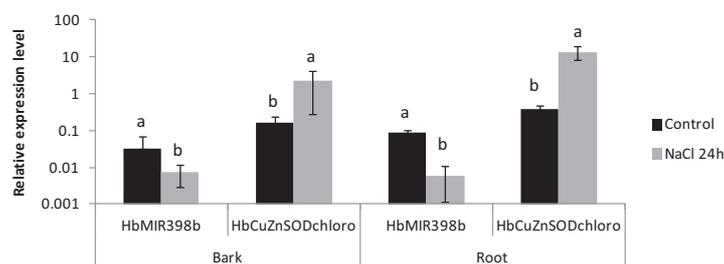
The relative accumulation of premiRs was analyzed for the three *MIR398* genes. The *MIR398c* transcript was only present in trace form under the majority of conditions. The relative expression of the *MIR398a* and *MIR398b* genes was significantly repressed in response to stress. In fact, in leaves, the relative expression of the *MIR398a* gene was repressed in response to 24 h of saline and ethylene treatment. Likewise, the relative expression of the *MIR398b* gene was repressed in bark and roots after 24 h of saline stress treatment.

When the relative expression level for the *HbMIR398* gene was compared with the expression data for the target, the significant reduction in the expression of the *HbMIR398b* gene in bark and roots in response to 24 h of saline stress was accompanied by a significant increase in the relative expression level for the non-cleaved *chloroplastic HbCuZnSOD* transcripts (Figure 3).

Table 4. Expression profiles of the *HbMIR398a*, *HbMIR398b* and *HbMIR398c* genes and their *chloroplastic HbCuZnSOD* target by real time-PCR in leaves, bark and roots of one-year-old *Hevea brasiliensis* *in vitro* plantlets in response to a positive cold, NaCl, ethylene, methyl jasmonate (MeJA) and wounding treatments. The statistically significant ratios (p -value < 0.05; Tables S1 and S2) with a value over 1 (**in red**) and under 1 (**in green**), represent over-expression and under-expression, respectively, in response to the treatments. The non-significantly regulated genes are shown **in yellow** (nc: not calculated).

| Treatment | Tissue | Ratio of relative transcript accumulation (Treated/Control) | | | | |
|-----------|--------|---|---------------------------|-------------------|-------------------|-------------------|
| | | chloro CuZnSOD (flanking miRNA site) | chloro CuZnSOD (3'UTR) | <i>preMIR398a</i> | <i>preMIR398b</i> | <i>preMIR398c</i> |
| Cold | leaf | 2.71 | 3.76 | 0.34 | 0.95 | nc |
| | bark | 8.98 | 8.90 | 0.23 | 0.52 | nc |
| | root | 3.41 | 1.46 | 0.47 | 0.62 | nc |
| NaCl | leaf | 0.84 | 0.29 | 0.21 | 0.87 | nc |
| | bark | 13.99 | 5.21 | 0.49 | 0.07 | nc |
| | root | 37.45 | 17.83 | 0.1 | 0.15 | nc |
| Ethylene | leaf | nc | 0.42 | 0.22 | 0.52 | 0.91 |
| | bark | 0.49 | 0.32 | 3.59 | 15.19 | nc |
| MeJA | leaf | 1.47 | nc | 0.69 | 0.09 | nc |
| | bark | 0.79 | 1.00 | 3.38 | 0.61 | nc |
| Wounding | leaf | 2.61 | 7.07 | 1.85 | 0.05 | 0.04 |
| | bark | 0.85 | 0.88 | 2.99 | 4.66 | nc |

Figure 3. Graph showing the expression values for the *chloroplastic HbCuZnSOD* and *HbMIR398b* genes in response to saline stress in bark and roots after 24 h of treatment. The expression values represent the mean and standard deviation of three treated or untreated plants, for each tissue. Data with the same letter are not significantly different at the level of 5%.



Co-expression of *MIR398* and of the gene encoding chloroplastic CuZnSOD was analyzed in response to several treatments. This enzyme takes part in ROS detoxification through the transformation of two superoxide ions H_2O_2 and H_2O . In *Arabidopsis*, miR398 targets cytosolic and chloroplastic superoxide dismutase (CSD1 and CSD2, [22]). Under oxidative stress conditions (heavy metal, strong sunlight and methyl viologen treatments), negative regulation of CSD1 and CSD2 by miR398 is halted in leaves, leading to an accumulation of enzymes and a drop in toxic free radicals for the cell. Tolerance of oxidative stress is consequently increased [22]. Interestingly, the *chloroplastic CuZnSOD* is the only one to have been validated in *Hevea* for miR398, the cytosolic form was not predicted to be

targeted by miR398 and the absence of cleavage was experimentally validated [26]. Some other targets have been predicted for this microRNA but have yet to be experimentally validated [26]. Co-regulation was only visible in response to saline stress. A reduction in the expression of the *HbMIR398b* gene was associated with an accumulation of *chloroplastic HbCuZnSOD* transcripts in bark and roots but not in leaves. It is noteworthy to specify that the bark of the young *Hevea in vitro* plantlets was also chlorophyllous. For root tissues, other plastid organelles might be present. In addition, the level of expression for the *MIR398* genes was not always negatively correlated with the level of *chloroplastic HbCuZnSOD* transcripts, which has already been found in other species and suggests the existence of two regulation mechanisms, dependent on and independent of miR398 [42].

3. Experimental Section

3.1. Plant Material and Treatments

In vitro plantlets were regenerated by indirect secondary somatic embryogenesis from maternal inner integument of immature fruits from rubber clone PB 260 [43]. These *in vitro* plantlets were acclimatized and reared in the greenhouse for one year at a temperature of 28 °C. At that stage, the treatments were applied on three biological replicates. For cold stress, the plants were placed at 4 °C for 12 h. Saline stress was applied to the *in vitro* plantlets by watering with a 300 mM solution of NaCl rather than just water twice a day for the duration of the treatment (24 h).

Three-month-old epicormic shoot from grafted plants were used for ethylene (ET), methyl-jasmonate (MeJA) and wounding treatments. Plants from the *Hevea* clone PB 260 were grafted onto rootstock GT1. This plant material included a mature growth unit corresponding to a three-month-old epicormic shoot after cutting back. The treatment was applied for 4 h on the whole plant at 8:00 a.m. For application of the ethylene and methyl jasmonate treatments, each plant was placed in a hermetically sealed plexiglas box with a volume of 300 L leaving the doors open for 24 h prior to treatment to limit the effects of transport or any wounds. For application of the ethylene treatment, the gas was injected to obtain an ethylene concentration of 1 ppm (300 µL of pure ethylene) in the box. For the methyl jasmonate treatment, 20 µL of methyl jasmonate solution (>95%) (Sigma, St. Louis, MO, USA) was diluted in 500 µL of absolute ethanol and placed on Whatman paper inside the box, in order to release the methyl jasmonate in gas form. For these two stress conditions, samples were taken 4 h after treatment. The plants used as controls for these treatments were also placed in a hermetically sealed box without gas injection. In that way, it was possible to eliminate the impact of the box when comparing the two treatments (control and treated). For the wounding treatment, the plants were wounded on their leaves by applying pressure every 2 cm with serrated tweezers and on the bark by making cuts with a scalpel 1 cm apart on the stem of the scion.

Table 1 summarizes the different treatments applied to the rubber *in vitro* and grafted plants. Some leaf, bark and root samples were taken after 12 h for the *in vitro* plantlets, and only leaves and bark for the budded plants whose root system was not clonal. All the samples were immersed directly in liquid nitrogen after being taken and were stored at -80 °C pending RNA extraction.

3.2. Extraction and Purification of Total RNAs

The samples (leaf, bark and root) were ground in liquid nitrogen. The resulting fine powder was placed on ice in an extraction buffer (25 mM guanidium isothiocyanate 5M NaAc, 0.88% of sarcosine, 0.9% of polyvinylpyrrolidone (PVP) and 1% of β -mercapto-ethanol). The samples were mixed for 30 s to optimize membrane lysis. An initial centrifugation (30 min at 4 °C at 10,000 g) was applied to precipitate cell debris. The supernatant was deposited on a caesium cushion (24 mM NaAc, 45.6 mM CsCl). The tubes were then balanced and centrifuged for 20 h at 20 °C at 25,000 rpm in a Beckmann Coulter ultracentrifuge (L7, Beckmann Coulter, Brea, CA, USA) using the SW41 rotor (Beckmann Coulter). Only dense RNAs were able to cross the caesium chloride cushion. After 20 h, the supernatant was removed by pipetting, leaving 1 cm height of solution, which was eliminated by tipping up the tube. This precaution was taken to avoid contamination by genomic DNA. The pellet was dried for 5 min then rinsed in 70% ethanol then taken up in 200 μ L of sterile water. The RNAs obtained were quantified then checked by electrophoresis (Mini-Sub Cell, Biorad, Richmond, CA, USA) on agarose gel (1% w/v, 1 \times Tris Acetate Ethylenediaminetetraacetic acid) and stained with ethidium bromide for visualization under ultraviolet. The RNAs were then aliquoted and stored at -80 °C.

3.3. Checking for the Presence of Genomic DNA and DNase Treatment

A polymerase chain reaction (PCR) was carried out to check for the presence of genomic DNA in the samples of extracted total RNA. Amplification was carried out in a final volume of 50 μ L. This was done with a volume of RNA corresponding to 50–100 ng to which was added a reaction mixture comprising 10 \times reaction buffer (10 mM Tris, pH = 8.3), MgCl₂ (1.5 mM), dNTP (0.04 mM), gene primers encoding actin (Forward: (5'-TCCATAATGAAGTGTGATGT-3', Reverse: (5'-GGACCTGACTCGTCATACTC-3', chosen astride an intron) (0.04 μ M) and Taq polymerase (1 U/ μ L). A sample of *Hevea* genomic DNA was used as the positive control. The samples were incubated for 1 min and 30 s at 94 °C followed by 40 cycles, each composed of 30 s of denaturing at 94 °C, 30 s of hybridization at 58 °C and 90 s of elongation at 72 °C, with the reaction ending on a final elongation of 10 min at 72 °C. The amplification products were deposited on an agarose gel (1% w/v, 1 \times Tris Acetate Ethylenediaminetetraacetic acid) then stained with ethidium bromide. The RNA samples that proved to be contaminated by genomic DNA were treated by digestion with DNase (Turbo DNA-FREE kit from Ambion, Life technologies, Carlsbad, CA, USA). The reaction took place from 8 ng of total RNAs in a final volume of 40 μ L composed of DNase buffer (10 \times , 4 μ L, Ambion), DTT (0.1 M, 2 μ L, Invitrogen, Life technologies, Carlsbad, CA, USA), RNase inhibitor (RiboLock, Fermentas, Burlington, ON, Canada, 40 U/ μ L) and TURBO DNase (2 U/ μ L). The samples were then incubated for 15 min at 37 °C. Next, a DNase inactivator (TURBO DNase Inactivation Reagent, Ambion) was added to the RNA samples. Enzyme inactivation took place for 2 min at ambient temperature by mixing the tubes. The supernatant containing the purified RNAs was recovered by centrifugation (12,000 g, 2 min, 4 °C).

3.4. cDNA Synthesis and Real-Time PCR

A microgram of total RNAs was used to synthesize some cDNAs in a reaction volume of 20 μ L using RevertAid™ M-MuLV reverse transcriptase (MBI, Fermentas, Burlington, Canada) following the supplier's recommendations. The primers used to amplify the stem-loop structures are listed in Table 2. The PCR conditions comprised a denaturing cycle at 95 °C for 2 min, then 45 cycles comprising denaturing at 95 °C for 20 s and a hybridization and elongation stage at 60 °C for 20 s. The PCR reaction took place in a volume of 6 μ L containing 2 μ L cDNA diluted 25 times, 1 μ L of primers at 5 μ M, and 3 μ L of 2 \times SYBR green mixture (LightCycler® 480 SYBR Green I Master, Roche Applied Sciences, Bâle, Switzerland). All the primers were validated by generating a standard curve from the dilution in series of 8 points, in triplicate, to calculate the efficiency of each primer pair (Table 3). In addition, amplification specificity was checked by generating the melting curves that needed to display a unique peak, and all PCR products were checked by sequencing. Relative quantification was carried out with Light Cycler 480 software (version 1.5.0, Roche Applied Sciences, Bâle, Switzerland), using *RH2b* [44] as the reference.

3.5. Statistical Analysis

The statistical analyses (analysis of variance followed by a Fisher comparison of means test) were done with XLSTAT (version 2011.4.02, Addinsoft, Paris, France) using expression data normalized by the LOG10(X) function. It was considered as up-regulation when the ratio was >1.0 and down-regulation when the ratio was <1.0. The p value corresponds to the Fisher test of the ANOVA.

4. Conclusions

Studying the regulation of *MIR* genes under harvesting stress (tapping, ethephon stimulation) and ROS-induced TPD in mature rubber trees remains difficult. However, some recent studies revealed that only one *MIR* genes is differentially regulated upon TPD. The *MIR159b* gene was shown up-regulated upon TPD occurrence [24]. The expression of this gene was increased in response to cold in leaves and bark, and in response to the jasmonic acid treatment in leaves of juvenile plantlets. Conversely, it was inhibited in all the tissues in response to salinity.

In order to get a full understanding of mechanisms involved in latex production and TPD syndrome, a complete validation of miRNA/target messenger pairs is first needed using by high throughput “degradome” analysis [45]. Combination of analyses on juvenile and mature plant materials will help developing model of *MIR* gene regulations under abiotic stress and further characterization of the TPD-regulated miRNAs and their targets.

Acknowledgments

This work was supported by the Institut Français du Caoutchouc. The authors thank Peter Biggins for his revision of the English, and Maryannick Rio, Florence Dessailly, Florence Martin and Eve Lorenzini for technical help.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Chen, S.; Peng, S.; Huang, G.; Wu, K.; Fu, X.; Chen, Z. Association of decreased expression of a Myb transcription factor with the TPD (tapping panel dryness) syndrome in *Hevea brasiliensis*. *Plant Mol. Biol.* **2003**, *51*, 51–58.
2. Strassner, J.; Schaller, F.; Frick, U.B.; Howe, G.A.; Weiler, E.W.; Amrhein, N.; Macheroux, P.; Schaller, A. Characterization and cDNA-microarray expression analysis of 12-oxophytodienoate reductases reveals differential roles for octadecanoid biosynthesis in the local *versus* the systemic wound response. *Plant J.* **2002**, *32*, 585–601.
3. Orozco-Cardenas, M.L.; Narvaez-Vasquez, J.; Ryan, C.A. Hydrogen peroxide acts as a second messenger for the induction of defense genes in tomato plants in response to wounding, systemin, and methyl jasmonate. *Plant Cell* **2001**, *13*, 179–191.
4. Chrestin, H. La stimulation à l'éthrel de l'hévéa jusqu'où ne pas aller trop loin. *Caoutchouc et Plastiques* janvier- février **1985**, *1985*, 75–78.
5. Faÿ, D. Histo- and cytopathology of trunk phloem necrosis, a form of rubber tree (*Hevea brasiliensis* Müll. Arg.) tapping panel dryness. *Aust. J. Bot.* **2011**, *59*, 563–574.
6. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297.
7. Chen, X. MicroRNA biogenesis and function in plants. *FEBS Lett* **2005**, *579*, 5923–5931.
8. Bonnet, E.; van de Peer, Y.; Rouze, P. The small RNA world of plants. *New Phytol.* **2006**, *171*, 451–468.
9. Zhang, B.; Pan, X.; Cobb, G.P.; Anderson, T.A. Plant microRNA: A small regulatory molecule with big impact. *Dev. Biol.* **2006**, *289*, 3–16.
10. Naqvi, A.; Sarwat, M.; Hasan, S.; Choudhury, N. Biogenesis, functions and fate of plant microRNAs. *J. Cell. Physiol.* **2012**, *227*, 3163–3168.
11. Mallory, A.C.; Bouche, N. MicroRNA-directed regulation: To cleave or not to cleave. *Trends Plant Sci.* **2008**, *13*, 359–367.
12. Mallory, A.C.; Elmayan, T.; Vaucheret, H. MicroRNA maturation and action—The expanding roles of ARGONAUTES. *Curr. Opin. Plant Biol.* **2008**, *11*, 560–566.
13. Baulcombe, D. RNA silencing in plants. *Kent* **2004**, *431*, 356–363.
14. Wu, L.; Zhou, H.; Zhang, Q.; Zhang, J.; Ni, F.; Liu, C.; Qi, Y. DNA methylation mediated by a microRNA pathway. *Mol. Cell* **2010**, *38*, 465–475.
15. Chellappan, P.; Xia, J.; Zhou, X.; Gao, S.; Zhang, X.; Coutino, G.; Vazquez, F.; Zhang, W.; Jin, H. siRNAs from miRNA sites mediate DNA methylation of target genes. *Nucleic Acids Res.* **2010**, *38*, 6883–6894.
16. Khraiweh, B.; Zhu, J.K.; Zhu, J. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim. Biophys. Acta* **2012**, *1819*, 137–148.
17. Rhoades, M.W.; Reinhart, B.J.; Lim, L.P.; Burge, C.B.; Bartel, B.; Bartel, D.P. Prediction of plant microRNA targets. *Cell* **2002**, *110*, 513–520.

18. Noctor, G.; Foyer, C.H. Ascorbate and glutathione: Keeping active oxygen under control. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1998**, *49*, 249–279.
19. Dat, J.; Vandenabeele, S.; Vranova, E.; van Montagu, M.; Inze, D.; van Breusegem, F. Dual action of the active oxygen species during plant stress responses. *Cell. Mol. Life Sci.* **2000**, *57*, 779–795.
20. Foyer, C.H.; Noctor, G. Redox homeostasis and antioxidant signaling: A metabolic interface between stress perception and physiological responses. *Plant Cell* **2005**, *17*, 1866–1875.
21. Scandalios, J.G. Oxidative stress: Molecular perception and transduction of signals triggering antioxidant gene defenses. *Braz. J. Med. Biol. Res.* **2005**, *38*, 995–1014.
22. Sunkar, R.; Kapoor, A.; Zhu, J.K. Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in *Arabidopsis* is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cell* **2006**, *18*, 2051–2065.
23. Kozomara, A.; Griffiths-Jones, S. miRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **2011**, *39*, D152–D157.
24. 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.
25. Griffiths-Jones, S.; Saini, H.K.; van Dongen, S.; Enright, A.J. miRBase: Tools for microRNA genomics. *Nucleic Acids Res.* **2008**, *36*, D154–D158.
26. Gébelin, V.; Argout, X.; Engchuan, W.; Pitollat, B.; Duan, C.; Montoro, P.; Leclercq, J. Identification of novel microRNAs in *Hevea brasiliensis* and computational prediction of their targets. *BMC Plant Biol.* **2012**, *12*, doi:10.1186/1471-2229-12-18.
27. Lertpanyasampatha, M.; Gao, L.; Kongsawadworakul, P.; Viboonjun, U.; Chrestin, H.; Liu, R.; Chen, X.; Narangajavana, J. Genome-wide analysis of microRNAs in rubber tree (*Hevea brasiliensis* L.) using high-throughput sequencing. *Planta* **2012**, *236*, 437–445.
28. Gébelin, V.; Leclercq, J.; Kuswanhadi; Argout, X.; Chaidamsari, T.; Hu, S.; Tang, C.; Sarah, G.; Yang, M.; Montoro, P. The small RNA profile in latex from *Hevea brasiliensis* trees is affected by Tapping Panel Dryness. *Tree Physiol.* **2013**, in press.
29. Voinnet, O. Origin, biogenesis, and activity of plant microRNAs. *Cell* **2009**, *136*, 669–687.
30. Duan, C.; Argout, X.; Gebelin, V.; Summo, M.; Dufayard, J.F.; Leclercq, J.; Hadi, K.; Piyatrakul, P.; Pirrello, J.; Rio, M.; *et al.* Identification of the *Hevea brasiliensis* AP2/ERF superfamily by RNA sequencing. *BMC Genomics* **2013**, *14*, doi:10.1186/1471-2164-14-30.
31. Duan, C.; Rio, M.; Leclercq, J.; Bonnot, F.; Oliver, G.; Montoro, P. Gene expression pattern in response to wounding, methyl jasmonate and ethylene in the bark of *Hevea brasiliensis*. *Tree Physiol.* **2010**, *30*, 1349–1359.
32. Reyes, J.L.; Chua, N.H. ABA induction of miR159 controls transcript levels of two MYB factors during *Arabidopsis* seed germination. *Plant J.* **2007**, *49*, 592–606.
33. Lu, S.; Sun, Y.H.; Shi, R.; Clark, C.; Li, L.; Chiang, V.L. Novel and mechanical stress-responsive microRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. *Plant Cell* **2005**, *17*, 2186–2203.
34. Sanan-Mishra, N.; Kumar, V.; Sopory, S.K.; Mukherjee, S.K. Cloning and validation of novel miRNA from basmati rice indicates cross talk between abiotic and biotic stresses. *Mol. Genet. Genomics* **2009**, *282*, 463–474.

35. Sunkar, R.; Zhu, J.K. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell* **2004**, *16*, 2001–2019.
36. Lu, S.; Sun, Y.H.; Chiang, V.L. Stress-responsive microRNAs in *Populus*. *Plant J.* **2008**, *55*, 131–151.
37. Liu, H.H.; Tian, X.; Li, Y.J.; Wu, C.A.; Zheng, C.C. Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. *RNA* **2008**, *14*, 836–843.
38. Ding, D.; Zhang, L.; Wang, H.; Liu, Z.; Zhang, Z.; Zheng, Y. Differential expression of miRNAs in response to salt stress in maize roots. *Ann. Bot.* **2009**, *103*, 29–38.
39. Ben Chaabane, S.; Liu, R.; Chinnusamy, V.; Kwon, Y.; Park, J.-H.; Kim, S.Y.; Zhu, J.-K.; Yang, S.W.; Lee, B.-H. STA1, an *Arabidopsis* pre-mRNA processing factor 6 homolog, is a new player involved in miRNA biogenesis. *Nucleic Acids Res.* **2013**, *41*, 1984–1997.
40. Cuperus, J.T.; Fahlgren, N.; Carrington, J.C. Evolution and functional diversification of *MIRNA* genes. *Plant Cell* **2011**, *23*, 431–442.
41. Axtell, A. JPM patient information. Depression in palliative care. *J. Palliat. Med.* **2008**, *11*, 529–530.
42. Zhu, C.; Ding, Y.; Liu, H. MiR398 and plant stress responses. *Physiol. Plant* **2011**, *143*, 1–9.
43. Lardet, L.; Martin, F.; Dessailly, F.; Carron, M.P.; Montoro, P. Effect of exogenous calcium on post-thaw growth recovery and subsequent plant regeneration of cryopreserved embryogenic calli of *Hevea brasiliensis* (Mull. Arg.). *Plant Cell Rep.* **2007**, *26*, 559–569.
44. Piyatrakul, P.; Putranto, R.A.; Martin, F.; Rio, M.; Dessailly, F.; Leclercq, J.; Dufayard, J.F.; Lardet, L.; Montoro, P. Some ethylene biosynthesis and *AP2/ERF* genes reveal a specific pattern of expression during somatic embryogenesis in *Hevea brasiliensis*. *BMC Plant Biol.* **2012**, *12*, doi:10.1186/1471-2229-12-244.
45. German, M.A.; Luo, S.; Schroth, G.; Meyers, B.C.; Green, P.J. Construction of Parallel Analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nat. Protoc.* **2009**, *4*, 356–362.

Reprinted from *IJMS*. Cite as: Matsui, A.; Nguyen, A.H.; Nakaminami, K.; Seki, M. *Arabidopsis* Non-Coding RNA Regulation in Abiotic Stress Responses. *Int. J. Mol. Sci.* **2013**, *14*, 22642-22654.

Review

***Arabidopsis* Non-Coding RNA Regulation in Abiotic Stress Responses**

Akihiro Matsui¹, Anh Hai Nguyen¹, Kentaro Nakaminami¹ and Motoaki Seki^{1,2,3,*}

¹ Plant Genomic Network Research Team, RIKEN Center for Sustainable Resource Science, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan; E-Mails: akihirom@psc.riken.jp (A.M.); nguyenhaianh@psc.riken.jp (A.H.N.); nm_ken@psc.riken.jp (K.N.)

² Kihara Institute for Biological Research, Yokohama City University, 641-12 Maioka-cho, Totsuka-ku, Yokohama, Kanagawa 244-0813, Japan

³ CREST, JST, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

* Author to whom correspondence should be addressed; E-Mail: motoaki.seki@riken.jp; Tel.: +81-45-503-9587.

Received: 15 October 2013; in revised form: 31 October 2013 / Accepted: 31 October 2013 / Published: 18 November 2013

Abstract: Plant growth and productivity are largely affected by environmental stresses. Therefore, plants have evolved unique adaptation mechanisms to abiotic stresses through fine-tuned adjustment of gene expression and metabolism. Recent advanced technologies, such as genome-wide transcriptome analysis, have revealed that a vast amount of non-coding RNAs (ncRNAs) apart from the well-known housekeeping ncRNAs such as rRNAs, tRNAs, small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs) are expressed under abiotic stress conditions. These various types of ncRNAs are involved in chromatin regulation, modulation of RNA stability and translational repression during abiotic stress response. In this review, we summarize recent progress that has been made on ncRNA research in plant abiotic stress response.

Keywords: non-coding RNA; small RNA; antisense RNA; abiotic stress response

1. Introduction

Environmental stresses, such as drought, heat, salinity and low temperature, are major limiting factors for plant growth and productivity. Under natural conditions, plants are exposed to a variety of environmental stresses. In order to adapt and survive under the stresses, plants have evolved various molecular mechanisms for a fine-tuned control of adaptive responses [1]. Post-transcriptional regulatory mechanisms, as well as epigenetic and post-translational modifications, like ubiquitination and sumoylation, have been implicated to play an important role in the regulation of gene expression during stress conditions.

Recent genome-wide transcriptome analysis, such as tiling arrays and next generation sequencing, has revealed a large number of stress-responsive ncRNAs. Emerging evidence has revealed that ncRNAs are major products of the plant transcriptome with significant regulatory importance [2,3]. ncRNAs are transcribed from intergenic regions, antisense strands of protein-coding genes and also pseudogenes. According to their size, ncRNAs are classified as small ncRNAs (sRNAs) (<40 nt) and long ncRNAs (lncRNAs) (>200 nt). These ncRNAs are involved in the transcriptional and posttranscriptional regulation of gene expression and the modulation of RNA stability and translation under stress conditions [1,4–7].

2. Small RNAs (sRNAs)

sRNAs are known to have major functional roles in eukaryotic gene regulation. In plants, knowledge regarding the biogenesis and mechanisms of action of sRNA classes including microRNAs (miRNAs), transcriptional gene silencing (TGS)-related heterochromatic small interfering RNAs (siRNAs), trans-acting siRNAs (ta-siRNAs) and natural antisense transcript siRNAs (nat-siRNA) has been primarily gained through *Arabidopsis* studies (Figure 1; Table 1). These sRNAs are loaded into RNA-induced silencing complexes (RISC) and negatively regulate the expression of their target genes by affecting mRNA levels, chromatin remodeling and DNA methylation.

Figure 1. Regulation of ncRNAs in abiotic stress responses. ncRNAs (lncRNAs, miRNAs and siRNAs, *etc.*) are generated in response to abiotic stress, such as drought, low-temperature, heat and high-salinity. The ncRNAs are involved in various types of regulation, such as chromatin regulation, transcriptional regulation, RNA degradation, protection of miRNA targets, translational repression.

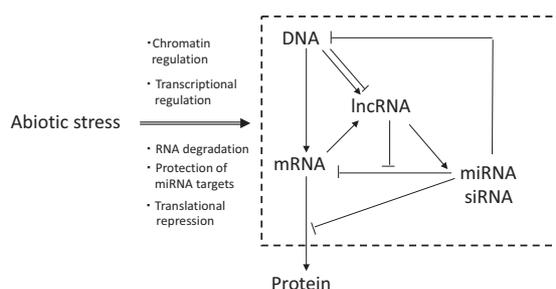


Table 1. Classification of non-coding RNAs (ncRNAs) ⁽¹⁾ involved in abiotic stress responses.

| Small ncRNAs (sRNAs) | |
|------------------------------|--|
| • | micro RNAs (miRNAs) |
| • | small interfering RNAs (siRNAs) |
| | TGS-related heterochromatic siRNAs |
| | <i>trans</i> -acting siRNAs (ta-siRNAs) |
| | natural antisense transcript siRNAs (nat-siRNA) |
| | others |
| Long ncRNAs (lncRNAs) | |
| • | long antisense non-coding RNAs (lancRNAs) |
| | lancRNAs generated during biogenesis of ta-siRNAs and S-PTGS-related siRNAs |
| | lancRNAs involved in epigenetic silencing |
| | lancRNAs co-expressed with sense-stranded protein-coding mRNAs |
| | others |
| • | long intergenic non-coding RNAs (lincRNAs) |
| | RNAs generated from small open reading frame regions |
| | RNA decoys mimicking miRNA targets |
| | RNAs transcribed by Pol III |
| | RNAs transcribed by Pol IV during biogenesis of TGS-related heterochromatic siRNAs |
| | others |

⁽¹⁾ ncRNAs are classified except for well-known housekeeping ncRNAs, such as rRNAs, tRNAs, snRNAs and snoRNAs.

2.1. *Micro RNAs (miRNAs)*

Many stress-responsive miRNAs have been identified in various plants [6,8]. miRNAs are generated from pri-miRNAs that are transcribed by RNA polymerase II and the mature miRNAs guide to cleavage target mRNAs [9]. Therefore, the expression levels of pri-miRNAs are regulated by *cis*-elements in a similar manner as protein-coding genes. An abundance of positive cold stress-related *cis*-regulatory elements, such as the Dehydration Responsive Element/Low Temperature Responsive Element (DRE/LTRE) -core (A/GCCGAC) [10], Abscisic Acid-Responsive Element (ABRE) -core (ACGTGG/TC) [11] and W-box (TTGAC) [12] are found in the promoter region of several cold-inducible *MIRNA* genes in *Arabidopsis* [13]. Several miRNAs that are upregulated in response to various abiotic stresses, including cold [14], dehydration [14,15], salinity [14] and nutrient deficiency [16] have been reported.

An *Arabidopsis* miRNA involved in the detoxification of reactive oxygen species (ROS) has been reported [17]. Expression of miR398 was found to be decreased by copper (Cu(II)), which is an essential nutrient in photosynthesis and response to oxidative stress and abiotic stresses [18]. miR398 was also shown to suppress superoxide dismutase (SOD) and Cu/Zinc (Zn) superoxide dismutases (CSDs) under low Cu. Interestingly, it was reported as a rare instance in plants that miR398 not only digests *CSD1* and *CSD2* mRNAs but also negatively regulated the translation of their protein products [19]. This miRNA-dependent translational repression is effected in part by the ARGONAUTE proteins AGO1 and AGO10 [19]. It also requires the activity of the microtubule-severing enzyme katanin and the de-capping component VARICOSE (VCS)/Ge-1, as recently suggested from animal studies. *katanin1* and *vcs1* mutants that did not affect *CSD2* mRNA accumulation, exhibited an

overaccumulation of CSD2 proteins under low Cu (II) conditions in comparison to WT plants. However, under high Cu (II) conditions, this trend was not observed [19]. This result suggested that the translational inhibition of miR398 had an important role in the regulation of CSD expression under low Cu (II) condition.

Several miRNAs function in the maintenance of phytohormone signaling during exposure to abiotic stress. An miR168-mediated feedback regulatory loop regulates AGO1 homeostasis in ABA and abiotic stress responses in *Arabidopsis* [20]. *MIR168a*-overexpressing plants and *ago1* loss-of-function mutants showed ABA hypersensitivity and drought tolerance, while the *mir168a* mutants showed ABA hyposensitivity and drought hypersensitivity [20]. The promoter of *MIR168a* has four ABREs, suggesting that expression of *MIR168a* was directly induced by ABA. Although promoter activity of *AGO1* was induced by ABA, *AGO1* transcripts were negatively regulated by ABA-induced miR168a; and as a result, mRNA was maintained at a steady level during the stress response. These results suggest that a complex crosstalk exists between the global regulation of miRNA metabolism and ABA signaling functions to enable fine-tuning of abiotic stress response.

In another example of miRNA-phytohormone crosstalk in abiotic stress response, miR160 regulates the expression of *Auxin Response Factors* (*ARF10*, *ARF16* and *ARF17*) [21]. *Arabidopsis* plants expressing miR160-resistant *ARF10* not only showed an abnormal leaf shape but also showed hypersensitivity to ABA. Plants overexpressing miR160 showed hyposensitivity to ABA. In addition, other miRNAs targeting auxin signaling factors were also induced in response to abiotic stress [14]. Expression of miR393 was up-regulated by dehydration, salt and cold stresses and ABA [8]. A miR393 target gene, *TIR1*, an auxin receptor, is involved in the response to salt and oxidative stresses [22]. The promoter of miR167 contains ABREs, indicating their own regulation by ABA [14]. *ARF6* and *ARF8*, which are targeted by miR167, are regulators of female and male reproduction [23]. *TAS3*-siRNA also regulates auxin signaling [24]. It is thought that miRNAs targeting auxin signaling function as mediators that connect abiotic signaling with development [7,25]. These results suggest that multistep regulation by miRNAs is required for the correct adjustment of gene expression under abiotic stress.

Nutrient deficiency under abiotic stress is known to induce or suppress various miRNAs that regulate nutrient metabolism. Previous studies have demonstrated that the expression of miR395 was increased by sulfate starvation [26]. This specific miRNA suppresses ATP sulfurylases as target mRNAs, thus resulting in catalysis of the first step of inorganic sulfate assimilation [8]. A phosphate starvation-inducible miRNA (miR399) regulates Pi homeostasis by regulating the expression of *UBC24* mRNA encoding an ubiquitin-conjugating E2 enzyme [16,27]. miR399 functions as a positive regulator of Pi uptake and translocation. In addition to miR399 regulating *UBC24* expression, the cleavage activity of miR399 is suppressed by a long intergenic ncRNA, Induced by Phosphate Starvation 1 (IPS1) [28]. A detailed description is provided at a later point in this chapter.

2.2. Small Interfering RNAs (siRNAs)

Recent studies have reported that small interfering RNAs (siRNAs) function in stress responses. These siRNAs are generated from long double strand RNAs through various biological processes [29].

A specific type of siRNA that is involved in RNA-directed DNA methylation suppresses the activation of retrotransposons under heat stress [30]. The transcriptional gene silencing (TGS)-related heterochromatic siRNAs are generated from RNA Polymerase IV (PolIV)-derived transcripts in repetitive DNA sequences and heterochromatin. After heat stress, a *Copia*-type retrotransposon in *Arabidopsis*, named *ONSEN*, becomes transcriptionally active and has been shown to result in the synthesis of extrachromosomal DNA copies in the siRNA-mediated silencing deficient mutant *nrpd1* (*nrpd1a*), which is the largest subcomponent of PolIV. Heat-induced expression and transgenerational retrotransposition of *ONSEN* were suppressed by siRNA-mediated silencing. It was also reported that abiotic stresses changed the genome-wide DNA methylation status across multiple generations [31].

Trans-acting siRNAs (ta-siRNA) can be classified as a specialized case of siRNAs in plants. These siRNAs are generated from dsRNAs, which are generated from miR173, miR390 and miR828 - cleaved lncRNAs [32,33]. The long antisense non-coding RNAs (lancRNAs) of these dsRNAs were synthesized by RNA-dependent RNA polymerase 6 (RDR6), those dsRNAs were positive candidates producing siRNAs [34]. The tight feedback regulation between miR390, *TAS3* ta-siRNA and ARF4, which is a target of *TAS3* ta-siRNA, was required for lateral root initiation [24]. Although ta-siRNA expression has not been found to change in response to abiotic stress in *Arabidopsis*, the expression of a rice *RDR6* homolog, which produces dsRNAs from cleaved *TAS* RNAs, was induced by ABA in rice [35]. These results imply that *TAS3* ta-siRNA is involved in dynamic changes of root architecture during exposure to abiotic stress [36].

Borsani *et al.* 2005 reported that natural antisense transcript small interfering RNA (nat-siRNAs) were generated from dsRNAs produced from natural *cis*-antisense gene pairs of $\Delta 1$ -*Pyrrroline-5-Carboxylate Dehydrogenase* (*P5CDH*) and a high-salinity-stress inducible gene of unknown function (*SRO5*) during high-salinity stress [37]. Recent genome-wide analysis reported an accumulation of sRNAs in their overlapping region, suggesting the occurrence of an RNA interference event [38]. However, the biological process of generating nat-siRNAs is not completely understood at this time [39,40].

3. Long Non-Coding RNAs (lncRNAs)

Genome-wide tiling arrays and high-throughput sequencing have identified a vast amount of lncRNAs in plants [4,41–45]. These lncRNAs may represent alternatively spliced forms of known genes [46], products of antisense RNAs [4,38,47,48], double stranded RNAs [49], retained introns [46,50], short open reading frame [34,51,52], RNA polymerase III-derived RNAs [53] and RNA decoys mimicking miRNA targets [28]. In this manuscript, we classified the lncRNAs into long antisense non-coding RNAs (lancRNAs) and long intergenic non-coding RNAs (lincRNAs) based on its genomic locations (Figure 1; Table 1). These RNAs have various modifications that depend on each biosynthetic process.

3.1. Long Antisense Non-Coding RNAs (lancRNAs)

Over the past decade, genome-wide transcriptome analyses confirmed that approximately 30% of all annotated genes exhibited significant lancRNA expression in *Arabidopsis* [4,41]. These data regarding lancRNA expression are consistent with results from other organisms such as fly, human,

and rice [54]. Expression profiles of lincRNAs in response to environmental stresses have been extensively characterized by an *Arabidopsis* tiling array analysis [4]. A certain type of lincRNAs belongs to a pair of fully overlapping sense–antisense transcripts (fSATs) in which the lincRNAs exist within the protein-coding gene regions in opposing orientations. The expression of these sense and antisense RNA transcripts are stress-responsive. On the other hand, partially overlapping sense–antisense transcripts (pSATs) do not exhibit synchronous expression patterns. These observations suggest that lincRNA were generated through multiple biosynthetic processes.

A type of lincRNA is co-expressed with sense protein-coding RNAs [4]. A large *Arabidopsis* tiling array analysis confirmed that more than 6000 lincRNAs were classified into the fSATs category. Interestingly, a significant linear correlation between the expression ratios (abiotic stress treated/untreated) of the sense transcripts and the ratios of the lincRNAs was observed in the fSATs. The *RD29A* and *CYP707A1* lincRNAs that were simultaneously accumulated with sense mRNAs, were accumulated by drought- and ABA treatments. Some of the lincRNAs that were identified contained complementary sequences to those of the sense mRNAs [4], indicating that lincRNA expression is dependent upon sense mRNAs.

Co-expression of sense RNA and lincRNA of a transgene was reported as a trigger for sense post-transcriptional gene silencing (S-PTGS) [55]. In the S-PTGS process, RNA-dependent RNA polymerase 6 (RDR6), one of six *Arabidopsis* RDRs, generates antisense RNAs from non-canonical sense RNAs of transgenes with aberrant features, such as non-cap structure of 5' end or poly(A) tail, to generate double stranded RNA (dsRNA) [56]. dsRNA-seq analysis showed that dsRNAs of more than 100 loci were reduced in *rdr6* in relative comparison to WT [57]. RNA-seq analysis of sRNAs also revealed that mutations of *ABH1* and *EIN5* (*XRN4*), which are involved in mRNA processing and mRNA degradation, respectively, affect the level of sRNAs mapped on the antisense strand of endogenous protein-coding genes [58]. The *xrn4* mutant was also screened as an enhancer of transgene silencing [59]. These results supported the observation that a certain type of lincRNA was generated from non-canonical sense RNA. Recent studies revealed that *Werner Exonuclease* (*WEX*), *Silencing Defective3* (*SDE3*), *DCL2*, *DCL4*, *Nuclear RNA Polymerase IVa* (*NRPD1a*), *RDR2* and *CLASSY1* were involved in S-PTGS and downstream PTGS [60–65]. Since S-PTGS-related siRNAs are generated from dsRNAs, these results also indicate that a certain type of lincRNA is synthesized from mRNA templates via a complex amplification pathway.

It is possible that the expression of lincRNA could serve as a functional link to the chromatin regulation of epigenetic silencing [48]. *Cold Induced Long Antisense Intragenic RNA* (*COOLAIR*) in the *FLC* locus is a well-characterized example of this in *Arabidopsis* [47,66]. Exposure of plants to low temperature treatment for a 2-week period resulted in a high level of *COOLAIR* expression. Several weeks after the induction of *COOLAIR*, the transcription of *FLC* was significantly decreased. During this period, tri-methylated histone H3 Lys27 (H3K27me3) levels progressively increased at a region around the transcription start site [67]. The level of H3K27me3 spreading the gene body was required to maintain the repression of *FLC* transcription after plants were returned to warm conditions [68]. *COOLAIR* has been suggested to be required for a plant homeodomain-polycomb repressive complex 2 (PHD-PRC2) located at a tightly localized nucleation region within *FLC*. Consequently, this results in an increase in H3K27me3 levels at the *FLC* locus [64]. Although further analysis is necessary to elucidate the role of *COOLAIR* and its epigenetic silencing of *FLC* during the short period

of vernalization [69,70], a reporter gene that was fused with *COOLAIR* was shown and confirmed to be capable of causing cold-induced silencing [47].

The second class of lncRNA, *COLDAIR*, is transcribed from a region within the first intron of *FLC* on the sense strand [71]. The *COLDAIR* transcript has been shown to interact with PRC2 and its abundance also increased during vernalization. Reduction of *COLDAIR* transcript levels by RNAi confirmed that it is not required for the initial repression of *FLC* but is necessary for subsequent maintenance of repression. These results showed that the interaction of lncRNAs and the chromatin modification complex mediates cold-inducible epigenetic regulation. Based upon bioinformatic analyses comparing lncRNAs and chromatin status, it was recently hypothesized that a type of lncRNAs from the genome was repressed by cytosine methylation and H3K36me [48]. This prediction proposed that an interaction between chromatin regulation and a certain type of lncRNAs functions in genome regulation.

A recent strand-specific RNA-seq study showed that approximately 1300 ncRNA loci exist in the antisense strand of protein-coding genes in *Arabidopsis* [48]. The global ratio between sense and antisense tags in exons in this study was 0.01, which was similar to a previous strand-specific RNA-seq analysis using floral tissue [43,48]. A comparison between RNA-seq and genome-wide tiling array data showed that one-half to two-thirds of the sense–antisense transcripts were only represented in one experiment [48]. Since the expression levels of lncRNAs are low compared to sense transcripts, it is possible that this difference between the two techniques may reflect technical limitations inherent to these methods. Future transcriptome analyses are required to increase our understanding of non-canonical transcripts and to clarify the types of lncRNAs that function in molecular signals, RNA decoys, guides, and scaffolds [72].

3.2. Long Intergenic Non-Coding RNAs (lincRNAs)

Several transcriptome analyses have reported that more than 1000 lincRNA loci exist between protein-coding genes in *Arabidopsis* [4,41–45]. A part of these lincRNAs was transcribed from the methylated DNA regions by RNA polymerase IV and are thought to be positive candidates that generate into TGS-related heterochromatic siRNAs that guide DNA methylation [64].

Another type of lincRNA is transcribed by RNA Polymerase III. *In silico* genome sequence analysis predicted 20 novel ncRNA candidates [53]. A specific Pol III-derived ncRNA (*AtR18*) responded negatively to hypoxic stress and this regulation was evidently different from that of *U6* snRNA. Specifically, *AtR18* was not processed into a smaller fragment and no small open reading frames (sORFs) were included. Short interspersed elements (SINEs) and 7SL (signal recognition particle) RNA for protein trafficking are known as Pol III-derived RNA, with exception of the canonical functional RNA [73,74].

Many sORFs that have not been annotated as protein-coding genes have been identified as expressed genes during developmental or environmental conditions [51,52]. A specific sORF (*npc536*) has a large dynamic variation of expression across a wide range of tissue and hormonal, biotic, or abiotic treatment [34]. *npc536* exists in the antisense strand of a Golgi-transport complex related protein. *35S::npc536* transformants displayed heightened root growth under salt stress conditions. However, the biological function of this RNA is not still clearly understood at this time.

LincRNA is capable of regulating the cleavage activity of miRNA as a target mRNA decoy. *IPSI* is a lincRNA, which was found to be induced by phosphate starvation [28,75]. Although *Arabidopsis IPSI*, *AT4-1*, *AT4-2* and family members in other species share little sequence conservation with the predicted ORFs, high conservation was observed with a 22-nt sequence located in the 3' half of the transcript [75]. This observation suggests that the *IPSI* family contains the sequence that is targeted by miRNA and it does not function as a short peptide. The 22-nt sequence of *IPSI* is not perfectly complementary to miR399 due to a mismatch occurring at the 10–11th position [28]. Although miR399 was capable of hybridizing to *IPSI* transcript, there has been no indication of cleavage via miR399. On a whole plant level, phosphate starvation resulted in the induction of *IPSI*. As a result of the phosphate starvation, the miR399 target (*UBC24* mRNA) accumulated. miR399 has been characterized as a phosphate starvation inducible-miRNA and a positive regulator of Pi-uptake. Interestingly, two studies implied that miR399 may function as a long-distance signal from shoots to suppress *UBC24* expression in roots [76,77]. Taken together, it is reasonable to consider that the complex interactions between *IPSI*, miR399 and *UBC24* function to maintain spatial homeostasis of Pi. Computational analyses predicted the occurrence of these possible target mimics in rice and *Arabidopsis*, suggesting that RNA decoys (miRNA target mimic) are conserved and functional contributors to the regulation of RNA [78].

4. Conclusions and Perspectives

Whole genome transcriptome analyses of high-density microarrays and high throughput sequencing have expanded our understanding of a novel research area pertaining to networks, which govern abiotic stress responses. Despite the extensive body of published information regarding various types of ncRNAs, their association to the adaptation and response of plants to abiotic stress is not completely understood at this time. Recent advances in whole transcriptome analyses have enabled us to gain a greater understanding regarding the mechanisms of transcriptional and posttranscriptional regulation through ncRNAs, as well as sRNAs and lincRNAs. Functional analyses of miRNAs have deciphered and unraveled complex feedback loop regulations between miRNA and target genes. In addition, the characterized interaction between miR399, *IPSI* and *UBC24* has been shown to be a novel type of RNA interaction system in which a lincRNA reduced a cleavage activity of phosphate starvation-inducible miRNA and regulated a target mRNA. In addition, ncRNAs are also linked to chromatin regulation. The studies regarding epigenetics in the *FLC* locus enabled a hypothesis to be made which suggests that an interaction of lincRNA and chromatin modification complex mediates cold-inducible epigenetic regulation. Overall, these data have revealed a complex interaction between transcriptional regulators which function to fine-tune responses to various environmental stimuli. For the future, it will be essential to investigate the biological regulations of ncRNAs in order to enable scientists to completely elucidate the entire picture of gene regulation networks in stress response.

Acknowledgments

This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan (Innovative Areas 25119724, Challenging Exploratory Research 24657041), Japan Science and Technology Agency (JST), Core Research for Evolutionary

Science and Technology (CREST) and grants from the RIKEN Research Institute, Japan, to Motoaki Seki.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Hirayama, T.; Shinozaki, K. Research on plant abiotic stress responses in the postgenome era: Past, present and future. *Plant J.* **2010**, *61*, 1041–1052.
2. Laporte, P.; Merchan, F.; Amor, B.B.; Wirth, S.; Crespi, M. Riboregulators in plant development. *Biochem. Soc. Trans.* **2007**, *35*, 1638–1642.
3. Rymarquis, L.A.; Kastenmayer, J.P.; Huttenhofer, A.G.; Green, P.J. Diamonds in the rough: mRNA-like non-coding RNAs. *Trends Plant Sci.* **2008**, *13*, 329–334.
4. Matsui, A.; Ishida, J.; Morosawa, T.; Mochizuki, Y.; Kaminuma, E.; Endo, T.A.; Okamoto, M.; Nambara, E.; Nakajima, M.; Kawashima, M.; *et al.* *Arabidopsis* transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. *Plant Cell Physiol.* **2008**, *49*, 1135–1149.
5. Mazzucotelli, E.; Mastrangelo, A.M.; Crosatti, C.; Guerra, D.; Stanca, A.M.; Cattivelli, L. Abiotic stress response in plants: When post-transcriptional and post-translational regulations control transcription. *Plant Sci.* **2008**, *174*, 420–431.
6. Khraiweh, B.; Zhu, J.K.; Zhu, J. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim. Biophys. Acta* **2012**, *1819*, 137–148.
7. De Lima, J.C.; Loss-Morais, G.; Margis, R. microRNAs play critical roles during plant development and in response to abiotic stresses. *Genet. Mol. Biol.* **2012**, *35*, 1069–1077.
8. Sunkar, R.; Zhu, J.K. Novel and stress-regulated MicroRNAs and other small RNA from *Arabidopsis*. *Plant Cell* **2004**, *16*, 2001–2019.
9. Voinnet, O. Origin, biogenesis, and activity of plant microRNAs. *Cell* **2009**, *136*, 669–687.
10. Yamaguchi-Shinozaki, K.; Shinozaki, K. A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness. *Plant Cell* **1994**, *6*, 251–264.
11. Mundy, J.; Yamaguchi-Shinozaki, K.; Chua, N.H. Nuclear proteins bind conserved elements in the abscisic acid-responsive promoter of a rice *rab* gene. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 1406–1410.
12. Shang, Y.; Yan, L.; Liu, Z.Q.; Zheng, C.; Chao, M.; Xin, Q.; Wu, F.Q.; Wang, X.F.; Du, S.Y.; Jiang, T.; *et al.* The Mg-chelatase H subunit of *Arabidopsis* antagonizes a group of transcription repressors to relieve ABA-responsive genes of inhibition. *Plant Cell* **2010**, *22*, 1909–1935.
13. Zhou, X.; Wang, G.; Sutoh, K.; Zhu, J.K.; Zhang, W. Identification of cold-inducible microRNAs in plants by transcriptome analysis. *Biochim. Biophys. Acta* **2008**, *1779*, 780–788.
14. Liu, H.H.; Tian, X.; Li, Y.U.; Wu, C.A.; Zheng, C.C. Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. *RNA* **2008**, *14*, 836–843.
15. Zhou, L.; Liu, Y.; Liu, Z.; Kong, D.; Duan, M.; Luo, L. Genome-wide identification and analysis of drought-responsive microRNAs in *Oryza sativa*. *J. Exp. Bot.* **2010**, *61*, 4157–4168.

16. Fuji, H.; Chiou, T.J.; Lin, S.I.; Aung, K.; Zhu, J.K. A miRNA involved in phosphate starvation response in *Arabidopsis*. *Curr. Biol.* **2005**, *15*, 2038–2043.
17. Sunkar, R.; Kapoor, A.; Zhu, J.K. Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in *Arabidopsis* is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cell* **2006**, *18*, 2051–2065.
18. Yamasaki, H.; Abdel-Ghany, S.E.; Cohu, C.M.; Kobayashi, Y.; Shikanai, T.; Pilon, M. Regulation of copper homeostasis by micro-RNA in *Arabidopsis*. *J. Biol. Chem.* **2007**, *282*, 16369–16378.
19. Brodersen, P.; Sakvarelidze-Achard, L.; Bruun-Rasmussen, M.; Dunoyer, P.; Yamamoto, Y.Y.; Sieburth, L.; Voinnet, O. Widespread translational inhibition by plant miRNAs and siRNAs. *Science* **2008**, *320*, 1185–1190.
20. Li, W.; Cui, X.; Meng, Z.; Huang, X.; Xie, Q.; Wu, H.; Jin, H.; Zhang, D.; Liang, W. Transcriptional regulation of *Arabidopsis* miR168a and ARGONAUTE1 homeostasis in abscisic acid and abiotic stress responses. *Plant Physiol.* **2013**, *158*, 1279–1292.
21. Liu, P.P.; Montgomery, T.A.; Fahlgren, N.; Kasschau, K.D.; Nonogaki, H.; Carrington, J.C. Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is critical for seed germination and post-germination stages. *Plant J.* **2007**, *52*, 133–146.
22. Blomster, T.; Salojarvi, J.; Sipari, N.; Brosche, M.; Ahlfors, R.; Keinanen, M.; Overmyer, K.; Kangasjarvi, J. Apoplastic reactive oxygen species transiently decrease auxin signaling and cause stress-induced morphogenic response in *Arabidopsis*. *Plant Physiol.* **2011**, *157*, 1866–1883.
23. Wu, M.F.; Tian, Q.; Reed, J.W. *Arabidopsis* microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. *Development* **2006**, *133*, 4211–4218.
24. Marin, E.; Jouannet, V.; Herz, A.; Lokerse, A.S.; Weijers, D.; Vaucheret, H.; Nussaume, L.; Crespi, M.D.; Maizel, A. miR390, *Arabidopsis* TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. *Plant Cell* **2010**, *22*, 1104–1117.
25. Iglesias, M.J.; Terrile, M.C.; Bartoli, C.G.; D’Ippolito, S.; Casalongue, C.A. Auxin signaling participates in the adaptative response against oxidative stress and salinity by interacting with redox metabolism in *Arabidopsis*. *Plant Mol. Biol.* **2010**, *74*, 215–222.
26. Jones-Rhoades, M.W.; Bartel, D.P. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol. Cell.* **2004**, *14*, 787–799.
27. Aung, K.; Lin, S.I.; Wu, C.C.; Huang, Y.T.; Su, C.L.; Chiou, T.J. *pho2*, a phosphate overaccumulator, is caused by a nonsense mutation in a microRNA399 target gene. *Plant Physiol.* **2006**, *141*, 1000–1011.
28. Franco-Zorrilla, J.M.; Valli, A.; Todesco, M.; Mateos, I.; Puga, M.I.; Rubio-Somoza, I.; Leyva, A.; Weigel, D.; Garcia, J.A.; Paz-Ares, J. Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat. Genet.* **2007**, *39*, 1033–1037.
29. Mallory, A.C.; Vaucheret, H. Functions of microRNAs and related small RNAs in plants. *Nat. Genet.* **2006**, *38*, S31–S36.
30. Ito, H.; Gaubert, H.; Bucher, E.; Mirouze, M.; Vaillant, I.; Paszkowski, J. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* **2011**, *472*, 115–119.

31. Boyko, A.; Blevins, T.; Yao, Y.; Golubov, A.; Bilichak, A.; Ilnytsky, Y.; Hollunder, J.; Meins, F., Jr.; Kovalchuk, I. Transgenerational adaptation of *Arabidopsis* to stress requires DNA methylation and the function of Dicer-like proteins. *PLoS One* **2010**, *5*, e9514.
32. Ronemus, M.; Vaughn, M.W.; Martienssen, R.A. MicroRNA-targeted and small interfering RNA-mediated mRNA degradation is regulated by argonaute, dicer, and RNA-dependent RNA polymerase in *Arabidopsis*. *Plant Cell* **2006**, *18*, 1559–1574.
33. 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.
34. Ben Amor, B.; Wirth, S.; Merchan, F.; Laporte, P.; d'Aubenton-Carafa, Y.; Hirsch, J.; Maizel, A.; Mallory, A.; Lucas, A.; Deragon J.M.; *et al.* Novel long non-protein coding RNAs involved in *Arabidopsis* differentiation and stress responses. *Genome Res.* **2009**, *19*, 57–69.
35. Yang, J.H.; Seo, H.H.; Han, S.J.; Yoon, E.K.; Yang, M.S.; Lee, W.S. Phytohormone abscisic acid control RNA-dependent RNA polymerase 6 gene expression and post-transcriptional gene silencing in rice cells. *Nucleic Acids Res.* **2008**, *36*, 1220–1226.
36. Bartels, D.; Sunkar, R. Drought and salt tolerance in plants. *Crit. Rev. Plant. Sci.* **2005**, *24*, 23–58.
37. 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.
38. Zhang, X.; Lii, Y.; Wu, Z.; Polishko, A.; Zhang, H.; Chinnusamy, V.; Lonardi, S.; Zhu, J.K.; Liu, R.; Jin, H. Mechanisms of small RNA generation from *cis*-NATs in response to environmental and developmental cues. *Mol. Plant* **2013**, *6*, 704–715.
39. Henz, S.R.; Cumbie, J.S.; Kasschau, K.D.; Lohmann, J.U.; Carrington, J.C.; Weigel, D.; Schmid, M. Distinct expression patterns of natural antisense transcripts in *Arabidopsis*. *Plant Physiol.* **2007**, *144*, 1247–1255.
40. Zhan, S.; Lukens, L. Protein-coding *cis*-natural antisense transcripts have high and broad expression in *Arabidopsis thaliana*. *Plant Physiol.* **2013**, *161*, 2171–2180.
41. Yamada, K.; Lim, J.; Dale, J.M.; Chen, H.; Shinn, P.; Palm, C.J.; Southwick, A.M.; Wu, H.C.; Kim, C.; Nguyen, M.; *et al.* Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* **2003**, *302*, 842–846.
42. Li, L.; Wang, X.F.; Stolc, V.; Li, X.Y.; Zhang, D.F.; Su, N.; Tongprasit, W.; Li, S.G.; Cheng, Z.K.; Wang, J.; *et al.* Genome-wide transcription analyses in rice using tiling microarrays. *Nat. Genet.* **2006**, *38*, 124–129.
43. Lister, R.; O'Malley, R.C.; Tonti-Filippini, J.; Gregory, B.D.; Berry, C.C.; Millar, A.H.; Ecker, J.R. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* **2008**, *133*, 523–536.
44. Laubinger, S.; Zeller, G.; Henz, S.R.; Sachsenberg, T.; Widmer, C.K.; Naouar, N.; Vuylsteke, M.; Scholkopf, B.; Ratsch, G.; Weigel, D. At-TAX: A whole genome tiling array resource for developmental expression analysis and transcript identification in *Arabidopsis thaliana*. *Genome Biol.* **2008**, *9*, R112.
45. Hazen, S.P.; Naef, F.; Quisel, T.; Gendron, J.M.; Chen, H.; Ecker, J.R.; Borevitz, J.O.; Kay, S.A. Exploring the transcriptional landscape of plant circadian rhythms using genome tiling arrays. *Genome Biol.* **2009**, *10*, R17.

46. Filichkin, S.A.; Priest, H.D.; Givan, S.A.; Shen, R.; Bryant, D.W.; Fox, S.E.; Wong, W.K.; Mockler, T.C. Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Genome Res.* **2010**, *20*, 45–58.
47. 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.
48. Luo, C.; Sidote, D.J.; Zhang, Y.; Kerstetter, R.A.; Michael, T.P.; Lam, E. Integrative analysis of chromatin states in *Arabidopsis* identified potential regulatory mechanisms for natural antisense transcript production. *Plant J.* **2013**, *73*, 77–90.
49. 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.
50. Ner-Gaon, H.; Halachmi, R.; Savaldi-Goldstein, S.; Rubin, E.; Ophir, R.; Fluhr, R. Intron retention is a major phenomenon in alternative splicing in *Arabidopsis*. *Plant J.* **2004**, *39*, 877–885.
51. Hanada, K.; Higuchi-Takeuchi, M.; Okamoto, M.; Yoshizumi, T.; Shimizu, M.; Nakaminami, K.; Nishi, R.; Ohashi, C.; Iida, K.; Tanaka, M.; *et al.* Small open reading frames associated with morphogenesis are hidden in plant genomes. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 2395–2400.
52. Moghe, G.D.; Lehti-Shiu, M.D.; Seddon, A.E.; Yin, S.; Chen, Y.; Juntawong, P.; Brandizzi, F.; Bailey-Serres, J.; Shiu, S.H. Characteristics and significance of intergenic polyadenylated RNA transcription in *Arabidopsis*. *Plant Physiol.* **2013**, *161*, 1210–1224.
53. Wu, J.; Okada, T.; Fukushima, T.; Tsudzuki, T.; Sugiura, M.; Yukawa, Y. A novel hypoxic stress-responsive long non-coding RNA transcribed by RNA polymerase III in *Arabidopsis*. *RNA Biol.* **2012**, *9*, 302–313.
54. Kapranov, P.; Willingham, A.T.; Gingeras, T.R. Genome-wide transcription and the implications for genomic organization. *Nat. Rev. Genet.* **2007**, *8*, 413–423.
55. Beclin, C.; Boutet, S.; Waterhouse, P.; Vaucheret, H. A branched pathway for transgene-induced RNA silencing in plants. *Curr. Biol.* **2002**, *12*, 684–688.
56. 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.
57. Zheng, Q.; Ryvkin, P.; Li, F.; Dragomir, I.; Valladares, O.; Yang, J.; Cao, K.; Wang, L.S.; Gregory, B.D. Genome-wide double-stranded RNA sequencing reveals the functional significance of base-paired RNAs in *Arabidopsis*. *PLoS Genet.* **2010**, *6*, e1001141.
58. Gregory, B.D.; O'Malley, R.C.; Lister, R.; Urich, M.A.; Tonti-Filippini, J.; Chen, H.; Millar, A.H.; Ecker, J.R. A link between RNA metabolism and silencing affecting *Arabidopsis* development. *Dev. Cell* **2008**, *14*, 854–866.
59. Gazzani, S.; Lawrenson, T.; Woodward, C.; Headon, D.; Sablowski, R. A link between mRNA turnover and RNA interference in *Arabidopsis*. *Science* **2004**, *306*, 1046–1048.
60. Dunoyer, P.; Himber, C.; Voinnet, O. DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. *Nat. Genet.* **2005**, *37*, 1356–1360.
61. Xie, Z.; Allen, E.; Fahlgren, N.; Calamar, A.; Givan, S.A.; Carrington, J.C. Expression of *Arabidopsis* miRNA genes. *Plant Physiol.* **2005**, *138*, 2145–2154.

62. Jauvion, V.; Rivard, M.; Bouteiller, N.; Elmayan, T.; Vaucheret, H. RDR2 partially antagonizes the production of RDR6-dependent siRNA in sense transgene-mediated PTGS. *PLoS One* **2012**, *7*, e29785.
63. Glazov, E.; Phillips, K.; Budziszewski, G.J.; Schob, H.; Meins, F., Jr.; Levin, J.Z. A gene encoding an RNase D exonuclease-like protein is required for post-transcriptional silencing in *Arabidopsis*. *Plant J.* **2003**, *35*, 342–349.
64. Herr, A.J.; Jensen, M.B.; Dalmay, T.; Baulcombe, D.C. RNA polymerase IV directs silencing of endogenous DNA. *Science* **2005**, *308*, 118–120.
65. Dalmay, T.; Horsefield, R.; Braunstein, T.H.; Baulcombe, D.C. SDE3 encodes an RNA helicase required for post-transcriptional gene silencing in *Arabidopsis*. *EMBO J.* **2001**, *20*, 2069–2078.
66. Liu, F.; Marquardt, S.; Lister, C.; Swiezewski, S.; Dean, C. Targeted 3' processing of antisense transcripts triggers *Arabidopsis* FLC chromatin silencing. *Science* **2010**, *327*, 94–97.
67. Angel, A.; Song, J.; Dean, C.; Howard, M. A Polycomb-based switch underlying quantitative epigenetic memory. *Nature* **2011**, *476*, 105–108.
68. De Lucia, F.; Crevillen, P.; Jones, A.M.; Greb, T.; Dean, C. A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of FLC during vernalization. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 16831–16836.
69. Helliwell, C.A.; Robertson, M.; Finnegan, E.J.; Buzas, D.M.; Dennis, E.S. Vernalization-repression of *Arabidopsis* FLC requires promoter sequences but not antisense transcripts. *PLoS One* **2011**, *6*, e2151.
70. 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.
71. Heo, J.B.; Sung, S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* **2011**, *331*, 76–79.
72. Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914.
73. Kramerov, D.A.; Vassetzky, N.S. Short retroposons in eukaryotic genomes. *Int. Rev. Cytol.* **2005**, *247*, 165–221.
74. Yukawa, Y.; Felis, M.; Englert, M.; Stojanov, M.; Matousek, J.; Beier, H.; Sugiura, M. Plant 7SL RNA genes belong to type 4 of RNA polymerase III- dependent genes that are composed of mixed promoters. *Plant J.* **2005**, *43*, 97–106.
75. Shin, H.; Shin, H.S.; Chen, R.; Harrison, M.J. Loss of At4 function impacts phosphate distribution between the roots and the shoots during phosphate starvation. *Plant J.* **2006**, *45*, 712–726.
76. Lin, S.I.; Chiang, S.F.; Lin, W.Y.; Chen, J.W.; Tseng, C.Y.; Wu, P.C.; Chiou, T.J. Regulatory network of microRNA399 and PHO2 by systemic signaling. *Plant Physiol.* **2008**, *147*, 732–746.
77. Pant, B.D.; Buhtz, A.; Kehr, J.; Scheible, W.R. MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. *Plant J.* **2008**, *53*, 731–738.
78. Wu, H.J.; Wang, Z.M.; Wang, M.; Wang, X.J. Wide-spread long non-coding RNAs (lncRNAs) as endogenous target mimics (eTMs) for microRNAs in plants. *Plant Physiol.* **2013**, *161*, 1875–1884.

Reprinted from *IJMS*. Cite as: Lin, L.-L.; Wu, C.-C.; Huang, H.-C.; Chen, H.-J.; Hsieh, H.-L.; Juan, H.-F. Identification of MicroRNA 395a in 24-Epibrassinolide-Regulated Root Growth of *Arabidopsis thaliana* Using MicroRNA Arrays. *Int. J. Mol. Sci.* **2013**, *14*, 14270-14286.

Article

Identification of MicroRNA 395a in 24-Epibrassinolide-Regulated Root Growth of *Arabidopsis thaliana* Using MicroRNA Arrays

Li-Ling Lin ¹, Chia-Chi Wu ², Hsuan-Cheng Huang ^{3,*}, Huai-Ju Chen ⁴, Hsu-Liang Hsieh ^{4,*} and Hsueh-Fen Juan ^{1,2,5,*}

¹ Department of Life Science, National Taiwan University, Taipei 106, Taiwan;
E-Mail: f94b43019@ntu.edu.tw

² Institute of Molecular and Cellular Biology, National Taiwan University, Taipei 106, Taiwan;
E-Mail: jcwu0417@hotmail.com

³ Institute of Biomedical Informatics, Center for Systems and Synthetic Biology,
National Yang-Ming University, Taipei 112, Taiwan

⁴ Institute of Plant Biology, National Taiwan University, Taipei 106, Taiwan;
E-Mail: d92621103@ntu.edu.tw

⁵ Graduate Institute of Biomedical Electronic and Bioinformatics, National Taiwan University,
Taipei 106, Taiwan

* Authors to whom correspondence should be addressed; E-Mails: hsuancheng@ym.edu.tw (H.-C.H.); hlhsieh@ntu.edu.tw (H.-L.H.); yukijuan@ntu.edu.tw (H.-F.J.); Tel.: +886-2-2826-7357 (H.-C.H.); +886-2-3366-2540 (H.-L.H.); +886-2-3366-4536 (H.-F.J.); Fax: +886-2-2367-3374 (H.-F.J.).

Received: 20 May 2013; in revised form: 18 June 2013 / Accepted: 28 June 2013 /

Published: 9 July 2013

Abstract: Brassinosteroids (BRs) are endogenous plant hormones and are essential for normal plant growth and development. MicroRNAs (miRNAs) of *Arabidopsis thaliana* are involved in mediating cell proliferation in leaves, stress tolerance, and root development. The specifics of BR mechanisms involving miRNAs are unknown. Using customized miRNA array analysis, we identified miRNAs from *A. thaliana* ecotype Columbia (Col-0) regulated by 24-epibrassinolide (EBR, a highly active BR). We found that miR395a was significantly up-regulated by EBR treatment and validated its expression under these conditions. miR395a was over expressed in leaf veins and root tissues in EBR-treated miR395a promoter::GUS plants. We integrated bioinformatics methods and publicly available DNA microarray data to predict potential targets of miR395a. GUN5—a multifunctional protein involved in plant metabolic functions such as chlorophyll synthesis and the abscisic acid

(ABA) pathway—was identified as a possible target. ABI4 and ABI5, both genes positively regulated by ABA, were down-regulated by EBR treatment. In summary, our results suggest that EBR regulates seedling development and root growth of *A. thaliana* through miR395a by suppressing GUN5 expression and its downstream signal transduction.

Keywords: brassinosteroids; miR395a; root growth; *Arabidopsis thaliana*; microRNA array

1. Introduction

In 1970, a new family of plant hormones, brassins, was reported but later found to be a mixture of multiple compounds [1]. In 1979, another steroid hormone named brassinolide was identified from rape pollen of *Brassica napus* and its structure determined [2,3]. A number of related steroid hormones have since been isolated and collectively classified under the general term brassinosteroids (BRs). To date, more than 50 BR forms including 24-epibrassinolide (EBR) have been identified in a wide variety of plant species [4]. In an attempt to understand how BRs act on plant growth and in what mechanisms they are involved, numerous studies have been and are being conducted in wide-ranging fields, including structural biology, plant physiology, molecular biology, and genetics [5]. Concurrent with biosynthetic research, a large number of BR-deficient or -insensitive mutants have been investigated, among them the *bri1* mutant, which enabled the exploration of the affected gene's role in BR receptor expression [6]. The components of the BRs signal transduction pathway have subsequently been studied in an effort to discriminate the relevant mechanisms [7].

BRs are structurally similar to animal and insect steroid hormones [8] and are products of the isoprenoid biosynthetic pathway; however, they differ in their subsequent metabolism of squalene-2,3-epoxide. While in animals the compound is converted to the precursor of cholesterol and steroid hormones, lanosterol, it is metabolized to cycloartenol in plants, which is the parent compound of all plant sterols [3,5,7].

In 2001, a study comparing *A. thaliana* ecotype Columbia (Col-0) with BR-deficient mutants demonstrated that BR stimulates seed germination by reversing ABA-induced dormancy [9]. A recent study showed that 2 μM of exogenous BR reduced the inhibitory effect of high salt concentrations on seed germination and promoted early stages of seedling growth in *Brassica napus* [10]. Another study indicated that overexpression of the gene *AtDWF4*, essential for BR biosynthesis, was able to overcome ABA-induced inhibition of seed germination [11]. It has been proposed that exogenous BR can regulate other endogenous hormones, and the effect of BR on other plant hormones has been explored in several studies [12].

Plant miRNAs occupy only a small number of functional genes. Currently, 299 *A. thaliana* miRNAs are recorded in miRBASE (release 19) [13]. miRNAs bind to complementary sequences on target mRNAs and in plants mostly act to degrade them [14]. Complementary features of plant miRNAs target their mRNAs by an almost perfect match; most miRNA binding sites exist in coding exons [14,15]. Recently, miRNAs have been reported to be hypersensitive ubiquitous stress regulators: *i.e.*, they function to mediate expression of their target genes when unbalanced nutrient conditions are

encountered [15]. miR399 and miR395 have been identified as being involved in sulfate- and phosphate-starvation responses [16,17].

While recent studies with a structural, genetic, molecular, transcriptomic and proteomic focus have helped elucidate the regulatory mechanisms of the BRs signaling pathway [18–21], the mechanisms of miRNA involvement with BRs are unknown. To gain insights into the mechanism of BR actions at the molecular level, we carried out global screening of miRNAs in *A. thaliana*, which responds rapidly to EBR treatment, successfully investigating potential targets of miRNA and their interaction in plant development.

2. Results and Discussion

2.1. EBR Regulates the Root Development of Arabidopsis

We monitored the morphology of EBR-treated seedlings for root length measurement and germination analysis. It has previously been shown that *Arabidopsis* with different levels of BRs display differences in root development [22,23]. Low concentrations (0.1 and 0.5 nM) of exogenous BRs promoted root elongation in wild-type strains and BR-deficient mutants [22]. In contrast, higher concentrations (1–100 nM) were inhibitory for primary root elongation, instead promoting lateral root formation [23,24]. In this study, we treated *Arabidopsis* with 10 nM EBR. Our results show that primary root length was significantly decreased ($p < 0.01$) and the number of lateral roots was significantly increased ($p < 0.05$) (Figure 1). Expression levels of *BRU6* and *SAURAC-1* genes in *Arabidopsis* were regulated by EBR stimulation (Figure S1), results that corroborate previous studies [21,24]. The development of germination was maintained after EBR treatment, but the root phenotype appeared obviously curved in our germination analysis (Figure 2). Based on these results, we confirm that the root development of *Arabidopsis* can be regulated by EBR at concentrations like those used in our treatments.

Figure 1. 24-epibrassinolide (EBR) regulates root development. (A) The development of lateral roots was enhanced in EBR-treated seedlings. Each plate contained 10 nM EBR or mock solution (control). The red line represents initial length before treatment; (B) Differences in primary root length between day 5 and day 11. Primary root length was significantly shorter under EBR treatment; and (C) Number of lateral roots on day 11. Lateral root number was significantly increased in the EBR supplement plate. Representative data from three independent experiments are presented as mean \pm SD. * $p < 0.05$, ** $p < 0.01$ vs. control treatment values.

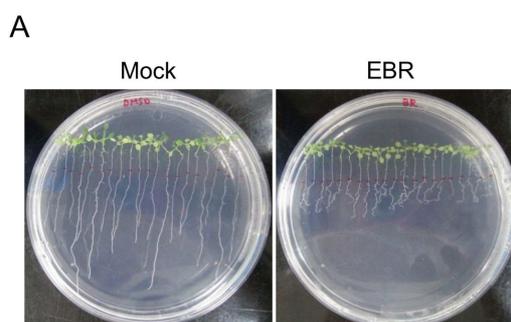


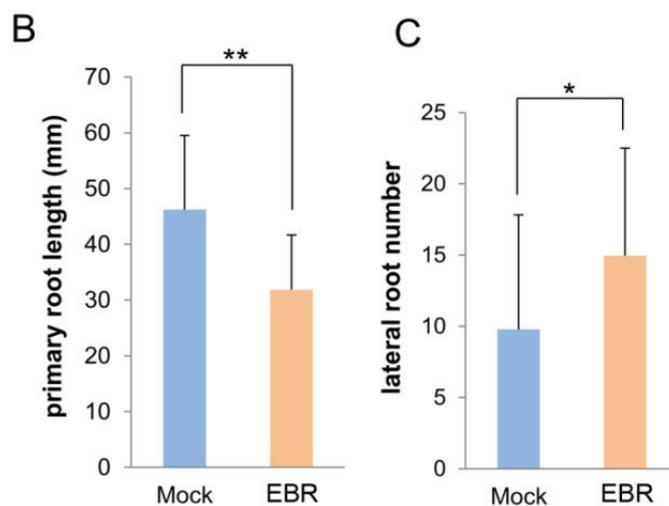
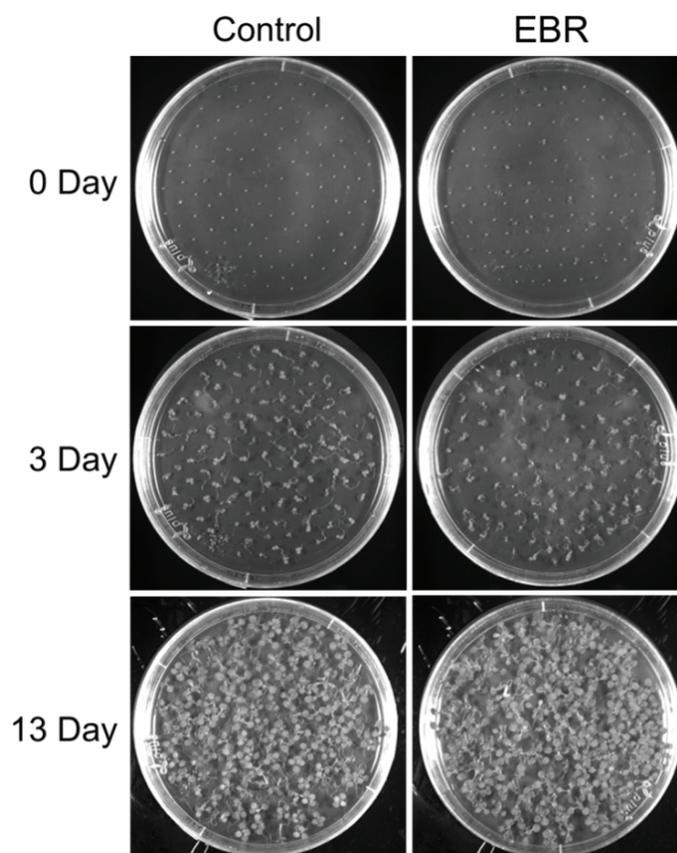
Figure 1. *Cont.*

Figure 2. EBR has no effect on germination. There was no significant difference in germination rate between control and EBR-treated seeds; however roots were shorter and more strongly curved in EBR-supplemented plates. Germination was recorded on day 3 and day 13 after imbibition.

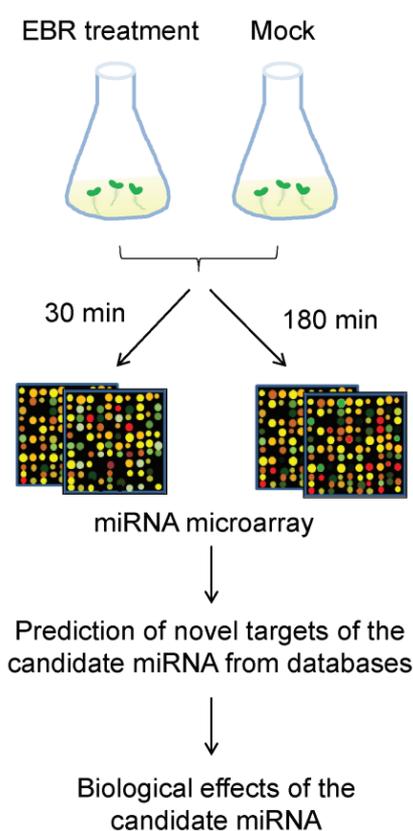


2.2. Identification of EBR-Regulated miRNAs in Arabidopsis

To explore the role of miRNAs in BR-mediated pathways, we analyzed differences in miRNA profiles between control (mock solution) and EBR treatments from customized miRNA microarrays.

Seeds were separately cultured under exogenous 10 nM EBR treatments for 30 (EBR30) or 180 (EBR180) minutes (Figure 3), and total RNA of all seedlings was extracted after seven days of growth. A scatter plot of probe intensities of duplicate microarrays shows no difference ($R^2 > 0.99$) among the duplicates, therefore validating the consistency of our microarray experiments (Figure S2).

Figure 3. Schematic flowchart of experimental design. After *Arabidopsis* Col-0 seeds had grown in MS liquid medium for 7 days, seedlings were treated for 30 or 180 min with MS medium supplemented with EBR or mock solution (DMSO), followed by RNA extraction, labeling, and hybridization. Candidate miRNAs were predicted using miRU, WMD3, and psRNATarget databases. The roles of candidate miRNAs in EBR-treated seedlings were investigated by further experiments as described in the text.

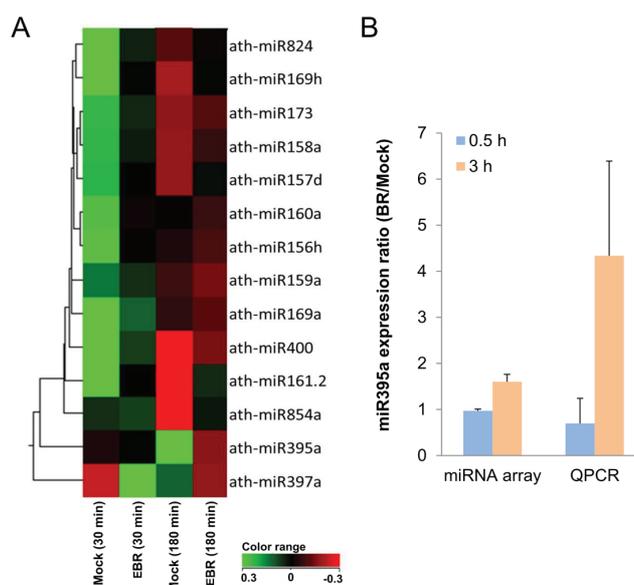


The expressed fold changes of miRNAs from EBR-treated seedlings were normalized to a DMSO-treated control. Fourteen miRNAs with significantly different expression ratios ($p < 0.05$) (Table 1) from both EBR30 and EBR180 treatments were selected for hierarchical clustering (Figure 4A). Among these, 11 miRNAs were up-regulated and three down-regulated for EBR30, and six up-regulated and eight down-regulated for EBR180. Of these, miR395a exhibited the highest fold change (1.6-fold) from microarray data at EBR180. Similarly, EBR-treated seedlings showed a higher expression of miR395a (4.3-fold) than control seedlings in real-time PCR (qPCR) analysis (Figure 4B). These results validate that miR395a expression is up-regulated by EBR.

Table 1. EBR regulates the expression of several microRNAs (miRNAs).

| microRNAs | Fold change (EBR/DMSO) | |
|--------------|------------------------|---------|
| | 30 min | 180 min |
| ath-miR824 | 1.27 | 0.93 |
| ath-miR169h | 1.37 | 0.86 |
| ath-miR173 | 1.15 | 0.95 |
| ath-miR158a | 1.16 | 0.92 |
| ath-miR157d | 1.18 | 0.87 |
| ath-miR160a | 1.25 | 1.06 |
| ath-miR156h | 1.25 | 1.04 |
| ath-miR159a | 1.07 | 1.05 |
| ath-miR169a | 1.26 | 1.04 |
| ath-miR400 | 1.21 | 0.89 |
| ath-miR161.2 | 1.31 | 0.71 |
| ath-miR854a | 0.98 | 0.78 |
| ath-miR395a | 0.97 | 1.60 |
| ath-miR397a | 0.53 | 1.25 |

Figure 4. EBR up-regulates miR395a in miRNA microarray analysis. **(A)** Hierarchical clustering of selected miRNAs expression regulated by EBR. Seedlings were treated with EBR for 30 or 180 min. miRNA expression was assessed with miRNA microarrays. Fourteen miRNAs had significantly different expression levels after EBR treatment ($p < 0.05$) and were further analyzed by a hierarchical clustering algorithm; and **(B)** Fold changes of miR395a in miRNA microarrays and qPCR analysis. miR395a was up-regulated after EBR treatment for 180 min. snoR85 was used as an internal control for normalization.



2.3. GUN5 Is a Novel Target of miR395a

To investigate the role of miR395a in *Arabidopsis* development, we explored its potential targets by the bioinformatics approach of complementary base-pairing. We obtained potential target genes of

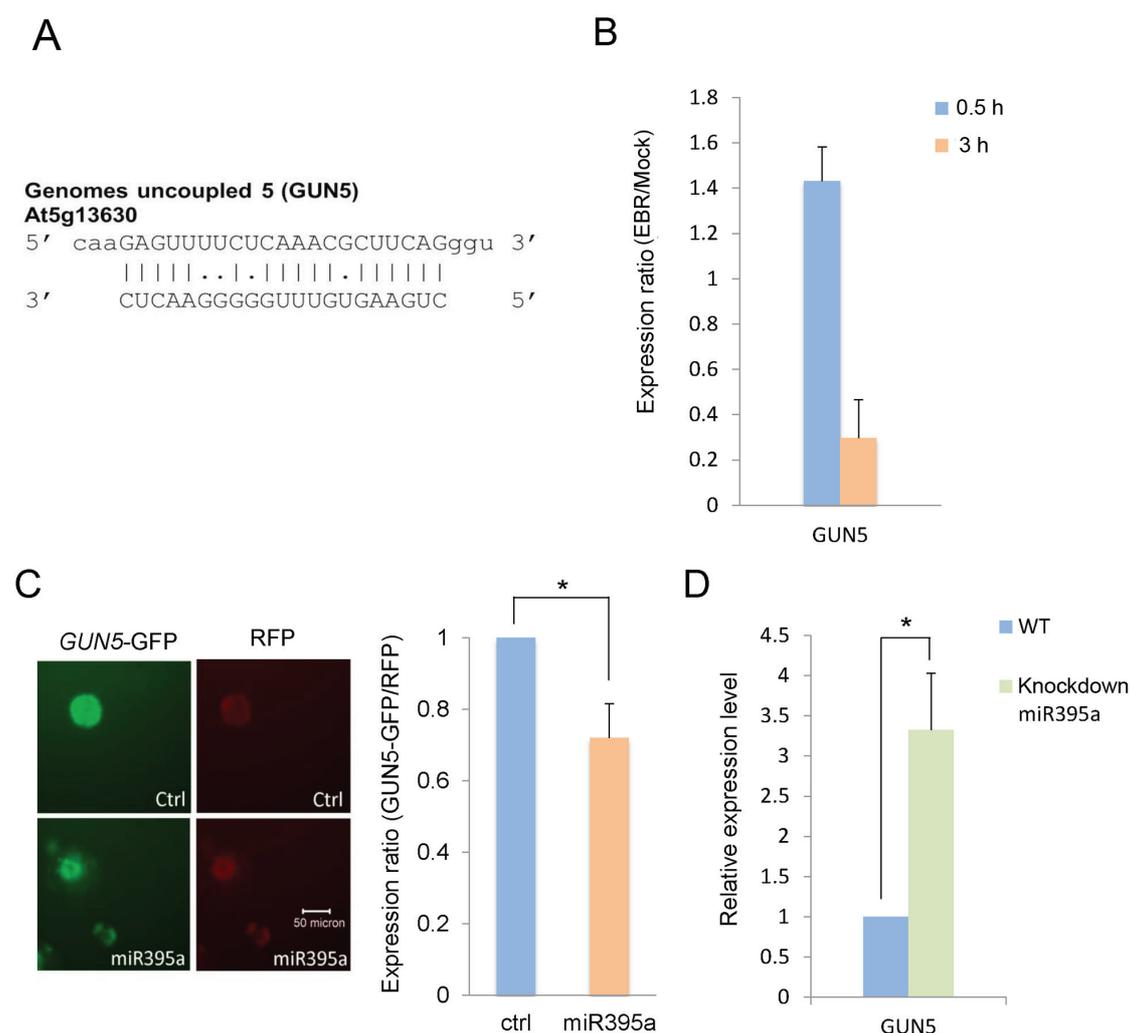
miR395a from three databases: miRU [25] (Table S1), which integrates most known plant miRNAs and target genes and can be employed to search potential plant miRNA targets and target-sites within mismatch and miRNA conservation thresholds in target recognition; WMD3 [26] (Table S2), which uses principles of artificial miRNA design to mimic natural plant miRNAs; and psRNATarget [27] (Table S3), which provides scoring schemata and evaluates target-site accessibility in miRNA target recognition. The *Arabidopsis* TAIR9 cDNA library and default parameters were used in predicting target sequences from these databases. Among the candidate targets, *GUN5* (At5g13630), the CHLH subunit of Mg-chelatase, has been reported as possibly involved in the abscisic acid (ABA) pathway [28,29]. The interaction between *GUN5* and miR395a is however still unknown. ABA, a plant hormone, mediates the development of plants by inhibiting seedling germination, maintaining primary root growth and reducing lateral root density [30–32]. We therefore decided to focus our investigation on whether *GUN5* is a target gene of miR395a; the target sequence of *GUN5* is shown in Figure 5A. We furthermore analyzed *GUN5* expressions in EBR- and DMSO-treated seedlings by qPCR. As shown in Figure 5B, these gene expressions were decreased in EBR-treated seedlings compared with DMSO-treated seedlings, indicating that *GUN5* might be associated with miR395a.

To further validate the interaction between miR395a and *GUN5* in *Arabidopsis*, we constructed pRTL2-miR395a and smGFP/pRTL2-*GUN5*. We co-transformed the plasmids into PSB-D cells and detected the fluorescence intensity of smGFP/pRTL2-*GUN5* and the internal control RFP. We also analyzed the levels of GFP-*GUN5* expression in transformed PSB-D cells. As shown in Figure 5C, lower *GUN5* expression was detected in PSB-D cells with miR395a overexpression than in PSB-D cells with control vectors. In plants with miR395a knockout, the gene expression of *GUN5* was also significantly increased (Figure 5D). These results suggest that *GUN5* is a target gene of miR395a.

2.4. Distribution of miR395a in Vascular Bundles, Leaf Veins and Roots of *Arabidopsis*

GUN5 has been reported as being instrumental in leaf greening [28], and a decrease in chlorophyll accumulation has been found in *gun5* mutants [33]. BR is also a crucial factor in the regulation of chloroplast development, playing a role as a negative regulator [34]. Based on these similarities, we explored the role of miR395a in EBR-treated *Arabidopsis*, following *GUN5* suppression. To clarify the expression sites of miR395a, we used T2 seeds from miR395a promoter::GUS plants to examine the miR395a expression pattern under EBR treatment (expression levels of miR395a promoter::GUS were also found to be up-regulated under EBR treatment). In leaf and root development, miR395a specifically was concentrated in leaf veins of the cotyledon (Figure 6A) and in partial vascular bundles of roots (Figure 6B) and was also distributed in chloroplasts around leaf veins (Figure S3). Figure 6B also shows that root diameter under EBR treatment was larger under mock treatment, indicating that miR395a might regulate root development through EBR signaling.

Figure 5. GUN5 is a target gene of miR395a. **(A)** The prediction was based on complementary base-pairing between miR395a and mRNA. Putative target genes for miR395a were predicted using web-based databases (miRU, psRNATarget, WMD3) and simultaneous comparisons to gene microarray data. The putative target genes of miR395a were down-regulated by EBR treatment; **(B)** Expression levels of GUN5 were analyzed by Q-PCR and normalized to 18S rRNA; **(C)** Fluorescence assay of miR395a and potential target GUN5 in the *Arabidopsis* PSB-D cell line. The fluorescent expression levels of GUN5 revealed significant down-regulation by miR395a; and **(D)** Gene expression of GUN5 measured by qPCR in miR395a knockout plants. GUN5 expression was significantly down-regulated by miR395a. * $p < 0.05$.



A previous study has shown that the transcription factors *ABI4* and *ABI5* are positive regulators of ABA signaling and can be considered downstream genes of *GUN5* [19]. We found that these genes could be suppressed by EBR treatment (Figure 7A). It has also been reported that a mutation of *ABI4* can increase the number of lateral roots [35] and that *ABI5* activity inhibits seedling germination and promotes primary root growth [36,37]. These results lead us to propose that EBR may maintain seedling germination, inhibit primary root growth, and increase the number of lateral roots through regulation of miR395a effects on *ABI4* and *ABI5* via *GUN5* (Figure 7B).

Figure 6. Histochemical GUS staining of miR395a expression in *A. thaliana*. Expression patterns of miR395a promoter::GUS plants in (A) leaf and (B) root tissue. After growing for 7 days, seedlings were grown under EBR treatment (10 nM EBR) or mock control for 3 h. The arrow indicates a high concentration of miR395a distributed in the vascular bundle compared with the mock treatment.

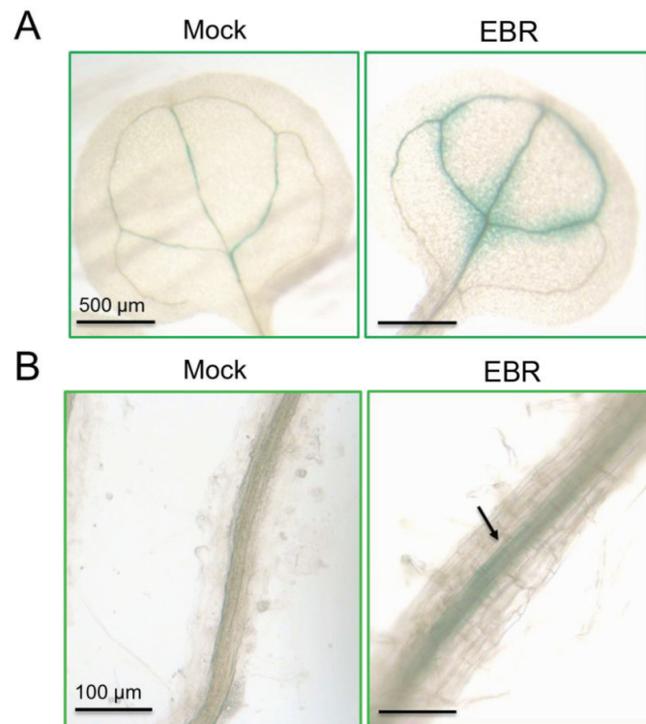
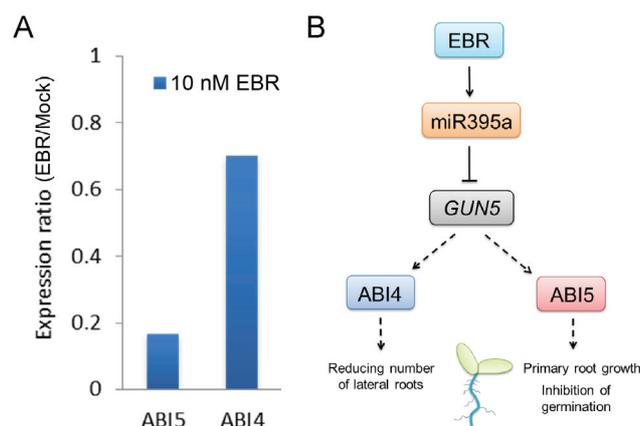


Figure 7. Expression of GUN5 downstream genes in the ABA pathway. (A) Relative expression ratio of GUN5 downstream genes; and (B) Diagram of GUN5-dependent ABA pathway.



2.5. Discussion

BRs can induce a wide range of physiological effects in cell elongation and division, photosynthesis, photomorphogenesis, flowering, senescence, seed germination, root development, male fertility, and abiotic and biotic stress resistance [8,19,38,39]. They are active at low concentrations throughout the plant kingdom and widely distributed in plants at varying levels of complexity [7,8]. Higher

concentrations of BRs are seen primarily in young growing tissues rather than in mature tissues [7]. BR-insensitive mutants in *Arabidopsis* exhibit phenotypes such as dwarfism, dark-green leaves, reduced fertility, prolonged life span, and abnormal skotomorphogenesis [3,8,39].

miRNAs also provide examples of regulation at various stages of plant development. Some miRNAs, such as miR159 and miR160, play roles during early development stages including seed germination. During post-germination stages, miR156 and miR172 mediate the emergence of vegetative leaves, a stage of transition to autotrophic growth [40,41]. miRNA-mediated signaling is also involved in the development of various tissues; several miRNA families such as miR160, miR164, miR167, and miR390 have been demonstrated to be involved in root cap formation and lateral root development [42]. However, the relationship between BRs and miRNAs is unknown.

In the present study, we screened different miRNA expression profiles in *Arabidopsis* with 10 nM EBR for periods of 30 and 180 min. The results show that in both cases the expression of miR395a was significantly up-regulated by EBR (Figure 4). Recent studies have indicated that miR395a is up-regulated in roots and expressed in cortex, phloem companion cells and epidermis under low-sulfur conditions [43]. miR395a is mostly expressed in roots when playing a role in homeostasis regulation [43]. As discussed above, the morphology of BR-treated plants showed a decrease in taproot length and an increase in lateral root formation [23]. These effects might be caused by miR395a-involved mechanisms, and miR395a might be among the factors affecting root growth and development. miR395 and miR397 play roles in sulfate metabolism and copper homeostasis, respectively [43–45]. The function of these miRNAs lies mostly in adapting to unbalanced conditions, which implies that the experimental concentration of 10 nM EBR might have been in excess of physiological levels and affected the homeostasis of the seedlings.

In addition to miR395a, several significantly different expressions of miRNAs may have potential functions relevant to BR-treated seedlings:

(a) miR824 was down-regulated in BR-treated seedlings. It is involved in stomatal development by targeting *AGL16*, through which it causes a decrease in the number of stomata [46]. This suggests an increased stomata number in BR-treated plants [46,47]. Proper amounts and distributions of stomata are essential for successful gas exchange [46], and so an increase in the stomata number might therefore contribute to greater metabolic efficiency in plants.

(b) miR169a, which can regulate adaptive responses to nutrient deprivation [48], was also up-regulated in our miRNA profiles. This suggests that miR169a might have acted in this capacity of adaptation to environmental change when we supplied exogenous BR.

(c) miR160 mediates agravitropic roots with disorganized root caps as well as lateral root development, primary root growth, floral organs in carpels, and germination [40,42,49]. Our miRNA arrays indicated that the up-regulation of miR160a might have resulted in the expression of the phenotype observed in the present study. Since the lateral root formation caused by miR160 was similar to the morphology of BR-treated seedlings, we suspect miR160 might play an important role in lateral root development in BR-supplied plants.

(d) miR156 has been shown in recent studies to increase leaf initiation, phase change, floral induction, and phosphate homeostasis, to decrease apical dominance, and to delay flowering

time [40,42,49]. As suspected, miR156h was up-regulated in the miRNA profiles, suggesting a crucial function in promoting growth and development.

(e) miR159 regulates germination, anthers, and flowering time by targeting the MYB transcription factor [49,50]. Overexpression of miR159 results in male sterility and delayed flowering time.

To further explore the role of miR395a, we predicted target genes of miR395a from several different databases and identified *GUN5* as a novel potential target of miR395a in *Arabidopsis*. We were able to show that the expression of *GUN5* was suppressible by miR395a (Figure 5C,D). Similar to the phenotype of the *gun4* mutant, the *gun5* mutant showed a decrease in chlorophyll accumulation, while the *gun4gun5* double mutant displayed the even more noticeable characteristic of albino leaves [28]. *GUN2/3/4/5* are also involved in communicating along plastid-to-nucleus retrograde signaling pathways with Mg-ProtoIX acting as a signaling molecule between chloroplast and nucleus [51]. In contrast, BR inhibits chloroplast development [34], and down-regulates *GUN5* expression (Figure 5B). After EBR treatment, miR395a was up-regulated and strongly expressed in cotyledon leaf veins and root vascular bundles (Figures 4B and 6). These results suggest that BR might enhance miR395a to suppress *GUN5* expression during plant development. However, the exact relationship between BR, miR395a and *GUN5* remains unknown. Recent studies have indicated that *ABI4* is a downstream regulator between chloroplast and nucleus that connects to ABA via retrograde signaling [52]. Hence, *GUN5* is likely to play a role in chlorophyll synthesis by connecting ABA to different pathways [33]. Additionally, we found that ABA regulatory genes were suppressed by EBR (Figure 7). These outcomes indicate that the interaction between miR395a and *GUN5* may regulate chlorophyll synthesis through the ABA signaling pathway.

3. Experimental Section

3.1. Plant Material and Growth Conditions

A. thaliana ecotype Columbia (Col-0) was used as plant material in this study. Before sowing, seeds were surface sterilized by rinsing them in 1% bleach (sodium hypochlorite) with 0.5% Tween 20 and vortexing for eight minutes, washed 5–6 times and then cold-treated for two days at 4 °C under dark conditions. Plants were sown in pots (containing 50% vermiculite and 50% soil mixture), medium, or agar plates and kept in a growth chamber operating at photoperiod conditions of 14 h light and 10 h darkness at 22 °C after stratification.

3.2. Germination Assay

Seeds were grown on half-strength Murashige and Skoog medium (1/2 MS medium; Duchefa Biochemie B.V., Haarlem, Netherlands) with 1.5% (w/v) sucrose (Sigma-Aldrich Co. LLC., Dorset, UK) and 0.8% (w/v) plant agar (Sigma-Aldrich, St. Louis, MO, USA) containing 10 nM 24-epibrassinolide (EBR, a highly active BR; Sigma-Aldrich) or mock solution (dimethyl sulfoxide, DMSO). Images were taken at zero, three, and thirteen days after sowing.

3.3. Root Growth Assay

For root elongation analysis, seedlings were grown vertically on 1/2 MS medium with 1.5% sucrose and 0.8% plant agar for five days after germination. Seedlings were then transferred to new plates containing MS medium supplemented with 10 nM EBR or mock solution for another six days, with images taken after five and 11 days. Differences in primary root length between the two images were measured, and number of lateral roots was calculated after 11 days, using ImageJ software [53].

3.4. MicroRNA Microarray Hybridization and Analysis

For the miRNA array experiments, seedlings were grown in 1/2 MS medium with 1.5% sucrose. After stratification, seeds were transferred into 50 mL flasks with 10 mL liquid medium and incubated for seven days at 50 rpm and 22 °C under continuous light conditions. Seedlings were then treated with medium containing 10 nM EBR or mock solution for 0.5 and 3 h, respectively. Total RNAs were extracted from complete frozen seedlings using TRIzol[®] Reagent (Invitrogen, Carlsbad, CA, USA), and RNA purity was confirmed by spectrophotometry (A_{260}/A_{280} ratio) and capillary electrophoresis (Agilent 2100 Bioanalyzer, Agilent Technologies, Palo Alto, CA, USA). Then, 100 ng total RNAs of each sample were prepared for labeling with Cyanine 3-pCp. RNA processing and hybridization were performed using miRBASE V14 arrays (Agilent Technologies, Palo Alto, CA, USA) according to the manufacturer's protocol; this version contains 161 *Arabidopsis thaliana* miRNA genes. Each plex on these customized eight-plex microarrays contained duplicate or triplicate probes for each miRNA, with 20 replicates for each probe. Microarray analysis was carried out in GeneSpring GX version 11 (Agilent, city, state, country). The data (covering the four conditions) were classified into groups by the averages of duplicates, and the median of all samples was set as a baseline. Differences in miRNA expression were tested using a one-way ANOVA. miRNAs with significant differences ($p < 0.05$) between mock control and EBR-treated seedlings were selected for clustering and those with the highest fold change were subjected to further analysis. Array data were submitted to the GEO database (series record number GSE46377).

3.5. Real-Time RT-PCR

All cDNA synthesis was carried out on total RNAs using the RevertAid H Minus Reverse Transcriptase Kit (Fermentas, Maryland, NY, USA) according to the manufacturer's instructions. Reactions for expression analysis of Col-0 genes treated with EBR or mock control were performed in triplicate and monitored using the iQ5 Real-time PCR Detection System (Bio-Rad, Philadelphia, PA, USA). Investigated genes and corresponding primers are listed in Table S5. Relative abundance of transcripts was normalized to the constitutive expression levels of 18S rRNA (At3g41768). For miRNA expression analysis, specific miRNAs were measured with TaqMan microRNA assays (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions. All reactions were run in triplicate and snoR85 was used as the internal control for normalization.

3.6. Prediction of Novel miRNA Target Genes

We obtained *Arabidopsis* miRNA sequences from the miRBASE dataset [13]. The mature sequences of all miRNA genes were used in this study. For predictions of miRNA targets, the programs miRU [25], psRNATarget [27] and WMD3 [26] were employed. The *Arabidopsis thaliana* full genome (TAIR9) was selected in the psRNATarget and WMD3 databases (other parameters were left at default).

3.7. Vector Construction

The genomic DNA of *Arabidopsis* leaves was extracted using QuickExtract™ Plant DNA Extraction Solution (Epicentre, Madison, WI, USA), following the manufacturer's instructions. Different vectors were used for specific purposes. For validating the interaction of miRNA and genes, vector pRTL2-mGFP (Biovector Co., LTD, Beijing, China) was used to construct miR395a and control. Vector pRTL2 was used to delete the mGFP gene from pRTL2-mGFP via the restriction enzymes *EcoR* I and *Xba* I of miR395a. Vector 326-RFP is an internal control for cell numbers in *Arabidopsis* cell lines. For cloning *GUN5* in translational fusion, the coding region of *GUN5* was inserted into smGFP/pRTL2 using the restriction site *Spe* I. Transcriptional fusions were created with *GUN5* and smGFP for activity analysis in protoplast system. Protoplasts were prepared following the protocol of Miao and Jiang [54]. Vector pZP221 was used in transgenic plant construction for the miR395a-overexpressing line; the insertion containing miR395a with CaMV 35S promoter and terminator in the pRTL2-miR395a vector was cloned into the *Pst* I site of pZP221. The binary vector pBI101 with the reporter gene β -glucuronidase (GUS) was used for promoter activity analysis of the miR395a promoter line. Transcriptional fusions for analysis of promoter activity in plants were generated using the miR395a promoter with built-in cloning sites *Sal* I and *Xba* I, and GUS.

The miR395a knockout line was purchased from the *Arabidopsis* Information Resource (TAIR). Transformations were performed with the Gene Pulser Xcell™ Electroporation System (Bio-Rad, Richmond, CA, USA) at pulse settings of 130 V at 1000 μ F.

3.8. Fluorescence Assay for Validating miR395a and GUN5

For fluorescence assays, 200 μ L of transformed protoplast cells were transferred to black opaque 96-well microplates (Greiner Bio-One, Wemmel, Belgium) and immediately measured in a multimode microplate reader (FlexStation 3 microplate reader; MDS Analytical Technologies, Sunnyvale, CA, USA). Excitation and emission wavelengths were 488 and 508 nm for green light (smGFP) and 558 and 583 nm for red light (DsRed).

3.9. Detection of the Expression Pattern of miR395a in *Arabidopsis thaliana*

When *Arabidopsis* plants had grown for four to six weeks, the first bolt was cut to induce the emergence of further bolts. About one week after clipping, plants containing numerous unopened floral buds were immersed in a buffer of *Agrobacterium tumefaciens*. The buffer was prepared as follows: Transformed *A. tumefaciens* cells were grown at 28 °C and 180 rpm shaking in LB medium with the appropriate antibiotics. A 10-mL pre-culture was grown for two days and then transferred to the

200-mL main culture. This was incubated until an OD600 value of 0.8 was reached, and then was centrifuged at $4000\times g$ for 15 min at 4 °C. The supernatant was discarded and sucrose and Silwet L-77 (Sigma, St. Louis, MO, USA) were added to the culture to obtain final concentrations of 5% and 0.05%, respectively.

Pots of plants were inverted and the inflorescence shoots dipped into suspension, then laid on a flat plastic surface and left covered and dark for the next 24 h, and afterwards returned to normal growing conditions. T1 plants were grown from selected transformants. The transgenic character of plants was confirmed by PCR and GUS staining.

When T2 transgenic lines were obtained, we used a GUS staining kit (GUSS; Sigma) to detect the expression pattern of miR395a. Seedlings were incubated at room temperature for 45 min with a fixation solution, which was then poured off. They were washed three times with wash solution for one minute, then left to incubate with staining solution for up to 24 h at 37 °C. Finally, the chlorophyll was removed by destaining the samples with ethanol. Tissues were stored in ethanol. Manufacturer's instructions were followed in performing the assay.

3.10. Statistical Analysis

Data were represented as mean \pm standard deviation (SD). Differences between independent groups were analyzed using a two-tailed Student's *t*-test. MicroRNA microarrays for miRNA expression were analyzed using a one-way ANOVA (GeneSpring 7.3.1, Agilent Technologies, Palo Alto, CA, USA). A *p* value < 0.05 was taken to indicate statistical significance.

4. Conclusions

Our results show that miR395a was significantly up-regulated by EBR in *Arabidopsis*, was expressed more strongly in leaf veins and roots of EBR-treated miR395a promoter::GUS plants, and targeted GUN5 with the effect of suppressing its expression. EBR was able to suppress GUN5 downstream genes to regulate seedling germination and the formation of primary and lateral roots. These results suggest that the reduced amount of chlorophyll in leaf veins and root growth of *Arabidopsis* might be attributable to the interaction between miR395a and GUN5. This study provides new insights into the function of miRNAs that will be useful in further research into the roles miRNAs play in the molecular mechanisms of plant development.

Acknowledgments

This work was supported by the National Science Council, Taiwan (NSC 99-2621-B-002-005-MY3, NSC 99-2621-B-010-001-MY3, NSC 101-2120-M-002-010) and the National Taiwan University Cutting-Edge Steering Research Project (NTU-CESRP-102R7602C3).

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Mitohell, J.W.; Mandava, N.; Worley, J.F.; Plimner, J.R.; Smith, M.V. Brassins—a new family of plant hormone from rape pollen. *Nature* **1970**, *225*, 1065–1066.
2. Grove, M.D.; Spencer, G.F.; Rohwedder, W.K. Brassinolide, a plant growth-promoting steroid isolated. *Nature* **1979**, *281*, 216–217.
3. Altmann, T. Molecular physiology of brassinosteroids revealed by the analysis of mutants. *Planta* **1999**, *208*, 1–11.
4. Hategan, L.; Godza, B.; Russinova, M. Regulation of Brassinosteroid Metabolism. In *Brassinosteroids: A Class of Plant Hormone*; Hayat, S., Ahmad, A., Eds.; Springer: Heidelberg, Germany, 2011; p. 67.
5. Yokota, T. The structure, biosynthesis and function of brassinosteroids. *Trends Plant Sci.* **1997**, *2*, 137–143.
6. Clouse, S.D.; Langford, M.; McMorris, T.C. A brassinosteroid-insensitive mutant in *Arabidopsis thaliana* exhibits multiple defects in growth and development. *Plant Physiol.* **1996**, *111*, 671–678.
7. Clouse, S.D.; Sasse, J.M. Brassinosteroids: Essential regulators of plant growth and development. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1998**, *49*, 427–251.
8. Clouse, S.D. Brassinosteroid signal transduction: Clarifying the pathway from ligand perception to gene expression. *Mol. Cell* **2002**, *10*, 973–982.
9. Steber, C.M.; McCourt, P. A role for brassinosteroids in germination in *Arabidopsis*. *Plant Physiol.* **2001**, *125*, 763–769.
10. Kagale, S.; Divi, U.K.; Krochko, J.E.; Keller, W.A.; Krishna, P. Brassinosteroid confers tolerance in *Arabidopsis thaliana* and *Brassica napus* to a range of abiotic stresses. *Planta* **2007**, *225*, 353–364.
11. Divi, U.K.; Krishna, P. Overexpression of the brassinosteroid biosynthetic gene AtDWF4 in *Arabidopsis* seeds overcomes abscisic acid-induced inhibition of germination and increases cold tolerance in transgenic seedlings. *J. Plant Growth Regul.* **2010**, *29*, 385–393.
12. Gomes, M.M.A. Physiological Effects Regulated to Brassinosteroid Application in Plants. In *Brassinosteroids: A class of Plant Hormone*; Hayat, S., Ahmad, A., Eds.; Springer: Heidelberg, Germany, 2011; pp. 193–204.
13. Kozomara, A.; Griffiths-Jones, S. miRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **2011**, *39*, D152–D157.
14. Voinnet, O. Origin, biogenesis, and activity of plant microRNAs. *Cell* **2009**, *136*, 669–687.
15. Bartel, D.P. microRNAs: Target recognition and regulatory functions. *Cell* **2009**, *136*, 215–233.
16. Liu, H.H.; Tian, X.; Li, Y.J.; Wu, C.A.; Zheng, C.C. Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. *RNA* **2008**, *14*, 836–843.
17. Fujii, H.; Chiou, T.J.; Lin, S.I.; Aung, K.; Zhu, J.K. A miRNA involved in phosphate-starvation response in *Arabidopsis*. *Curr. Biol.* **2005**, *15*, 2038–2043.
18. Deng, Z.; Zhang, X.; Tang, W.; Osés-Prieto, J.A.; Suzuki, N.; Gendron, J.M.; Chen, H.; Guan, S.; Chalkley, R.J.; Peterman, T.K.; *et al.* A proteomics study of brassinosteroid response in *Arabidopsis*. *Mol. Cell. Proteomics* **2007**, *6*, 2058–2071.
19. Tang, W.; Deng, Z.; Osés-Prieto, J.A.; Suzuki, N.; Zhu, S.; Zhang, X.; Burlingame, A.L.; Wang, Z.-Y. Proteomics studies of brassinosteroid signal transduction using prefractionation and two-dimensional DIGE. *Mol. Cell. Proteomics* **2008**, *7*, 728–738.

20. Nemhauser, J.L.; Hong, F.; Chory, J. Different plant hormones regulate similar processes through largely nonoverlapping transcriptional responses. *Cell* **2006**, *126*, 467–475.
21. Goda, H.; Shimada, Y.; Asami, T.; Fujioka, S.; Yoshida, S. Microarray analysis of brassinosteroid-regulated genes in *Arabidopsis*. *Plant Physiol.* **2002**, *130*, 1319–1334.
22. Mussig, C.; Shin, G.H.; Altmann, T. Brassinosteroids promote root growth in *Arabidopsis*. *Plant Physiol.* **2003**, *133*, 1261–1271.
23. Bao, F.; Shen, J.; Brady, S.R.; Muday, G.K.; Asami, T.; Yang, Z. Brassinosteroids interact with auxin to promote lateral root development in *Arabidopsis*. *Plant Physiol.* **2004**, *134*, 1624–1631.
24. Fukaki, H.; Tasaka, M. Hormone interactions during lateral root formation. *Plant Mol. Biol.* **2009**, *69*, 437–499.
25. Zhang, Y. miRU: An automated plant miRNA target prediction server. *Nucleic Acids Res.* **2005**, *33*, W701–W704.
26. Schwab, R.; Ossowski, S.; Riester, M.; Warthmann, N.; Weigel, D. Highly specific gene silencing by artificial microRNAs in *Arabidopsis*. *Plant Cell* **2006**, *18*, 1121–1133.
27. Dai, X.; Zhao, P.X. psRNATarget: A plant small RNA target analysis server. *Nucleic Acids Res.* **2011**, *39*, W155–W159.
28. Mochizuki, N.; Brusslan, J.A.; Larkin, R.; Nagatani, A.; Chory, J. *Arabidopsis* genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. *Proc. Natl. Acad. Sci. USA* **2000**, *98*, 2053–2058.
29. Rodermel, S.; Park, S. Pathways of intracellular communication: Tetrapyrroles and plastid-to-nucleus signaling. *BioEssays* **2003**, *25*, 631–636.
30. Finkelstein, R.R.; Gampala, S.S.; Rock, C.D. Abscisic acid signaling in seeds and seedlings. *Plant Cell* **2002**, *14*, S15–S45.
31. Zhang, H.; Han, W.; de Smet, I.; Talboys, P.; Loya, R.; Hassan, A.; Rong, H.; Jurgens, G.; Paul Knox, J.; Wang, M.H. ABA promotes quiescence of the quiescent centre and suppresses stem cell differentiation in the *Arabidopsis* primary root meristem. *Plant J.* **2010**, *64*, 764–774.
32. De Smet, I.; Signora, L.; Beeckman, T.; Inze, D.; Foyer, C.H.; Zhang, H. An abscisic acid-sensitive checkpoint in lateral root development of *Arabidopsis*. *Plant J.* **2003**, *33*, 543–555.
33. Jenks, M.A.; Wood, A.J. *Genes for Plant Abiotic Stress*, 1st ed.; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2009.
34. Nakano, T.; Kimura, T.; Kaneko, I.; Nagata, N.; Matsuyama, T.; Asami, T.; Yoshida, S. Molecular mechanism of chloroplast development regulated by plant hormones. *RIKEN Rev.* **2001**, *41*, 86–87.
35. Shkolnik-Inbar, D.; Bar-Zvi, D. ABI4 mediates abscisic acid and cytokinin inhibition of lateral root formation by reducing polar auxin transport in *Arabidopsis*. *Plant Cell* **2010**, *22*, 3560–3573.
36. Lopez-Molina, L.; Mongrand, S.; McLachlin, D.T.; Chait, B.T.; Chua, N.H. ABI5 acts downstream of ABI3 to execute an ABA-dependent growth arrest during germination. *Plant J.* **2002**, *32*, 317–328.
37. Miura, K.; Lee, J.; Jin, J.B.; Yoo, C.Y.; Miura, T.; Hasegawa, P.M. Sumoylation of ABI5 by the *Arabidopsis* SUMO E3 ligase SIZ1 negatively regulates abscisic acid signaling. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5418–5423.

38. Ye, Q.; Zhu, W.; Li, L.; Zhang, S.; Yin, Y.; Ma, H.; Wang, X. Brassinosteroids control male fertility by regulating the expression of key genes involved in *Arabidopsis* anther and pollen development. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 6100–6105.
39. Clouse, S.D. Molecular genetic analysis of brassinosteroid action. *Physiol. Plant* **1997**, *100*, 702–709.
40. Nonogaki, H. microRNA gene regulation cascades during early stages of plant development. *Plant Cell Physiol.* **2010**, *51*, 1840–1846.
41. Poethig, R.S. Small RNAs and developmental timing in plants. *Curr. Opin. Genet. Dev.* **2009**, *19*, 374–378.
42. Meng, Y.; Ma, X.; Chen, D.; Wu, P.; Chen, M. microRNA-mediated signaling involved in plant root development. *Biochem. Biophys. Res. Commun.* **2010**, *393*, 345–349.
43. Kawashima, C.G.; Yoshimoto, N.; Maruyama-Nakashita, A.; Tsuchiya, Y.N.; Saito, K.; Takahashi, H.; Dalmay, T. Sulphur starvation induces the expression of microRNA-395 and one of its target genes but in different cell types. *Plant J.* **2009**, *57*, 313–321.
44. Liang, G.; Yang, F.; Yu, D. microRNA395 mediates regulation of sulfate accumulation and allocation in *Arabidopsis thaliana*. *Plant J.* **2010**, *62*, 1046–1057.
45. Yamasaki, H.; Hayashi, M.; Fukazawa, M.; Kobayashi, Y.; Shikanai, T. SQUAMOSA promoter binding protein-like7 is a central regulator for copper homeostasis in *Arabidopsis*. *Plant Cell* **2009**, *21*, 347–361.
46. Kutter, C.; Schob, H.; Stadler, M.; Meins, F., Jr.; Si-Ammour, A. microRNA-mediated regulation of stomatal development in *Arabidopsis*. *Plant Cell* **2007**, *19*, 2417–2429.
47. Petti, F.B.; Liguori, A.; Ippoliti, F. Study on cytokines IL-2, IL-6, IL-10 in patients of chronic allergic rhinitis treated with acupuncture. *J. Tradit. Chin. Med.* **2002**, *22*, 104–111.
48. Zhao, M.; Ding, H.; Zhu, J.-K.; Zhang, F.; Li, W.-X. Involvement of miR169 in the nitrogen-starvation responses in *Arabidopsis*. *New Phytol.* **2011**, *190*, 906–915.
49. Jones-Rhoades, M.W.; Bartel, D.P.; Bartel, B. microRNAs and their regulatory roles in plants. *Annu. Rev. Plant Biol.* **2006**, *57*, 19–53.
50. Achard, P.; Herr, A.; Baulcombe, D.C.; Harberd, N.P. Modulation of floral development by a gibberellin-regulated microRNA. *Development* **2004**, *131*, 3357–3365.
51. Strand, A.; Asami, T.; Alonso, J.; Ecker, J.R.; Chory, J. Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. *Nature* **2003**, *421*, 79–83.
52. McCourt, P.; Creelman, R. The ABA receptors—We report you decide. *Curr. Opin. Plant Biol.* **2008**, *11*, 474–478.
53. Rasband, W.S.; ImageJ, U.S. ImageJ software; National Institutes of Health: Bethesda, MD, USA, 1997–2012.
54. Miao, Y.; Jiang, L. Transient expression of fluorescent fusion proteins in protoplasts of suspension cultured cells. *Nat. Protoc.* **2007**, *2*, 2348–2353.

Reprinted from *IJMS*. Cite as: Zhuang, Y.; Zhou, X.-H.; Liu, J. Conserved miRNAs and Their Response to Salt Stress in Wild Eggplant *Solanum linnaeanum* Roots. *Int. J. Mol. Sci.* **2014**, *15*, 839-849.

Article

Conserved miRNAs and Their Response to Salt Stress in Wild Eggplant *Solanum linnaeanum* Roots

Yong Zhuang *, Xiao-Hui Zhou and Jun Liu

Institute of Vegetable Crops, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China; E-Mails: xhzhou1984@sina.com (X.-H.Z.); Kehl_lau@foxmail.com (J.L.)

* Author to whom correspondence should be addressed; E-Mail: jaaszy@163.com; Tel.: +86-25-8439-1752; Fax: +86-25-8439-0262.

Received: 12 November 2013; in revised form: 19 November 2013 / Accepted: 30 December 2013 / Published: 9 January 2014

Abstract: The Solanaceae family includes some important vegetable crops, and they often suffer from salinity stress. Some miRNAs have been identified to regulate gene expression in plant response to salt stress; however, little is known about the involvement of miRNAs in Solanaceae species. To identify salt-responsive miRNAs, high-throughput sequencing was used to sequence libraries constructed from roots of the salt tolerant species, *Solanum linnaeanum*, treated with and without NaCl. The sequencing identified 98 conserved miRNAs corresponding to 37 families, and some of these miRNAs and their expression were verified by quantitative real-time PCR. Under the salt stress, 11 of the miRNAs were down-regulated, and 3 of the miRNAs were up-regulated. Potential targets of the salt-responsive miRNAs were predicted to be involved in diverse cellular processes in plants. This investigation provides valuable information for functional characterization of miRNAs in *S. linnaeanum*, and would be useful for developing strategies for the genetic improvement of the Solanaceae crops.

Keywords: salt stress; miRNA; *Solanum linnaeanum*; high-throughput sequencing

1. Introduction

Salt stress is one of the most common abiotic stresses of crops. It was estimated that salt stress may affect half of all arable lands and will be a major factor of agriculture production for the coming decades [1]. Unlike other abiotic stresses, salt stress brings both osmotic stress and ion toxicity to crops. Under salt stress, crops can respond via cascades of molecular networks to change gene

expression profile and posttranslational modifications involved in a broad spectrum of biochemical, cellular and physiological processes [2,3]. Therefore, an understanding of the basis of the salt stress response is important for strategies aimed at improving crop tolerance to salt stress.

miRNAs are endogenous non-coding small RNAs that are regulators of gene expression in organisms. They are known to play negative regulatory functions at the post-transcription level by inhibiting gene translation or cleaving target mRNAs via base-pairing their target mRNAs [4–6]. Many investigations indicated that plant miRNAs are involved in various important physiological processes, such as seed germination and root development [7–9]. In addition, increasing evidence has shown that miRNAs play important roles in the response of plants to biotic and abiotic stresses [10]; the expression levels of miRNAs were changed in plants infected with virus and fungus [11–13], and miRNAs were identified to be involved in plant response to abiotic stresses such as temperature [14,15], drought [16,17], metals [18,19], and salt [20–22].

The Solanaceae family includes some agriculturally important crops such as potato (*Solanum tuberosum*), eggplant (*S. melongena*), tomato (*S. lycopersicum*), and pepper (*Capsicum annuum*), and they often suffer from salt stress that can cause reduction of production, especially in greenhouse production. *S. linnaeanum*, which was used to construct a comparative genetic linkage map of eggplant, has tolerance to salt stress [23,24], however, little is known about the mechanism in response to salt stress. Comparative genomic studies revealed that relatively few genome rearrangements and duplications occurred in the evolutionary history of the Solanaceae species [25–28]. Although little information is known about the genomes of *S. linnaeanum* and *S. melongena*, the published data of other plants, especially those from Solanaceae family, may provide sufficient reference.

In the present study, using high-throughput sequencing, a large number of miRNAs and their response to salt stress in *S. linnaeanum* roots are identified and characterized. The results lay the foundation for further investigation and better understanding of the regulatory mechanisms for the plant response to salt stress. In addition, it also provides important information for genetic improvement of Solanaceae crops to salt stress.

2. Results and Discussion

2.1. Deep Sequencing Results of Small RNAs from *S. linnaeanum* Roots

To identify the miRNAs and their response to salt in *S. linnaeanum*, two small RNA libraries were generated from roots of NaCl-free (CK) and NaCl-treated (TR). Deep sequencing generates 21,284,496 and 13,989,100 raw reads in two libraries. After removal of low-quality and corrupted adapter sequences, 8,462,890 and 8,999,145 mappable reads remain in two libraries. The size distribution of mappable reads is assessed (Figure 1, Table S1). The data show that 24 nt small RNA is the major size class, followed by 21, 23, 30 and 22 nt small RNA. Similar results were reported in some other plant species, such as *Arabidopsis thaliana* [29,30], *Medicago truncatula* [31], *Oryza sativa* [32], *Arachis hypogaea* [33], *Cucumis Sativus* [34], *Nicotiana tabacum* [35], and *Citrus trifoliata* [36].

Because details of *S. linnaeanum* genome are limited, these mappable reads are analyzed with genome information of tomato and other plants. The results show that 5.51% reads of CK and 4.86% reads of TR are mapped to known plant pre-miRNAs in miRbase. Reads from CK (24.17%) and TR

(24.29%) are mapped to plant repeats, mRNA, and other RNAs including tRNA, rRNA, snRNA and snoRNA. In addition, some reads that cannot be mapped to pre-miRNAs in miRbase and other RNAs are mapped to tomato genome sequences, and a fraction of them potentially form hairpins. Also, nearly half of these reads have no mapping information (Table 1). To eliminate possible sequencing errors, only those sequences with more than five reads in either of the two libraries are further analyzed.

Figure 1. Length distribution of mappable small RNAs in two databases of *S. linnaeanum* roots. TR represents library of NaCl treatment, and CK represents library of control. The number in vertical axis is the total reads of all small RNAs in a certain length.

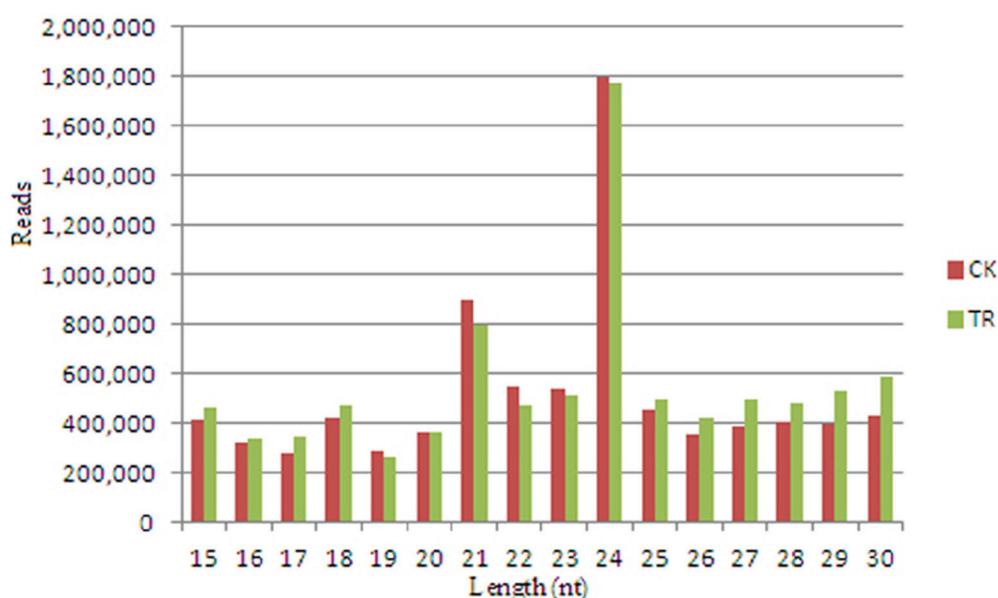


Table 1. Statistical analysis of sequencing reads in the two libraries.

| Category | CK | Percent (%) | TR | Percent (%) |
|-------------------|------------|-------------|------------|-------------|
| Raw reads | 13,989,100 | | 21,284,496 | |
| Mappable reads | 8,462,890 | 100.00 | 8,999,145 | 100.00 |
| Mapped to miRNA | 466,136 | 5.51 | 437,144 | 4.86 |
| Mapped to mRNA | 723,297 | 8.55 | 793,475 | 8.82 |
| Mapped to RFam | 1,315,886 | 15.55 | 1,387,942 | 15.42 |
| Mapped to Repbase | 5,674 | 0.07 | 4,745 | 0.05 |
| Mapped to genome | 1,763,934 | 20.84 | 2,063,801 | 22.93 |
| No hit | 4,187,963 | 49.49 | 4,312,038 | 47.92 |

2.2. Conserved miRNAs in *S. linnaeanum* Roots

To identify the conserved miRNAs in *S. linnaeanum* roots, small RNA sequences are mapped to tomato and other plant miRNAs in miRBase. Based on sequence homology (number of mismatch < 3), 98 known miRNAs and 7 miRNAs* are found (Table S2). The majority of these miRNAs are 20–22 nt long, and 56 of them are 21 nt long. These identified conserved miRNAs correspond to 37 families. The number of miRNA members in each known family shows significant divergence. The miR166 family is the largest one with 11 members, and for the family of miR171, miR396, and miR156, each

of them has 7, 6, and 5 members respectively. Six families including miR159, miR162, miR167, miR168, miR319, and miR390 contain four members, and the remaining 27 miRNA families contain one to three members.

The read counts of miRNAs in sequencing libraries can be used as an index to estimate their relative abundance. In this study, the read counts differ among the miRNAs, which indicate that their expressions varied. Counting redundant miRNA reads reveals that 18 out of 98 known miRNAs and 2 miRNAs* are represented by more than 1000 reads in both libraries, and 5 of them, sli-miR166e (201,378 reads), sli-miR2911c (48,948 reads), sli-miR396d (29,823 reads), sli-miR166f (29,594 reads), and sli-miR403a (28,676 reads) are the most frequent. In addition, sequence analysis shows that the relative abundance of certain member within the miRNA families varies greatly, suggesting functional divergence within the family. For instance, reads of the sli-miR166 family vary from 10 reads (sli-miR166k) to 201,378 reads (sli-miR166e). Similar results are observed in some other miRNA families, such as sli-miR396 (7-29,823 reads) and sli-miR2911 (256-48,948 reads). The above results indicate the different expression levels of different miRNAs in roots, and may be the result of tissue specific or developmental expression.

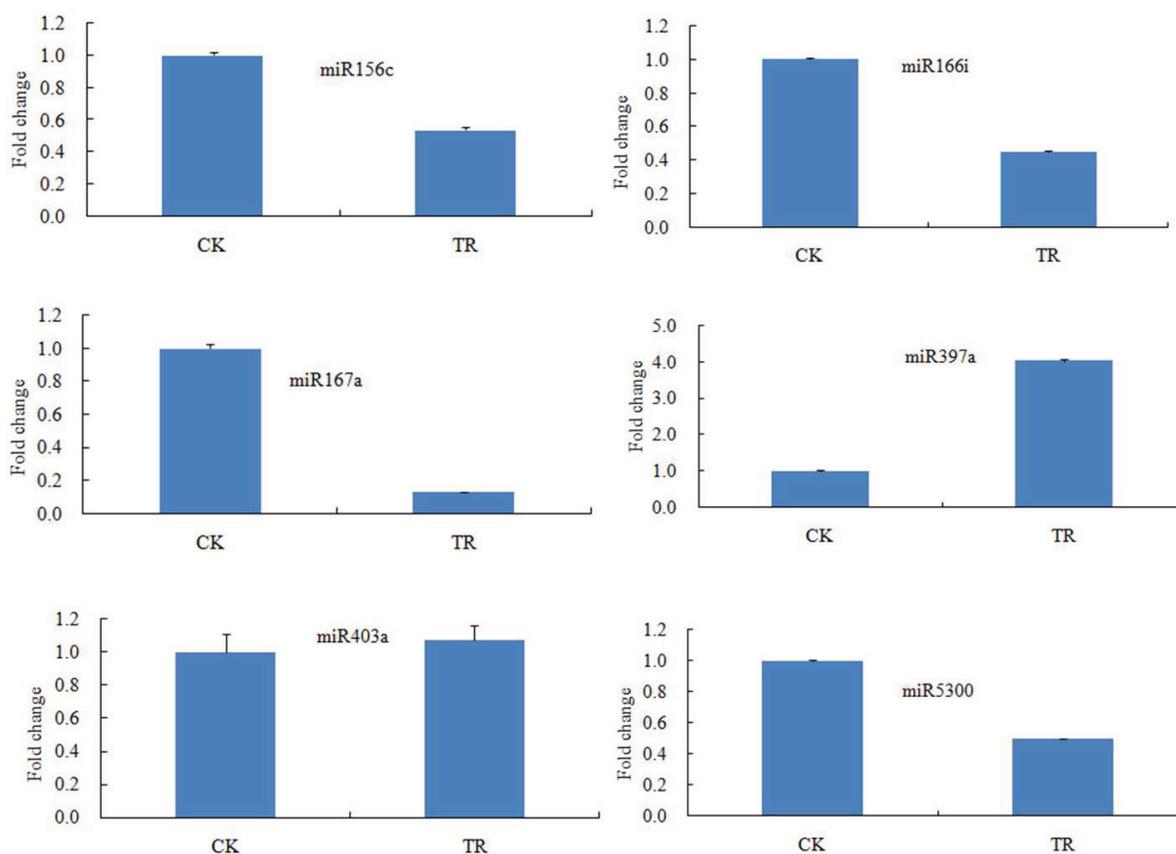
2.3. Validation of miRNAs in *S. linnaeanum* Roots

To verify the results of RNA sequencing and bioinformatics analysis, six miRNAs (sli-miR156c, sli-miR166i, sli-miR167a, sli-miR397a, sli-miR403a and sli-miR5300) are selected randomly for validation by qRT-PCR. According to the Illumina sequencing results, these miRNAs are four down-regulated miRNAs, one up-regulated miRNA and one no responsive miRNA. As shown in the Figure 2 and Table S3, the expression changes detected by qRT-PCR for 4 miRNAs (sli-miR156c, sli-miR166i, sli-miR397a and sli-miR403a) are similar to the results of Illumina sequencing. For sli-miR167a and sli-miR5300, the results have small differences, but they all show down regulation. This may be induced by sequencing error or sampling difference. Above results suggest that miRNAs and their expression changes under NaCl stress have been successfully discovered from *S. linnaeanum* roots by Illumina sequencing.

2.4. NaCl-Responsive miRNAs in *S. linnaeanum* Roots

A deep sequencing approach can be used as a powerful tool for profiling miRNA expression [15,31]. The changes in the frequency of miRNAs between the NaCl-treated and control libraries might indicate that their expression is regulated in response to NaCl stress. To minimize noise and improve accuracy, only the 18–24 nt miRNAs with normalized sequence reads over 10 in at least one library are selected for comparison. miRNAs with $\log_2(\text{TR}/\text{CK}) > 1$ and $p < 0.05$ are designated as up-regulated. Similarly, miRNAs with $\log_2(\text{TR}/\text{CK}) < -1$ and $p < 0.05$ are designated as down-regulated. As showed in Table 2, under the stress of NaCl treatment, 11 miRNAs belonging to eight families are down-regulated, and three miRNAs belonging to three families are up-regulated. The above results indicate that the number of NaCl-induced down-regulated miRNAs is more than that of up-regulated miRNAs.

Figure 2. Validation of selected miRNAs in roots by qRT-PCR. The data are the average of three qRT-PCR replicates for each sample from three biological repeats. Small nuclear RNA U6 is used as an internal reference. Error bars indicate one standard deviation of three different biological replicates. The expression changes of six miRNAs detected by qRT-PCR are consistent with the Illumina sequencing results.



To understand the potential functions of NaCl-responsive miRNAs, 31 target genes (Table S4) for these miRNAs are predicted, and the representative results are listed in Table 2. These genes were reported to be involved in many plant physiological processes, such as plant development, metabolism, and defense. Interestingly, different members in a miRNA family may target the same or different genes. For example, both sli-miRNA156b and sli-miRNA156c can target genes encoding squamosa promoter-binding protein-like, which indicates that they are functionally conservative. The same results are observed for sli-miR171b and sli-miR171e, and they all target the gene encoding scarecrow transcription factor family protein. However, for sli-miR167a and sli-miR167b, which can target different genes, their functions may be differentiated by sequence variation.

As *S. linnaeanum* is salt-tolerant, these salt responsive miRNAs may play an important role for salt tolerance. Some miRNAs, such as *Zea mays* miR166, miR159, miR156 and miR319, and *Arabidopsis* miR393, miR397b, and miR402, have been reported to show altered expression profile under salt stress [21,37]. In the present study, one of the up-regulated miRNA, sli-miR397a, is predicted to target a laccase gene, which was reported to reduce root growth under dehydration [38]. Similarly, another up-regulated miRNA, sli-miR166d, is predicted to target a DNA repair protein RAD4 family gene which was previously found as a key repair factor that directly recognizes DNA damage and initiates

DNA repair, and recently it was found to regulate protein turnover at a postubiquitylation step [39]. Because of the salt tolerance of *S. linnaeanum*, it is possible that the roots do not suffer with serious injuries such as reduced growth and DNA damage. Therefore, the two potential targets do not need to show high expression to alleviate the injuries. However, further investigations are needed to confirm the above hypothesis.

Table 2. NaCl-responsive miRNAs and their targets. The value of TR/CK is the ratio between normalized count from library TR and CK.

| miRNA | Log ₂ (TR/CK) | <i>p</i> value | Predicted target | Putative function of target |
|-------------|--------------------------|-------------------------|--------------------|---|
| sli-miR156b | -1.01 | 2.46×10^{-127} | SGN-U325281 | Squamosa promoter-binding protein |
| sli-miR156c | -1.18 | 2.72×10^{-52} | SGN-U317176 | Squamosa promoter-binding protein |
| sli-miR162b | -1.25 | 8.35×10^{-99} | Solyc10g005130.2.1 | Ribonuclease 3-like protein 3 |
| sli-miR164c | 1.03 | 3.57×10^{-16} | SGN-U327571 | Lipase-related |
| sli-miR166d | 1.97 | 1.02×10^{-45} | Solyc11g011150.1.1 | DNA repair protein Rad4 family |
| sli-miR167a | -1.45 | 4.26×10^{-166} | SGN-U313907 | Annexin 1 |
| sli-miR167b | -1.25 | 5.57×10^{-15} | Solyc03g095940.1.1 | LOB domain family protein |
| sli-miR171b | -1.18 | 1.89×10^{-47} | SGN-U333058 | Scarecrow transcription factor family protein |
| sli-miR171e | -1.08 | 1.24×10^{-12} | SGN-U333058 | Scarecrow transcription factor family protein |
| sli-miR172a | -1.66 | 1.94×10^{-44} | SGN-U563871 | Floral homeotic protein APETALA2 |
| sli-miR319a | -1.07 | 2.28×10^{-28} | SGN-U31990 | TCP family transcription factor |
| sli-miR397a | 1.91 | 1.04×10^{-43} | SGN-U327694 | Laccase |
| sli-miR399b | -1.14 | 9.82×10^{-14} | Solyc03g031410.1.1 | Unknown Protein |
| sli-miR5300 | -1.92 | 1.55×10^{-155} | SGN-U336733 | CC-NBS-LRR protein |

Unlike the targets of up-regulated miRNAs in *S. linnaeanum*, some of the down-regulated miRNAs target mRNAs of transcription factors, indicating an upstream regulation of miRNAs during the response to salt stress. sli-miR171b and sli-miR171e are predicted to target a scarecrow transcription factor gene which was reported to be involved in ground tissue formation in Arabidopsis root [40]. sli-miR172a is predicted to target an Floral homeotic protein APETALA2 gene. SlAP2a, the true ortholog of AP2 in tomato has been found to control fruit ripening via regulation of ethylene biosynthesis and signaling [41]. However, the role of AP2 in response to salt stress has not been described in detail. sli-miR319a is predicted to target a TCP family transcription factor gene which was reported to play a pivotal role in the control of morphogenesis of shoot organs by negatively regulating the expression of boundary-specific genes in Arabidopsis [42]. The above results indicate that the function involved in the response to salt stress of these potential targets needs to be explored in depth. The identification of salt-responsive miRNAs that target these genes may suggest additional roles for the defense against salt stress.

3. Experimental Section

3.1. Plant Materials and NaCl Treatment

A wild eggplant species, *S. linnaeanum* (PI388846) is used in this study. The seeds are surface-sterilized with 70% ethanol, and allowed to germinate in 30 °C. The uniform germinated seeds are sown in pots containing commercial nursery substrate. The seedlings are grown in an incubator with a 16 h photoperiod at a temperature regime of 25 °C. When the seedlings develop five true leaves, uniform seedlings are picked out and irrigated with 150 mM NaCl for salt treatment or the distilled water as a control. The roots of the NaCl treated and control plants are harvested after 24 h. The collected roots are pooled with ten plants and immediately frozen in liquid nitrogen for RNA extraction.

3.2. Small RNA Library Construction and Sequencing

Total RNA is extracted with the Total RNA Purification Kit (Norgen Biotek, Thorold, Canada) and treated with *DNase I* according to the manufacturer's instructions. The small RNA libraries are constructed using the Truseq™ Small RNA Preparation kit (Illumina, San Diego, CA, USA). The purified cDNA library from 15 to 32 nt small RNAs is used for cluster generation on Illumina's Cluster Station and then sequenced on Illumina GAIIx (San Diego, CA, USA). Raw sequencing reads are obtained using Illumina's Sequencing Control Studio software version 2.8 (SCS v2.8, San Diego, CA, USA) following real-time sequencing image analysis and base-calling by Illumina's Real-Time Analysis version 1.8.70 (RTA v1.8.70).2.1.1 (Illumina, San Diego, CA, USA).

3.3. Analysis of Small RNA Sequencing Data

A proprietary pipeline script, ACGT101-miR v4.2 (LC Sciences, Houston, TX, USA), is used for sequencing data analysis. The "impurity" reads due to sample preparation, sequencing chemistry and processes, and the optical digital resolution of the sequencer detector are removed. Those remaining sequences are grouped by families (unique sequences). Thereafter, families that match known plant repeats, mRNA, rRNAs, tRNAs, snRNAs, and snoRNAs were removed. The remaining unique sequences are mapped to known plant miRNAs from miRBase and Pre-miRBase (Version 17.0, <ftp://mirbase.org/pub/mirbase/CURRENT>, University of Manchester, Manchester, UK) and *S. lycopersicum* genome database (PlantGDB, ftp://ftp.plantgdb.org/download/Genomes/SIGDB/ITAG2_genomic.fasta).

The number of read copies from each sample is tracked during mapping and normalized for comparison. The normalization of sequence counts in each sample is achieved by dividing the counts by a library size parameter of the corresponding sample. The library size parameter is a median value of the ratio between the counts a specific sample and a pseudo-reference sample. A count number in the pseudo-reference sample is the count geometric mean across two samples. For miRNA expression analysis, *p* value calculation is performed with the method introduced by Audic and Claverie [43].

3.4. miRNA Validation by Quantitative Real-Time PCR

The identified *S. linnaeanum* miRNAs are validated by using quantitative real time PCR (qRT-PCR). In this study, six conserved miRNAs (sli-miR156c, sli-miR166i, sli-miR167a, sli-miR397a,

sli-miR403a and sli-miR5300) are validated. Total RNA is isolated from roots of CK and TR, which are samples of parallel experiments for RNA sequencing. For determination of miRNA expression, RNAs are reverse-transcribed by miScript II Reverse Transcription Kit (Qiagen, Germantown, MD, USA), which adds a poly (A) tail to the 3'-end of miRNA and with transcription led by a known oligo-dT ligase. SuperReal PreMix (SYBR Green, TIANGEN, Beijing, China) is used for qRT-PCR. Small nuclear RNA U6 is used as an internal reference. The primers for the 6 miRNAs are universal primers (QIAGEN, Germantown, MD, USA) and corresponding miRNA sequences. qRT-PCR experiments are performed on Roche LightCycler 480 II. PCR program is set as: (1) 95 °C, 15 min; (2) 95 °C, 10 s, thereafter 60 °C, 30 s, 40 cycles. All reactions are run in three replicates for each sample from three biological repeats.

3.5. Prediction of miRNA Target Genes

The putative target sites of miRNA are identified using the psRNATarget program (<http://plantgrn.noble.org/psRNATarget/>) with default parameters [44]. Because there is not enough genome information for *S. melongena* and *S. linnaeanum*, the database of tomato *S. lycopersicum* is used as the sequence library for target search.

4. Conclusions

In this study, by using high-throughput sequencing and taking advantage of the genome information of other plants, 98 known miRNAs were discovered in *S. linnaeanum* roots, and 14 of them show response to salt stress. The potential targets of the identified salt responsive miRNAs are also predicted based on sequence homology search. However, the further investigation for the function of potential target genes still needs to be performed. As more salt tolerance related miRNAs are confirmed, artificial miRNA will be a powerful tool to create elite plant germplasm with salt tolerance.

Acknowledgments

This research was partially supported by the National Natural Science Foundation (31101542); Natural Science Foundation (BK2011675); and Independent Innovation Foundation of Agricultural Sciences (CX (13)2003) of Jiangsu Province, China.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Wang, W.; Vinocur, B.; Altman, A. Plant responses to drought, salinity and extreme temperatures: Towards genetic engineering for stress tolerance. *Planta* **2003**, *218*, 1–14.
2. 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.

3. Vinocur, B.; Altman, A. Recent advances in engineering plant tolerance to abiotic stress: Achievements and limitations. *Curr. Opin. Biotechnol.* **2005**, *16*, 123–132.
4. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297.
5. Jones-Rhoades, M.W.; Bartel, D.P.; Bartel, B. MicroRNAs and their regulatory roles in plants. *Annu. Rev. Plant Biol.* **2006**, *57*, 19–53.
6. Voinnet, O. Origin, biogenesis, and activity of plant microRNAs. *Cell* **2009**, *136*, 669–687.
7. Liu, P.P.; Montgomery, T.A.; Fahlgren, N.; Kasschau, K.D.; Nonogaki, H.; Carrington, J.C. Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is critical for seed germination and post-germination stages. *Plant J.* **2007**, *52*, 133–146.
8. Reyes, J.L.; Chua, N.H. ABA induction of miR159 controls transcript levels of two MYB factors during Arabidopsis seed germination. *Plant J.* **2007**, *49*, 592–606.
9. Gutierrez, L.; Bussell, J.D.; Pacurar, D.I.; Schwambach, J.; Pacurar, M.; Bellini, C. Phenotypic plasticity of adventitious rooting in Arabidopsis is controlled by complex regulation of AUXIN RESPONSE FACTOR transcripts and microRNA abundance. *Plant Cell* **2009**, *21*, 3119–3132.
10. Sunkar, R.; Chinnusamy, V.; Zhu, J.; Zhu, J.K. Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. *Trends Plant Sci.* **2007**, *12*, 301–309.
11. Bazzini, A.A.; Hopp, H.E.; Beachy, R.N.; Asurmendi, S. Infection and coaccumulation of tobacco mosaic virus proteins alter microRNA levels, correlating with symptom and plant development. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 12157–12162.
12. Lu, S.; Sun, Y.H.; Amerson, H.; Chiang, V.L. MicroRNAs in loblolly pine (*Pinus taeda* L.) and their association with fusiform rust gall development. *Plant J.* **2007**, *51*, 1077–1098.
13. Yang, L.; Jue, D.; Li, W.; Zhang, R.; Chen, M.; Yang, Q. Identification of MiRNA from eggplant (*Solanum melongena* L.) by small RNA deep sequencing and their response to *Verticillium dahliae* infection. *PLoS One* **2013**, *8*, e72840.
14. Tang, Z.; Zhang, L.; Xu, C.; Yuan, S.; Zhang, F.; Zheng, Y.; Zhao, C. Uncovering small RNA-mediated responses to cold stress in a wheat thermosensitive genic male-sterile line by deep sequencing. *Plant Physiol.* **2012**, *159*, 721–738.
15. Yu, X.; Wang, H.; Lu, Y.; de Ruiter, M.; Cariaso, M.; Prins, M.; van Tunen, A.; He, Y. Identification of conserved and novel microRNAs that are responsive to heat stress in *Brassica rapa*. *J. Exp. Bot.* **2012**, *63*, 1025–1038.
16. Barrera-Figueroa, B.E.; Gao, L.; Diop, N.N.; Wu, Z.; Ehlers, J.D.; Roberts, P.A.; Close, T.J.; Zhu, J.K.; Liu, R. Identification and comparative analysis of drought-associated microRNAs in two cowpea genotypes. *BMC Plant Biol.* **2011**, *11*, 127.
17. Li, B.; Qin, Y.; Duan, H.; Yin, W.; Xia, X. Genome-wide characterization of new and drought stress responsive microRNAs in *Populus euphratica*. *J. Exp. Bot.* **2011**, *62*, 3765–3779.
18. Ding, Y.; Chen, Z.; Zhu, C. Microarray-based analysis of cadmium-responsive microRNAs in rice (*Oryza sativa*). *J. Exp. Bot.* **2011**, *62*, 3563–3573.
19. Chen, L.; Wang, T.; Zhao, M.; Tian, Q.; Zhang, W.H. Identification of aluminum-responsive microRNAs in *Medicago truncatula* by genome-wide high-throughput sequencing. *Planta* **2012**, *235*, 375–386.

20. Liu, H.H.; Tian, X.; Li, Y.J.; Wu, C.A.; Zheng, C.C. Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis Thaliana*. *RNA* **2008**, *14*, 836–843.
21. Ding, D.; Zhang, L.; Wang, H.; Liu, Z.; Zhang, Z.; Zheng, Y. Differential expression of miRNAs in response to salt stress in maize roots. *Ann. Bot.* **2009**, *103*, 29–38.
22. Covarrubias, A.A.; Reyes, J.L. Post-transcriptional gene regulation of salinity and drought responses by plant microRNAs. *Plant Cell Environ.* **2010**, *33*, 481–489.
23. Collonnier, C.; Fock, I.; Kashyap, V.; Rotino, G.L.; Daunay, M.C.; Lian, Y.; Mariska, I.K.; Rajam, M.V.; Servaes, A.; Ducreux, G.; *et al.* Applications of biotechnology in eggplant. *Plant Cell Tissue Organ* **2001**, *65*, 91–107.
24. Doganlar, S.; Frary, A.; Daunay, M.C.; Lester, R.N.; Tanksley, S.D. A comparative genetic linkage map of eggplant (*Solanum melongena*) and its implications for genome evolution in the solanaceae. *Genetics* **2002**, *161*, 1697–1711.
25. Mueller, L.A.; Solow, T.H.; Taylor, N.; Skwarecki, B.; Buels, R.; Binns, J.; Lin, C.; Wright, M.H.; Ahrens, R.; Wang, Y.; *et al.* The SOL Genomics Network: A comparative resource for solanaceae biology and beyond. *Plant Physiol.* **2005**, *138*, 1310–1317.
26. Daniell, H.; Lee, S.B.; Grevich, J.; Saski, C.; Quesada-Vargas, T.; Guda, C.; Tomkins, J.; Jansen, R.K. Complete chloroplast genome sequences of *Solanum bulbocastanum*, *Solanum lycopersicum* and comparative analyses with other Solanaceae genomes. *Theor. Appl. Genet.* **2006**, *112*, 1503–1518.
27. Wu, F.; Eannetta, N.T.; Xu, Y.; Durrett, R.; Mazourek, M.; Jahn, M.M.; Tanksley, S.D. A COSII genetic map of the pepper genome provides a detailed picture of synteny with tomato and new insights into recent chromosome evolution in the genus *Capsicum*. *Theor. Appl. Genet.* **2009**, *118*, 1279–1293.
28. Wu, F.; Eannetta, N.T.; Xu, Y.; Tanksley, S.D. A detailed synteny map of the eggplant genome based on conserved ortholog set II (COSII) markers. *Theor. Appl. Genet.* **2009**, *118*, 927–935.
29. 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.
30. Fahlgren, N.; Howell, M.D.; Kasschau, K.D.; Chapman, E.J.; Sullivan, C.M.; Cumbie, J.S.; Givan, S.A.; Law, T.F.; Grant, S.R.; Dangl, J.L.; *et al.* High-throughput sequencing of *Arabidopsis* microRNAs: Evidence for frequent birth and death of MIRNA genes. *PLoS One* **2007**, *2*, e219.
31. Szittyá, G.; Moxon, S.; Santos, D.M.; Jing, R.; Fevereiro, M.P.; Moulton, V.; Dalmay, T. High-throughput sequencing of *Medicago truncatula* short RNAs identifies eight new miRNA families. *BMC Genomics* **2008**, *9*, 593.
32. Morin, R.D.; Aksay, G.; Dolgosheina, E.; Ebhardt, H.A.; Magrini, V.; Mardis, E.R.; Sahinalp, S.C.; Unrau, P.J. Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Res.* **2008**, *18*, 571–584.
33. Chi, X.; Yang, Q.; Chen, X.; Wang, J.; Pan, L.; Chen, M.; Yang, Z.; He, Y.; Liang, X.; Yu, S. Identification and characterization of microRNAs from peanut (*Arachis hypogaea* L.) by high-throughput sequencing. *PLoS One* **2011**, *6*, e27530.
34. Martínez, G.; Forment, J.; Llave, C.; Pallás, V.; Gómez, G. High-throughput sequencing, characterization and detection of new and conserved cucumber miRNAs. *PLoS One* **2011**, *6*, e19523.

35. Guo, H.; Kan, Y.; Liu, W. Differential expression of miRNAs in response to topping in flue-cured tobacco (*Nicotiana tabacum*) roots. *PLoS One* **2011**, *6*, e28565.
36. Song, C.; Wang, C.; Zhang, C.; Korir, N.K.; Yu, H.; Ma, Z.; Fang, J. Deep sequencing discovery of novel and conserved microRNAs in trifoliolate orange (*Citrus trifoliata*). *BMC Genomics* **2010**, *11*, 431.
37. Sunkar, R.; Zhu, J.K. Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. *Plant Cell* **2004**, *16*, 2001–2019.
38. Cai, X.; Davis, E.J.; Ballif, J.; Liang, M.; Bushman, E.; Haroldsen, V.; Torabinejad, J.; Wu, Y. Mutant identification and characterization of the laccase gene family in Arabidopsis. *J. Exp. Bot.* **2006**, *57*, 2563–2569.
39. Li, Y.; Yan, J.; Kim, I.; Liu, C.; Huo, K.; Rao, H. Rad4 regulates protein turnover at a postubiquitylation step. *Mol. Biol. Cell* **2010**, *21*, 177–185.
40. Pauluzzi, G.; Divol, F.; Puig, J.; Guiderdoni, E.; Dievart, A.; Périn, C. Surfing along the root ground tissue gene network. *Dev. Biol.* **2012**, *365*, 14–22.
41. Chung, M.Y.; Vrebalov, J.; Alba, R.; Lee, J.; McQuinn, R.; Chung, J.D.; Klein, P.; Giovannoni, J. A tomato (*Solanum lycopersicum*) *APETALA2/ERF* gene, *SLAP2a*, is a negative regulator of fruit ripening. *Plant J.* **2010**, *64*, 936–947.
42. Koyama, T.; Furutani, M.; Tasaka, M.; Ohme-Takagi, M. TCP transcription factors control the morphology of shoot lateral organs via negative regulation of the expression of boundary-specific genes in Arabidopsis. *Plant Cell* **2007**, *19*, 473–484.
43. Audic, S.; Claverie, J.M. The significance of digital gene expression profiles. *Genome Res.* **1997**, *7*, 986–995.
44. Dai, X.; Zhao, P.X. psRNATarget: A plant small RNA target analysis server. *Nucleic Acids Res.* **2011**, *39*, W155–W159.

Reprinted from *IJMS*. Cite as: Nicolás, F.E.; Ruiz-Vázquez, R.M. Functional Diversity of RNAi-Associated sRNAs in Fungi. *Int. J. Mol. Sci.* **2013**, *14*, 15348-15360.

Review

Functional Diversity of RNAi-Associated sRNAs in Fungi

Francisco E. Nicolás^{1,2,*} and Rosa M. Ruiz-Vázquez¹

¹ Department of Genetics and Microbiology, Faculty of Biology, University of Murcia, Murcia 30100, Spain; E-Mail: rmruiz@um.es

² Regional Campus of International Excellence “Campus Mare Nostrum”, Murcia 30100, Spain

* Author to whom correspondence should be addressed; E-Mail: fnicolas@um.es; Tel.: +34-868-887135; Fax: +34-868-883963.

Received: 30 May 2013; in revised form: 5 July 2013 / Accepted: 9 July 2013 /

Published: 24 July 2013

Abstract: Yeast and filamentous fungi have been essential model systems for unveiling the secrets of RNA interference (RNAi). Research on these organisms has contributed to identifying general mechanisms and conserved eukaryotic RNAi machinery that can be found from fungi to mammals. The development of deep sequencing technologies has brought on the last wave of studies on RNAi in fungi, which has been focused on the identification of new types of functional small RNAs (sRNAs). These studies have discovered an unexpected diversity of sRNA, biogenesis pathways and new functions that are the focus of this review.

Keywords: fungi; sRNA; siRNA; masiRNAs; esRNA; ex-siRNA; miRNA; qiRNA

1. Introduction

RNAi is a negative regulatory mechanism that represses the expression of target RNAs. It was firstly described as a mechanism triggered by double-stranded RNA (dsRNA) in the worm *Caenorhabditis elegans* [1]. However, this phenomenon was previously discovered in plants and fungi, in which it was named with the terms co-suppression and quelling, respectively [2,3]. Both in plants and fungi, it was observed after transforming a wild type strain with exogenous gene sequences that are required for the biosynthesis of different pigments. The result of these transformations was the lack of expression of both the transgene and the endogenous homolog sequence, which produced an albino phenotype instead of the expected overproduction of pigments [3]. In the case of fungi, these

experiments were carried out in *Neurospora crassa*, a robust study model with many available genetic tools that allowed the cloning of several genes involved in RNAi and helped to unveil the main core of the RNAi machinery. Using a transformed strain that showed a stable albino phenotype for mutagenesis, several quelling deficient mutants (*qde*) were isolated [4]. Complementation analyses of these mutants concluded in the identification of three different genes: *qde-1*, *qde-2* and *qde-3*. An RNA dependent RNA polymerase (RDRP), encoded by *qde-1*, was the first component of the RNAi machinery that was cloned in these screenings [4]. The role of this enzyme is to generate dsRNA from the aberrant RNAs (aRNAs) that are hypothetically produced from the transgene. The essential role of dsRNA as the main trigger of RNAi was demonstrated soon after in *C. elegans* in a wonderful work that was awarded the Nobel prize in 2006 [1]. The next enzyme required in the RNAi pathway is a ribonuclease type III that processes the dsRNA into the small interfering RNAs (siRNAs), a special type of sRNAs that exhibits a fixed size between 19 and 25 nt and has 5' phosphates and two nucleotide overhangs on the 3' ends [5]. This ribonuclease is known as the Dicer enzyme and despite its essential role in RNAi, it was not one of the three *qde* genes that were initially identified in this fungus [6]. This was due to the partially redundant activity of the two Dicer-like proteins that were found later in *N. crassa*, when the genome sequence was available [6]. The third essential enzyme of the RNAi core machinery is the Argonaute protein (Ago) which is encoded by *qde-2* gene [4]. The Ago protein acts downstream of Dicer, incorporating the siRNAs into the so-called RNA-induced silencing complex (RISC), a protein complex containing Ago, the siRNA, and other accessory proteins. The slicing activity of Ago produces a nicked siRNA duplex, which helps the exonuclease QIP to remove the passenger strand of the siRNA duplex, provoking the activation of the RISC complex [7,8]. Once the RISC complex has been activated, it uses the guide strand of the siRNA to identify complementary mRNA molecules, which are degraded by endonucleolytic and exonucleolytic cleavages. The last quelling deficient mutant that was cloned in *N. crassa*, *qde-3*, encodes a RecQ DNA helicase which could be involved in the generation of aRNAs/dsRNA [4]. Mutants in the *qde-3* gene present defective mechanisms of both RNAi and DNA repair in *N. crassa* [9,10]. It has been suggested that QDE-3 could be involved in the recognition of exogenous DNA and the generation of aRNA by recruiting QDE-1 to the ssDNA resulting from aberrant DNA structures, which could be produced during replication or recombination of repetitive sequences [11]. The synthesis of the aRNA has been traditionally assigned to RNA polymerase II, which would transcribe the ssDNA template to produce the aRNA used by QDE-1 to generate dsRNA. However, the crystal structure of QDE-1 has revealed that its catalytic core is structurally similar to eukaryotic DNA-dependent RNA polymerases (DdRPs) rather than viral RdRPs [11,12]. These findings suggested that QDE-1 could be both an RdRP and a DdRP, being required both for the synthesis of the aRNA and the subsequent dsRNA production; this dual activity of QDE-1 has been experimentally demonstrated [11].

The RNAi mechanism and its canonical pathway were initially described as a host defense mechanism that protects the genome from invasive nucleic acids, such as viruses and transposons. This defensive role supports that RNAi is an essential mechanism that has been evolutionary conserved through the entire eukaryotic domain. However, the existence of several eukaryotic microbes that lack an active RNAi pathway raises the question of how they can survive without the protective role of RNAi. It has been proposed that the RNAi mechanism may represent a liability rather than an advantage in some special evolutionary scenarios, forcing the selection of RNAi-deficient

species [13,14]. This is the case of several yeast species, including *Saccharomyces cerevisiae*, in which the absence of functional RNAi has been correlated with the presence of endemic dsRNA viruses that confer selective advantages to the host cells [13]. Besides the defensive role of RNAi, the discovery in recent years of several new pathways that are endogenously triggered has unveiled RNAi as a new regulatory mechanism that controls different cell functions. The different fungal sRNAs related to both aspects of RNAi, host defense and regulation of gene expression, are the focus of this review (Table 1).

Table 1. Classes of RNAi-dependent sRNAs in fungi.

| Function | Name | Acronym | Inducer | Firstly described | Reference | | |
|----------------------------|------------------------------|------------|--|---------------------------------|-------------------------------|----------------------------------|------|
| Host Defense | Small Interfering RNAs | siRNAs | Integrative transgenes | <i>Neurospora crassa</i> | [7] | | |
| | | | Non integrative transgenes | <i>Mucor circinelloides</i> | [15] | | |
| | | | Transposons | <i>Neurospora crassa</i> | [16] | | |
| | | | Viruses | <i>Cryphonectria parasitica</i> | [17] | | |
| | | | MSUD-associated small interfering RNAs | masiRNAs | Unpaired DNA | <i>Neurospora crassa</i> | [18] |
| Endogenous Gene Regulation | Sex Induced Silencing siRNAs | SIS siRNAs | Repetitive transgenes | <i>Cryptococcus neoformans</i> | [19] | | |
| | | | Exonic-siRNAs | ex-siRNA | Regular transcription | <i>Mucor circinelloides</i> | [20] |
| | | | MiRNA-like RNAs | milRNAs | Regular transcription | <i>Neurospora crassa</i> | [21] |
| | | | Heterochromatin derived siRNAs | siRNAs | Heterochromatin transcription | <i>Schizosaccharomyces pombe</i> | [22] |
| | | | QDE-2-interacting sRNAs | qiRNAs | DNA damage | <i>Neurospora crassa</i> | [23] |

2. Host Defense sRNAs in Fungi

The RNAi mechanism can be triggered by a wide variety of exogenous nucleic acids that represent a threat for genome integrity. Thus, different exogenous nucleic acids such as integrative transgenes, plasmids, viruses and transposons have been found to trigger the RNAi mechanism against whatever is transcribed from them. All the pathways and different types of sRNAs that are produced during this response are reviewed in this section.

2.1. siRNAs Triggered by Randomly Integrated Exogenous Sequences

The first type of sRNAs found in fungi were siRNAs produced to silence exogenous sequences that are homologous to an endogenous gene [7]. In particular, these siRNAs were found in *N. crassa* after the introduction of exogenous sequences, which were randomly integrated at ectopic locations of the genome of this fungus [7]. Wild type mycelium of *N. crassa* shows a bright orange phenotype due to the synthesis of carotenoids. Mutations in any of the three structural genes involved in the carotenoid biosynthesis, *albino-1* (*al-1*), *albino-2* (*al-2*) and *albino-3* (*al-3*), result in an easily detectable albino phenotype. Transformation of the wild type strain of *N. crassa* with different constructs of the *al-1*

gene, encoding the phytoene dehydrogenase gene, produced 36% albino transformants [3]. The mechanism that initiates the production of these siRNAs is still unclear. They are processed from dsRNA, as it is a common step for all the siRNAs associated to RNAi. The production of this dsRNA from the transgene and how the cell distinguishes between the endogenous gene and the exogenous transgene to initiate the RNAi mechanism are not certainly known. The analysis of these albino transformants has revealed that the copy number of the integrated exogenous sequences ranged from 2 to 20, with no gene rearrangements in any of the analyzed transformants that could explain the production of dsRNA. The analysis of the revertants, which are albino transformants that turn to wild type in the following generation, showed a reduced copy number of exogenous sequences, although this cannot be the direct reason of the albino phenotype, as indicated by the existence of albino transformants containing only two copies of the transgene. The current most accepted hypothesis to explain the early steps in the biogenesis of transgene-induced siRNAs is the production of aRNA from the repetitive transgene sequences integrated into the genome (see Introduction). Besides *N. crassa*, RNAi triggered by the integration of exogenous sequences has been described in others fungi such as *Cladosporium fulvum* and *Schizophyllum commune* [24,25].

2.2. Two Classes of siRNAs Triggered by Non-Integrative Transgenes in *Mucor circinelloides*

M. circinelloides is an outstanding study model among zygomycetes due to the existence of an efficient transformation protocol that allows genetics analysis in this fungus [26]. An added value of genetic transformation of *M. circinelloides* is that a non-integrative transgene can be delivered in self replicative plasmids. These plasmids behave as extrachromosomal DNA molecules inside the nucleus of this fungus. They exhibit all the features of self-replicative molecules, such as a high frequency of transformation, mitotic instability, easy re-isolation in a non-modified state from undigested transformant DNA and detection as discrete DNA molecules in Southern hybridization experiments [27]. An advantage of triggering RNAi using this kind of plasmid is that the expression of the non-integrative transgene is not affected by position effects or host regulatory sequences at insertion sites, such as inverted promoters or repeated transgene integrations in different orientations. This clean system to trigger RNAi enables a correlation between the transgene copy number and the strength of the silenced phenotype in *M. circinelloides*, and added new evidence supporting the hypothesis of aRNA as the link molecule between the invasive DNA and the necessary dsRNA [15]. However, the most interesting feature of RNA silencing in *M. circinelloides* is the existence of two different classes of antisense siRNAs that are differentially accumulated during vegetative growth. A long siRNA of 25 nt is more abundant at the beginning of the growth cycle, whereas a shorter 21 nt siRNA is accumulated at the end of the vegetative cycle and transmitted to the next generation through the spore [15]. The biological function of these two different classes of siRNAs and the differential role that they could play during growth and development is still unknown.

2.3. Transposon Control siRNAs

The mechanism of RNAi was initially observed after the integration of repetitive transgenes in the genome of plant and fungal cells [2,3]. It is obvious that this complex mechanism has not evolved to protect the genome against transgenes that are artificially delivered into the cell by a laboratory

protocol. However, the repetitive nature of transgene integration in silenced strains immediately suggested that RNAi could be a defense mechanism against invasive repetitive sequences like transposons. There are numerous examples through the eukaryotic kingdom showing the important role of RNAi in the maintenance of genome integrity by silencing these transposable elements [28]. The suggested mechanism proposes fortuitous transcription of the terminal inverted repeats of the transposon, which leads to the formation of dsRNA and the consequent activation of RNAi [29]. In fungi, genetic analysis with an African strain of *N. crassa* that harbors a LINE-like transposon showed that the RNAi mechanism is required to suppress transposon replication [16]. This analysis revealed that QDE-2 and Dicer, but not QDE-1 and QDE-3, are essential for silencing the transposable elements and suggested that transposons trigger the RNAi mechanism through the direct synthesis of dsRNA from inverted repeats generated by transposition. Later, the role of siRNAs in transposition control has been demonstrated in other fungi. The introduction of the LTR-retrotransposon MAGGY into *Magnaporthe oryzae* strains harboring mutations in the *dicer* genes resulted in a deficient siRNA accumulation, higher MAGGY mRNA production and faster MAGGY copy number increase than did the wild-type. These results indicated that RNA silencing functioned as an effective defense mechanism against transposable elements [30]. Moreover, deep sequencing of small RNA molecules has identified siRNAs derived from transposable sequences in *M. circinelloides* [20] and *M. oryzae* [31].

2.4. Antiviral siRNAs

Defense against invasive viruses along with the control of transposable elements were the first functions associated to RNAi. This defensive role has been identified in a diverse range of distant eukaryotes, including plants [32], worms [33], flies [34] and mammals [35]. All these organisms produce specific siRNAs that target and destroy the invasive viral genome. In fungi, the defensive action of siRNAs against mycoviruses was experimentally demonstrated in the chestnut blight fungus *Cryphonectria parasitica*, in which one of the two dicer-like genes, *dcl2*, and only one of the four argonaute-like genes, *agl2*, are required for antiviral defense response [17,36]. Mutants in one of these two genes lack the ability to avoid viral infections, becoming debilitated strains that are highly susceptible to mycovirus infections and present a hypovirulent phenotype when they infect their host [37]. The role of these two genes in the viral defense mechanism is similar to other RNAi pathways; the Dicer protein produces virus-specific siRNAs that are transferred to Argonaute containing complexes that target and destroy viral sequences. In response to this RNAi-based viral immunity system, viruses have developed RNAi suppression mechanisms. This is the case for the hypovirus CHV1-EP713, a mycovirus that infects *C. parasitica*, and that has developed an RNAi suppression mechanism based on the protein p29. This protein is a RNAi suppressor that acts in a promoter-dependent manner, mediating the repression of an *argonaute*-like gene (*agl2*) and therefore blocking the activation of the RNAi based antiviral response [38]. Similarly, *Aspergillus nidulans* exhibits a fully functional RNAi pathway whose efficiency is suppressed after mycovirus infections, which confirms the existence of an RNAi suppressor encoded by the virus and the antiviral purpose of the RNAi machinery [39].

2.5. *siRNAs Associated with Meiotic Silencing by Unpaired DNA*

Meiotic silencing by unpaired DNA (MSUD) is a genome surveillance system identified in *N. crassa* and *Gibberella zeae* that transiently silences genes unpaired during the pairing stage of the meiotic prophase I, along with any other DNA homologous to the unpaired sequences [40,41]. This system protects the genome from any trouble that could represent unpaired DNA segments, such as viruses and transposons on the move. The mechanism of action is similar to other RNAi pathways, thus, an aRNA is transcribed from unpaired DNA regions and this aRNA is used as a template to synthesize dsRNA, which is processed into MSUD-associated small interfering RNAs (masiRNAs) [18]. However, the machinery involved in this process presents some differences when compared with the canonical RNAi pathway. For instance, two new components, *sad-1* (*suppressor of ascus dominance 1*) and *sad-2*, have been identified in MSUD [40–42]. The first gene, *sad-1*, encodes an RdRP that is a paralog of *qde-1* [40,43]. SAD-2 is a new element in this pathway that does not contain any conserved domains and that is required for the proper localization of SAD-1 [42]. DCL-1 and QIP are common elements shared between MSUD and canonical RNAi pathways, but the Argonaute protein is encoded by *sms-2* (*suppressor of meiotic silencing 2*) in MSUD. Another element required for MSUD is SAD-3, a putative RNA/DNA helicase that is homologous to *S. pombe* Hrr1. The protein Hrr-1 is required for RNAi-mediated heterochromatin formation in fission yeast, which suggests that heterochromatinization and MSUD could be two processes that are mechanistically related [44,45].

2.6. *Sex Induced siRNAs in Cryptococcus neoformans*

The human fungal pathogen *C. neoformans* has the usual RNAi machinery with Argonaute, Dicer and RdRP as the central components, similarly to other fungi in which RNAi has been described. However, in this fungus the induction of RNAi by tandem integration of transgenes results in a special sex-induced silencing (SIS), as the repetitive transgene is silenced at an ~250-fold lower frequency during vegetative mitotic growth compared with sexual reproduction [19]. Regular siRNAs have been found by deep sequencing during SIS, but the most interesting result is that along with siRNAs from the transgene, many other siRNAs from transposons were identified. The production of these siRNAs was impaired in *rdrp* mutant strains in which the expression of a group of retrotransposons was notably increased during mating. The consequence of this uncontrolled transposon activity was a higher transposition/mutation rate. The interpretation of these results suggests that the function of siRNAs during SIS is to protect the genome of the progeny by reducing transposon activity during the sexual cycle [19].

3. Regulatory Endogenous sRNAs in Fungi

Endogenous short RNAs (esRNAs) are similar to siRNAs in their biogenesis, with the main difference based in the fact that they are directly produced from an endogenous precursor rather than an exogenous trigger molecule. Most classes of esRNAs are produced from precursor double-stranded RNAs, which are directly transcribed from the genome or generated after the action of RdRP enzymes. As it happens with exogenously triggered siRNAs, esRNAs are usually processed by a member of the Dicer family and are also incorporated into an effector complex containing a member of the Argonaute

family, similar to the RISC complex. The most relevant example of esRNAs were found in animals and plants, the so called microRNAs (miRNAs), a class of esRNAs that are produced from hairpin structured RNAs and target mRNAs for their repression [46]. miRNAs play a role in diverse processes such as development, cell differentiation, adaptation to environmental changes and disease [46]. For a long time, miRNAs have been widely considered to be absent in fungi, although *dicer* mutants of several fungi have been reported to be affected in vegetative and developmental processes [47], suggesting the existence of esRNAs with regulatory functions in this kingdom. The application in fungi of the new deep sequencing technologies has revealed the existence of several new regulatory esRNAs that are described in this section.

3.1. Exonic-siRNAs in *M. circinelloides*

The exonic-siRNAs (ex-siRNAs) were the first esRNAs found in fungi that regulate the expression of endogenous target genes through the repression of the corresponding mRNA [20]. These ex-siRNAs were discovered in the basal fungus *M. circinelloides* in which previous results had suggested a role of the RNAi machinery in the regulation of several processes such as asexual sporulation, vegetative development and hyphal morphology [47,48]. Deep sequencing of the short RNA content in the wild type and several RNAi mutants of this fungus showed the existence of ex-siRNAs as a new type of esRNAs. They are produced from exons of the same genes that are later regulated through the repression of the corresponding mRNA [20]. Hundreds of ex-siRNAs-producing exons have been identified, which correspond to a total of 276 genes, since some genes contain more than one exon producing these ex-siRNAs. There are four different classes of ex-siRNAs (classes 1–4) that have been classified based on the differential RNAi machinery involved in their biogenesis. The first two classes (classes 1 and 2) include all ex-siRNAs that are DCL2-dependent and present a strong preference for uracil (92%) in the first position of the molecule, a preference that is shared among Argonaute-bound guide RNAs of animal, plants and other fungi. The ex-siRNAs belonging to class 2 showed reduced levels in the *rdp1*⁻ mutant but not in the *rdp2*⁻ strain, which specifically defines this class. Class 1 contains a few ex-siRNAs that do not require RdRP1 but most of them depend on RdRP2. Class 3 covers a significant group of ex-siRNAs that are processed either by DCL1 or DCL2, since they are down-regulated only in the double *dicer* mutant but not in *dcl1*⁻ or *dcl2*⁻ single mutants, indicating a redundant function of the two *dicer* genes in the production of this class of ex-siRNAs. Oppositely, both RdRP enzymes are required for the biogenesis of class 3, as these ex-siRNAs are down-regulated in mutants of either *rdp1* or *rdp2*. The class 4 contains ex-siRNAs produced from only five exons and they are down-regulated in *dcl1*⁻ but not in *dcl2*⁻. One of the exons included in this class encodes a conserved protein that is involved in polarized growth (at the tip of the hypha), along with other proteins involved in the mitochondria metabolism and ribosome function, which could explain the abnormal hyphal morphology and lower growth rate described in *dcl1*⁻ mutants [47]. All ex-siRNAs are down regulated in the *ago-1*⁻ mutant, although only those of classes 1 and 2 are specifically bound to the Ago-1 protein, revealing the complexity of the esRNA biogenesis pathways in fungi [49].

3.2. *MicroRNA-like RNAs in N. crassa*

Analysis of esRNAs associated with the *N. crassa* QDE-2 protein identified several types of esRNAs, including those that share some similarities with conventional miRNAs from animals and plants [21]. These miRNA-like RNAs (milRNAs) are produced from stem-loop RNA precursors and most of them require Dicer in their biogenesis. This analysis also suggested that milRNAs could target endogenous RNA transcripts with imperfect complementarity, like regular miRNAs do. The biogenesis of these milRNAs, as with ex-siRNAs in *M. circinelloides*, shows a diverse use of the components of the RNAi machinery in order to produce four different classes of milRNAs (milR-1, milR-2, milR-3 and milR-4). The production of the first type of milRNA, milR-1, is completely dependent on Dicers, QDE-2 (but not its catalytic activity) and the exonuclease QIP. milR-2 milRNAs do not depend on Dicers but require QDE-2 and its catalytic activity. The production of milR-3 miRNAs only depends on Dicers activity, being the most similar pathway to plants miRNA synthesis. The biogenesis of the last milRNA class, milRNA-4, suggests the involvement of an unknown nuclease, as it is only partially dependent on Dicers. Reconstitution of the QDE-2-dependent milR-1 biogenesis *in vitro* has revealed the role of the RNA exosome, a 3' to 5' exonuclease complex, in determining the size of milR-1, demonstrating the importance of the exosome in esRNA processing [50]. Unlike in plants and animals, in which miRNAs are produced by RNA Pol II, the four major types of milRNAs identified in *N. crassa* are transcribed by the RNA Pol III, although Pol II was found to be associated with some milR loci, suggesting collaboration between the two polymerases in milRNAs production [51].

3.3. *Regulatory esRNAs in Magnaporthe oryzae*

The third study that analyzed the esRNAs content in a fungus by deep sequencing was carried out in *M. oryzae*, a model organism for the study of pathogen-host interactions in plants [31]. In this study, conversely to *N. crassa*, no putative miRNAs were found, and similarly to *M. circinelloides*, a profile composed of protein coding genes, intergenic regions and repetitive elements derived esRNAs was described. Interestingly, there were differences between the spectrum of esRNAs accumulated in vegetative and specialized-infection tissues. Whereas the esRNAs accumulated in vegetative mycelia were enriched for sequences that mapped to transposable elements, tRNA-derived fragments (tRFs) were the most abundant sRNA species identified in the appressorium, a specialized hypha that is involved in the invasion of the host plant cell. The specific presence of tRFs in the appressoria has been proposed to be part of a mechanism that restricts protein biosynthesis in order to direct cellular metabolism towards infection [31]. However, the biogenesis of tRFs is still unknown and there is no experimental evidence on the involvement of the RNAi machinery in their production, suggesting that they cannot be considered as *bona fide* esRNAs.

3.4. *siRNA-Mediated Regulation of Heterochromatin in Schizosaccharomyces pombe*

Regulation of heterochromatin formation depends on the components of the RNAi pathway in *S. pombe* [22]. Heterochromatin is transcriptionally inactive DNA with a highly condensed structure that can be found at three different regions of the *S. pombe* genome: centromeres, telomeres and mating-type loci. Some of these regions can lose their condensed structure and become

transcriptionally active. Later, the re-heterochromatinization of these regions occurs in an RNAi-dependent manner and is mediated by the production of specific siRNAs. The mechanism of RNAi in transcriptional silencing uses a similar machinery core to the post-transcriptional gene silencing mechanism. Thus, an RdRP is required to produce dsRNA from the nascent transcripts that are synthesized by RNA Pol II during the S phase of the cell cycle. The dsRNA is processed by Dicer into siRNAs that are then loaded onto *Ago1* in the so called RNA-induced transcriptional silencing (RITS) complex [52]. One difference between RITS and RISC is the presence in RITS of Chp1, a chromodomain-containing protein that binds the histone H3 [53]. Besides the binding properties of Chp1, RITS uses the guide strand to target the nascent transcript and reinforce the interaction with the region that is initiating heterochromatinization. Recruitment of RITS to these regions allows it to interact with other chromatin-modifying components, forming a major protein complex that spreads heterochromatinization through the region. In this interaction, RITS associates with the RNA-directed RNA polymerase complex (RDRC), a protein complex which contains an RNA polymerase, Rdrp1, a putative helicase termed Hrr1, and Cid12, a member of the Trf4 and Trf5 family [45]. The association between RITS and RDRC is siRNA and Clr4 dependent, which suggests that this association requires histone H3K9 methylation and occurs by a mechanism that involves tethering the nascent transcript to the chromatin. Several lines of evidence suggest that the siRNA guides RITS to target the nascent RNA transcript rather than the DNA complement strand [54]. This initiates the synthesis of dsRNA from the nascent transcript by the Rdrp1 present in RDRC, which leads to the processing of the resultant dsRNA by Dicer at the same location, as the Dicer enzyme is also required for the association between RITS and RDRC [45,55]. The activity of Dicer produces new siRNAs that amplifies the mechanism and helps to maintain and spread the heterochromatinization.

3.5. QDE-2-Interacting sRNAs Induced by DNA Damage

QDE-2-interacting sRNAs (qiRNAs) are a special type of esRNAs identified in *N. crassa* that are produced after treating this fungus with DNA-damaging agents [23]. The basic structure and biogenesis of qiRNAs are quite similar to regular siRNAs that are found interacting with Argonaute proteins. Thus, qiRNAs are 21–23 nt long, they require RdRP, Dicer and Argonaute proteins for their biogenesis, as well as the QDE-3 helicase, and they usually exhibit a 5' uridine, like regular siRNAs. The special feature of qiRNAs is their production from repetitive sequences of rDNA as a response to DNA damage. It was suggested that the triggering signal that activates the production of qiRNA is the synthesis of aRNA from damaged DNA sequences, in which double-stranded breaks and replication stress induce the formation of aberrant DNA structures. However, genetic screens designed to identify genes required for qiRNA biogenesis have revealed that homologous recombination is the only process required for qiRNA production, suggesting that *N. crassa* utilizes homologous recombination triggered by DNA damage to identify repetitive DNA loci [56]. The proposed model suggests that DNA damage promotes the formation of aberrant forms of recombination intermediates of repetitive DNA, which are recognized by QDE-3 and QDE-1 to produce aRNA and dsRNA [56]. The same mechanism has been suggested to operate in the initiation of silencing by repetitive transgenes, since homologous recombination is also required for quelling in *N. crassa*, indicating that quelling and qiRNA production share a common mechanism [56].

4. Conclusions

The kingdom Fungi has deeply contributed to our understanding of RNAi and its functions. Initially discovered as a defense mechanism against transposon and viral invasion, the RNAi mechanism has now emerged as a complex mechanism of gene regulation. Recent discoveries show several new sRNAs and different biogenesis pathways involved in a high diversity of new functions associated to the RNAi mechanism. Future studies on RNAi in fungi will reveal the whole picture of the role of these new regulatory sRNAs, enlightening the evolutionary origin of RNAi in eukaryotes.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Fire, A.; Xu, S.; Montgomery, M.K.; Kostas, S.A.; Driver, S.E.; Mello, C.C. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **1998**, *19*, 806–811.
2. Napoli, C.; Lemieux, C.; Jorgensen, R. Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *Plant Cell* **1990**, *2*, 279–289.
3. Romano, N.; Macino, G. Quelling, transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. *Mol. Microbiol.* **1992**, *6*, 3343–3353.
4. Cogoni, C.; Macino, G. Isolation of quelling-defective (qde) mutants impaired in posttranscriptional transgene-induced gene silencing in *Neurospora crassa*. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10233–10238.
5. 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**, *8*, 363–366.
6. Catalanotto, C.; Pallotta, M.; ReFalo, P.; Sachs, M.S.; Vayssie, L.; Macino, G.; Cogoni, C. Redundancy of the two dicer genes in transgene-induced posttranscriptional gene silencing in *Neurospora crassa*. *Mol. Cell. Biol.* **2004**, *24*, 2536–2545.
7. Catalanotto, C.; Azzalin, G.; Macino, G.; Cogoni, C. Involvement of small RNAs and role of the *qde* genes in the gene silencing pathway in *Neurospora*. *Genes Dev.* **2002**, *16*, 790–795.
8. Maiti, M.; Lee, H.C.; Liu, Y. QIP; a putative exonuclease; interacts with the *Neurospora* Argonaute protein and facilitates conversion of duplex siRNA into single strands. *Genes Dev.* **2007**, *21*, 590–600.
9. Cogoni, C.; Macino, G. Posttranscriptional gene silencing in *Neurospora* by a RecQ DNA helicase. *Science* **1999**, *17*, 2342–2344.
10. Pickford, A.; Braccini, L.; Macino, G.; Cogoni, C. The QDE-3 homologue RecQ-2 co-operates with QDE-3 in DNA repair in *Neurospora crassa*. *Curr. Genet.* **2003**, *42*, 220–227.
11. Lee, H.C.; Aalto, A.P.; Yang, Q.; Chang, S.S.; Huang, G.; Fisher, D.; Cha, J.; Poranen, M.M.; Bamford, D.H.; Liu, Y. The DNA/RNA-dependent RNA polymerase QDE-1 generates aberrant RNA and dsRNA for RNAi in a process requiring replication protein A and a DNA helicase. *PLoS Biol.* **2010**, *8*, e1000496.

12. Laurila, M.R.; Salgado, P.S.; Makeyev, E.V.; Nettelship, J.; Stuart, D.I.; Grimes, J.M.; Bamford, D.H. Gene silencing pathway RNA-dependent RNA polymerase of *Neurospora crassa*, yeast expression and crystallization of selenomethionated QDE-1 protein. *J. Struct. Biol.* **2005**, *1*, 111–115.
13. Drinnenberg, I.A.; Fink, G.R.; Bartel, D.P. Compatibility with killer explains the rise of RNAi-deficient fungi. *Science* **2011**, *333*, 1592.
14. Nicolas, F.E.; Torres-Martínez, S.; Ruiz-Vázquez, R.M. Loss and retention of RNA interference in fungi and parasites. *PLoS Pathog.* **2013**, *9*, e1003089.
15. Nicolas, F.E.; Torres-Martínez, S.; Ruiz-Vázquez, R.M. Two classes of small antisense RNAs in fungal RNA silencing triggered by non-integrative transgenes. *EMBO J.* **2003**, *22*, 3983–3991.
16. Nolan, T.; Braccini, L.; Azzalin, G.; de Toni, A.; Macino, G.; Cogoni, C. The post-transcriptional gene silencing machinery functions independently of DNA methylation to repress a LINE1-like retrotransposon in *Neurospora crassa*. *Nucleic Acids Res.* **2005**, *5*, 1564–1573.
17. Zhang, X.; Segers, G.C.; Sun, Q.; Deng, F.; Nuss, D.L. Characterization of hypovirus-derived small RNAs generated in the chestnut blight fungus by an inducible DCL-2-dependent pathway. *J. Virol.* **2008**, *6*, 2613–2619.
18. Hammond, T.M.; Xiao, H.; Boone, E.C.; Decker, L.M.; Lee, S.A.; Perdue, T.D.; Pukkila, P.J.; Shiu, P.K.T. Novel proteins required for meiotic silencing by unpaired DNA and siRNA generation in *Neurospora crassa*. *Genetics* **2013**, *1*, 91–100.
19. Wang, X.; Hsueh, Y.P.; Li, W.; Floyd, A.; Skalsky, R.; Heitman, J. Sex-induced silencing defends the genome of *Cryptococcus neoformans* via RNAi. *Genes Dev.* **2010**, *22*, 2566–2582.
20. Nicolas, F.E.; Moxon, S.; de Haro, J.P.; Calo, S.; Grigoriev, I.V.; Torres-Martínez, S.; Moulton, V.; Ruiz-Vázquez, R.M.; Dalmay, T. Endogenous short RNAs generated by Dicer 2 and RNA-dependent RNA polymerase 1 regulate mRNAs in the basal fungus *Mucor circinelloides*. *Nucleic Acids Res.* **2010**, *16*, 5535–5541.
21. Lee, H.C.; Li, L.; Gu, W.; Xue, Z.; Crosthwaite, S.K.; Pertsemlidis, A.; Lewis, Z.A.; Freitag, M.; Selker, E.U.; Mello, C.C.; *et al.* Diverse pathways generate microRNA-like RNAs and Dicer-independent small interfering RNAs in fungi. *Mol. Cell* **2010**, *6*, 803–814.
22. Volpe, T.A.; Kidner, C.; Hall, I.M.; Teng, G.; Grewal, S.I.; Martienssen, R.A. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **2002**, *13*, 1833–1837.
23. Lee, H.C.; Chang, S.S.; Choudhary, S.; Aalto, A.P.; Maiti, M.; Bamford, D.H.; Liu, Y. qiRNA is a new type of small interfering RNA induced by DNA damage. *Nature* **2009**, *459*, 274–277.
24. Hamada, W.; Spanu, P.D. Co-suppression of the hydrophobin gene Hcf-1 is correlated with antisense RNA biosynthesis in *Cladosporium fulvum*. *Mol. Gen. Genet.* **1998**, *6*, 630–638.
25. Schuur, T.A.; Schaeffer, E.A.; Wessels, J.G. Homology-dependent silencing of the SC3 gene in *Schizophyllum commune*. *Genetics* **1997**, *2*, 589–596.
26. Gutierrez, A.; Lopez-Garcia, S.; Garre, V. High reliability transformation of the basal fungus *Mucor circinelloides* by electroporation. *J. Microbiol. Methods* **2011**, *3*, 442–446.
27. Roncero, M.I.; Jepsen, L.P.; Stroman, P.; van Heeswijk, R. Characterization of a leuA gene and an ARS element from *Mucor circinelloides*. *Gene* **1989**, *2*, 335–343.
28. Buchon, N.; Vaury, C. RNAi, a defensive RNA-silencing against viruses and transposable elements. *Heredity* **2006**, *2*, 195–202.

29. Sijen, T.; Plasterk, R.H. Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature* **2003**, *20*, 310–314.
30. Murata, T.; Kadotani, N.; Yamaguchi, M.; Tosa, Y.; Mayama, S.; Nakayashiki, H. siRNA-dependent and -independent post-transcriptional cosuppression of the LTR-retrotransposon MAGGY in the phytopathogenic fungus *Magnaporthe oryzae*. *Nucleic Acids Res.* **2007**, *35*, 5987–5994.
31. Nunes, C.C.; Gowda, M.; Sailsbery, J.; Xue, M.; Chen, F.; Brown, D.E.; Oh, Y.Y.; Mitchell, T.K.; Dean, R.A. Diverse and tissue-enriched small RNAs in the plant pathogenic fungus; *Magnaporthe oryzae*. *BMC Genomics* **2011**, *12*, 288.
32. Harvey, J.J.; Lewsey, M.G.; Patel, K.; Westwood, J.; Heimstadt, S.; Carr, J.P.; Baulcombe, D.C. An antiviral defense role of AGO2 in plants. *PLoS One* **2011**, *6*, e14639.
33. Wilkins, C.; Dishongh, R.; Moore, S.C.; Whitt, M.A.; Chow, M.; Machaca, K. RNA interference is an antiviral defence mechanism in *Caenorhabditis elegans*. *Nature* **2005**, doi:10.1038/nature03957.
34. Zambon, R.A.; Vakharia, V.N.; Wu, L.P. RNAi is an antiviral immune response against a dsRNA virus in *Drosophila melanogaster*. *Cell. Microbiol.* **2006**, *8*, 880–889.
35. Jeang, K.T. RNAi in the regulation of mammalian viral infections. *BMC Biol.* **2012**, *10*, 58.
36. Sun, Q.; Choi, G.H.; Nuss, D.L. A single Argonaute gene is required for induction of RNA silencing antiviral defense and promotes viral RNA recombination. *Proc. Natl. Acad. Sci. USA* **2009**, *42*, 17927–17932.
37. Zhang, X.; Nuss, D.L. A host dicer is required for defective viral RNA production and recombinant virus vector RNA instability for a positive sense RNA virus. *Proc. Natl. Acad. Sci. USA* **2008**, *43*, 16749–16754.
38. Segers, G.C.; van Wezel, R.; Zhang, X.; Hong, Y.; Nuss, D.L. Hypovirus papain-like protease p29 suppresses RNA silencing in the natural fungal host and in a heterologous plant system. *Eukaryot. Cell* **2006**, *6*, 896–904.
39. Hammond, T.M.; Andrews, M.D.; Roossinck, M.J.; Keller, N.P. *Aspergillus* mycoviruses are targets and suppressors of RNA silencing. *Eukaryot. Cell* **2008**, *7*, 350–357.
40. Shiu, P.K.; Raju, N.B.; Zickler, D.; Metzzenberg, R.L. Meiotic silencing by unpaired DNA. *Cell* **2001**, *7*, 905–916.
41. Son, H.; Min, K.; Lee, J.; Raju, N.B.; Lee, Y.W. Meiotic silencing in the homothallic fungus *Gibberella zeae*. *Fungal Biol.* **2011**, *12*, 1290–1302.
42. Shiu, P.K.; Zickler, D.; Raju, N.B.; Ruprich-Robert, G.; Metzzenberg, R.L. SAD-2 is required for meiotic silencing by unpaired DNA and perinuclear localization of SAD-1 RNA-directed RNA polymerase. *Proc. Natl. Acad. Sci. USA* **2006**, *7*, 2243–2248.
43. Shiu, P.K.; Metzzenberg, R.L. Meiotic silencing by unpaired DNA, properties; regulation and suppression. *Genetics* **2002**, *4*, 1483–1495.
44. Hammond, T.M.; Xiao, H.; Boone, E.C.; Perdue, T.D.; Pukkila, P.J.; Shiu, P.K. SAD-3, a putative helicase required for meiotic silencing by unpaired DNA; interacts with other components of the silencing machinery. *G3 (Bethesda)* **2011**, *5*, 369–376.
45. Motamedi, M.R.; Verdel, A.; Colmenares, S.U.; Gerber, S.A.; Gygi, S.P.; Moazed, D. Two RNAi complexes; RITS and RDRC; physically interact and localize to noncoding centromeric RNAs. *Cell* **2004**, *6*, 789–802.

46. Bartel, D.P. MicroRNAs, genomics; biogenesis; mechanism; and function. *Cell* **2004**, *2*, 281–297.
47. Nicolas, F.E.; de Haro, J.P.; Torres-Martinez, S.; Ruiz-Vazquez, R.M. Mutants defective in a *Mucor circinelloides* dicer-like gene are not compromised in siRNA silencing but display developmental defects. *Fungal Genet. Biol.* **2007**, *6*, 504–516.
48. De Haro, J.P.; Calo, S.; Cervantes, M.; Nicolas, F.E.; Torres-Martinez, S.; Ruiz-Vazquez, R.M. A single dicer gene is required for efficient gene silencing associated with two classes of small antisense RNAs in *Mucor circinelloides*. *Eukaryot. Cell* **2009**, *10*, 1486–1497.
49. Cervantes, M.; Vila, A.; Nicolás, F.E.; Moxon, S.; de Haro, J.P.; Dalmay, T.; Torres-Martínez, S.; Ruiz-Vázquez, R.M. A single *argonaute* gene participates in exogenous and endogenous RNAi and controls cellular functions in the basal fungus *Mucor circinelloides*. *PLoS One* **2013**, doi:10.1371/journal.pone.0069283.
50. Zhang, Z.; Chang, S.S.; Zhang, Z.; Xue, Z.; Zhang, H.; Li, S.; Liu, Y. Homologous recombination as a mechanism to recognize repetitive DNA sequences in an RNAi pathway. *Genes Dev.* **2013**, *27*, 145–150.
51. Yang, Q.; Li, L.; Xue, Z.; Ye, Q.; Zhang, L.; Li, S.; Liu, Y. Transcription of the major *Neurospora crassa* microRNA-like small RNAs relies on RNA polymerase III. *PLoS Genet.* **2013**, *9*, e1003227.
52. Colmenares, S.U.; Buker, S.M.; Buhler, M.; Dlakic, M.; Moazed, D. Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. *Mol. Cell* **2007**, *3*, 449–461.
53. 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.
54. White, S.A.; Allshire, R.C. RNAi-mediated chromatin silencing in fission yeast. *Curr. Top. Microbiol. Immunol.* **2008**, *320*, 157–183.
55. Emmerth, S.; Schober, H.; Gaidatzis, D.; Roloff, T.; Jacobeit, K.; Buhler, M. Nuclear retention of fission yeast *dicer* is a prerequisite for RNAi-mediated heterochromatin assembly. *Dev. Cell* **2010**, *18*, 102–113.
56. Xue, Z.; Yuan, H.; Guo, J.; Liu, Y. Reconstitution of an Argonaute-dependent small RNA biogenesis pathway reveals a handover mechanism involving the RNA exosome and the exonuclease QIP. *Mol. Cell* **2012**, *46*, 299–310.

Reprinted from *IJMS*. Cite as: Michalova, E.; Vojtesek, B.; Hrstka, R. Impaired Pre-mRNA Processing and Altered Architecture of 3' Untranslated Regions Contribute to the Development of Human Disorders. *Int. J. Mol. Sci.* **2013**, *14*, 15861-15694.

Review

Impaired Pre-mRNA Processing and Altered Architecture of 3' Untranslated Regions Contribute to the Development of Human Disorders

Eva Michalova, Borivoj Vojtesek and Roman Hrstka *

Regional Centre for Applied Molecular Oncology, Masaryk Memorial Cancer Institute, Zlutý kopec 7, Brno 656 53, Czech Republic; E-Mails: michalova@mou.cz (E.M.); vojtesek@mou.cz (B.V.)

* Author to whom correspondence should be addressed; E-Mail: hrstka@mou.cz;
Tel.: +420-5-4313-3306; Fax: +420-5-4321-1169.

Received: 31 May 2013; in revised form: 21 June 2013 / Accepted: 24 June 2013 /

Published: 26 July 2013

Abstract: The biological fate of each mRNA and consequently, the protein to be synthesised, is highly dependent on the nature of the 3' untranslated region. Despite its non-coding character, the 3' UTR may affect the final mRNA stability, the localisation, the export from the nucleus and the translation efficiency. The conserved regulatory sequences within 3' UTRs and the specific elements binding to them enable gene expression control at the posttranscriptional level and all these processes reflect the actual state of the cell including proliferation, differentiation, cellular stress or tumourigenesis. Through this article, we briefly outline how the alterations in the establishment and final architecture of 3' UTRs may contribute to the development of various disorders in humans.

Keywords: 3' UTR; trinucleotide repeat expansion; polymorphism; mutation; cancer

1. Introduction

According to the central dogma of molecular biology, proteosynthesis proceeds from the genetic information carried by a DNA sequence through its transcription to RNA that functions as a template for polypeptide synthesis during the follow-up translation step [1]. Therefore, RNA might represent an interface between the coding DNA and the final protein. Recently it has also become known that

RNAs hold various functions in the cell and not all the RNAs are synthesised as mRNA templates for future polypeptides.

Transcription transfers the genetic information from a coding gene to a primary transcript—a precursor-mRNA (pre-mRNA). Normally, pre-mRNAs undergo a series of posttranscriptional modifications in the nucleus which bring forth the mRNA. These processing steps include capping of the 5' end, removal of introns by splicing, endonucleolytical cleavage and polyadenylation of the 3' end and editing. mRNAs are finally assembled by a central coding region which is translated to the final polypeptide and by non-coding regions at 5' and 3' ends that are not translated (untranslated regions; UTRs).

During the last decade, UTRs have been shown to harbour various sequence motifs (*cis*-acting elements, *cis*-elements) that in cooperation with specific binding proteins or RNAs (*trans*-acting elements, *trans*-elements) regulate the proper posttranscriptional modifications and proteosynthesis. Together with the pre-mRNA processing mechanism, these regulations play important roles in maintaining cellular functions reflecting cellular proliferation or differentiation and participating in cellular reactions to stress stimuli. Therefore, any alteration in the processing steps, mutation or polymorphism can influence the final character of the mRNA 3' end, the functional properties of the 3' UTR and the fate of the mRNA leading to variety of disorders in humans including cancer.

2. Processing of Pre-mRNA 3' End

2.1. Polyadenylation Signal

The cleavage and the addition of a poly(A) tail to the 3' end of a pre-mRNA are crucial for an efficient transcription termination, mRNA stability and export to the cytoplasm where the polypeptide chain is synthesised on the ribosomes [2,3]. The polyadenylation signal (PAS), a hexamer “AAUAAA” (less frequently—app. 15%—“AUUAAA”), located approximately 10–30 nucleotides (nt) upstream of the cleavage site, was identified as a highly conserved signal for the endonucleolytic cleavage at the 3' end [4]. Although additional sequences are recognised to stimulate the cleavage reaction [5], the PAS remains crucial for this process. Therefore, alterations in PAS sequence (SNP, insertions/deletions) disrupt the cleavage and polyadenylation steps resulting in various pathologies in humans, including an association with different malignancies. For instance, modified hexamers were found to inactivate gene expression in globin coding genes (“AAUAAG” in α 2-globin, “AACAAA” in β -globin) of thalassaemia patients [6,7]. Similarly, a transition “AAUAAA” \rightarrow “AAUGAA” within the PAS of *Foxp3* (forkhead box P3) transcript contributes to the IPEX syndrome (Immunodysregulation, Polyendocrinopathy, and Enteropathy, X-linked), a fatal autoimmune disease by reduced levels of Foxp3 transcription factor leading to the dysfunction of regulatory T cells [8]. The polymorphism of PAS in the human *N*-acetyltransferase coding gene (*NAT1*) influences the acetylation of carcinogens and administered drugs. Conversely, the T \rightarrow A transversion modifying the polyadenylation signal (“AAUAAA” \rightarrow “AAAAAA”) within the *NAT1*10* pre-mRNA does not significantly change the final protein level or the catalytic activity as shown previously in bladder and colon tissues [9]. However, the insertion of “AAA” to the 3' side of the PAS in the mutant allele *NAT1*16* results in a significant decrease of the protein level and of the catalytic activity measured *in vitro* suggesting that the cause lies in the disrupted secondary structure of the mRNA [10]. The human serotonin transporter

terminates the neurotransmission by the reuptake of serotonin and the alteration in the transporter coding gene (*hSERT*) is potentially involved in the development of the affective disorder [11]. Two different polyadenylation sites “AATGAA” and “AG/TTAAC” were identified within *hSERT*, and this polymorphism was considered to be playing role in this event, however it did not correlate with the susceptibility to the affective disorder [12]. Human lysosomal alpha-galactosidase A, the enzyme responsible for glycosphingolipid catabolism, bears the PAS within the coding sequence of the *GAL* gene and its mRNA thus lacks the 3' UTR [13]. The deletion of the “AA” dinucleotide within the PAS results in deficient enzymatic activity (residual or null) of the protein and in the development of Fabry disease, an inborn X-linked disorder characterised by the accumulation of globotriaosylceramide (GL-3), particularly, in vascular endothelial cells throughout the body [14]. Fabry disease patients develop vasculopathy and their life expectancy is shortened due to the renal insufficiency, cardiac disease and stroke [15,16].

2.2. Cleavage and Polyadenylation

Numerous proteins grouped into functional protein complexes participate in 3' end processing. The cleavage and polyadenylation specificity factor (CPSF) consists of five subunits: CPSF-160, -100, -73, -30 and hFip1. It recognises the PAS by the CPSF-160 subunit and catalyses the cleavage reaction through the CPSF-73 subunit [2,17]. Cleavage factors I and II recognise the additional sequence elements required for 3' end processing. They provide an interaction with poly(A) polymerase (PAP) and with nuclear poly(A) binding protein (PABPN1) and stimulate the cleavage reaction [18,19]. PAP catalyses the addition of a poly(A) tail under the control of PABPN1 [20].

The pre-mRNA 3' end is cleaved at the pA site, preferentially after the “CA” dinucleotide, however variations were observed, e.g., in the prothrombin gene (coagulation factor II, F2) where the cleavage normally occurs after the “CG” dinucleotide [21]. The cleavage reaction after the “CG” dinucleotide was observed to be less productive *in vitro* [22]. When mutation “CG”→“CA” occurs in F2 mRNA, this more effective 3' end processing leads to an increased concentration of F2 factor in plasma, resulting in a higher risk for thrombosis development [23].

The poly(A) tail comprising approximately 250 A-nucleotides in mammals is attached to the primary transcript at the cleaved pA site by a protein complex with PAP. The poly(A) sequence protects the 3' end against degrading exonucleases and enables the export of mRNA to the cytoplasm. It is also essential for the transcription termination, and together with the 5' cap and related binding proteins it enhances the translation process. The emerging poly(A) sequence is recognised and bound by PABPN1, which needs at least 27 A-nucleotides for a stable protection of the 3' end against nucleases [24]. PABPN1 is an ubiquitously expressed protein which binds to and regulates the processivity of PAP (that normally shows a low affinity for RNA substrates) and thus controls the elongation of poly(A) tail [25]. Normally, 10× “GCN” repeats coding for alanines within the N-terminus of the protein are present at the 5' end of the first exon of *PABPN1* gene. The expansion to 12–17× “GCN” repeats results in the synthesis of a misfolded protein that aggregates as filaments in nuclear inclusions in skeletal muscle fibres leading to cell death [26]. Patients with this expansion of a polyalanine stretch develop oculopharyngeal muscular dystrophy (OPMD), an autosomal dominant muscle disease. OPMD usually occurs at age fifty and manifests by eyelid drooping, mild

ophthalmoplegia, dysphagia and generalised muscle weakness and atrophy. Mechanistically, the deficit of functional PABPN1 affects other pre-mRNA processing steps: polyadenylation, mRNA export to cytoplasm and mRNA stability. Recently PABPN1 was identified to be involved in the regulation of alternative cleavage and polyadenylation (APA) [27]. It is becoming evident that a large portion of human genes contain multiple cleavage sites and PASs in their 3' UTRs generating multiple mRNA isoforms with different 3' UTRs [28–30]. The choice of an alternative pA site determines the length of the 3' UTR and furthermore the stability, localisation and translation efficiency of mRNA [31]. On the basis of the results of multiple studies, the length of 3' UTRs is inversely correlated with mRNA stability, gene expression and cellular proliferation. Mutated PABPN1 strongly correlates with usage of an alternative pA site and therefore deregulated gene expression [32]. An increased level of mutated protein causes cellular stress *in vitro*, enhances the expression of pro-apoptotic proteins and induces apoptosis in a p53 dependent manner [33].

3. Repeat Expansion Disorders

Human repeat expansion disorders represent an extensive heterogeneous group of diseases that are caused by the pathological expansion of repeats (mostly trinucleotide) in the coding or non-coding sequences of specific loci. The non-coding repeat expansions cause multisystem diseases and the extent of the repetitions often reflects disease severity and age of onset. The basis of the pathological mechanism of the non-coding repeats expansions thus lies at the posttranscriptional control of the gene expression and the resulting disorders thus mostly show similar aspects. As our review is focused on the role of 3' UTRs, we describe only selected disorders related to this topic.

3.1. Myotonic Dystrophy Type 1

Myotonic dystrophy type 1 (DM1) is an inherited autosomal dominant disease that is caused by the presence of multiple “CUG” repeats within the 3' UTR of myotonic dystrophy protein kinase (*DMPK*) pre-mRNA [34,35]. The number of “CUG” repeats strongly affects the final manifestation of the disease and the age-at-onset and varies from 5 to 35 repeats representing normal alleles, through 50–150 repeats in unaffected individuals or patients with mild to classical syndromes, up to more than 150 repeats in patients with severe DM1 [36]. When more than 1000 repeats are present, foetal development is disrupted and congenital DM1 occurs [37]. Moreover, a higher number of repeats implies an increased inherited instability of the mutant locus and further increase in copies transmitted across generations [38]

“CUG” transcripts were originally thought to cause DM1 pathogenesis when detected in a form of RNA foci in the nuclei of DM1 cells due to a blocked export of mutated transcripts to the cytoplasm [39]. However, a disrupted alternative splicing process in multiple pre-mRNAs was further observed and altered proteins were related to clinical symptoms of DM1 such as insulin resistance, myotonia muscle wasting, cardiac abnormalities or cognitive deficits [40]. The splicing alterations in multiple related transcripts result mainly from an imbalance in the levels of RNA-binding splicing factors CUGBP1 (CUG binding protein 1, also named CUGBP and ETR3-like factor 1, CELF1) and MBNL1 (muscleblind-like 1). There are three types of MBNL in mammals: MBNL1, 2 and 3, with MBNL1 being the best characterised to date. In DM1, CUGBP1 is hyper-phosphorylated by protein

kinase C (PKC) and its level increased, whereas MBNL1 and MBNL2 are bound abundantly to mutated transcripts and retained in the nucleus, co-localised with RNA foci and therefore are not accessible for a correct splicing process in other pre-mRNAs [41,42]. This “sequestration model” for MBNL is strongly supported by the results of the experiments in mouse models where the MBNL1 levels were increased after the transduction and the myotonia was reduced in HSA^{LR} poly(CUG) mice (human skeletal beta-actin long repeats = 250 “CUG” repeats) [43]. A similar effect manifested by the splicing aberration was observed *in vitro* in cells transfected with high concentrations of short synthetic oligoribonucleotides composed of “CUG” repeats [44]. The CUGBP1 protein interacts with different pre-mRNAs during muscle development and regulates the processing when phosphorylated by different kinases. The cyclin D3-cdk4/6 complex plays a key role in this process. In DM1 cells, the phosphorylation status of CUGBP1 and its interaction properties are modified due to a decreased level of cyclin D3 [45]. Jones *et al.* demonstrated that cyclin D3 is directed for proteasomal degradation by the GSK3 β -mediated phosphorylation when GSK3 β (glycogen synthase kinase 3 β) is stabilised by an increased autophosphorylation in DM1 cells induced by the presence of “CUG” transcripts [46]. Additionally, there is an increased and sequestered RNA binding protein hnRNP H, which regulates the alternative splicing leading to aberrant splicing of insulin receptor (INSR) pre-mRNA in DM1 myoblasts [47].

Despite the length of the “CTG” repeat region, its transcription is not blocked and the expanded “CUG” sequences form hairpins or longer dsRNA structures within the transcripts. These RNA secondary structures were shown to be digested *in vitro* with ribonuclease Dicer to shorter (CUG)_n sequences which may function as endogenous silencers (siRNA) of transcripts containing “CAG” repeats [48]. Moreover, the mutant transcripts interact with several transcription factors (SP1, STAT) disrupting cellular signalling and the transcription of target genes [49].

3.2. Huntington's Disease Like 2

The expansion of “CUG” repeats (up to 41 repeats compared to normal cells harbouring 6–28 repeats) within the alternative exon 2a of Junctophilin-3 (*JPH-3*) mRNA leads to the progression of Huntington's Disease Like 2 (HDL-2), an autosomal dominant disorder. The alternative splicing of exon 2a produces transcripts containing repetitions either within the coding region (translated into polyleucine or polyalanine tracts) or within the 3' UTR. HDL-2, similarly to Huntington's disease, manifests by motor defects, neurodegeneration and dementia [50,51]. Nevertheless, HDL-2 patients represent only a minor group of HD-like patients and are mostly Africans or with African remote ancestors [52]. The affected HDL-2 cells are characterised by the presence of RNA foci accumulating the *JPH-3* transcripts with bound MBNL1 resulting in an altered level of JPH-3 protein and in the neurodegeneration in the striatum [53]. Furthermore, a polyglutamine protein translated from antisense “(CAG)_n” transcripts and accumulated into nuclear inclusions contributes to the pathogenesis via neuronal dysfunction in a mouse model [54]. However, this phenomenon may play a minor role in the progression of HDL-2 and the polyglutamine expanded proteins do not need to be detected in brain samples. Therefore, the sole reduced level of Junctophilin-3, perhaps in association with other cellular processes, might seriously contribute to HDL-2 progression as shown in *JPH-3* knockout mice [55].

4. Altered mRNA Secondary Structure

Mutations within the 3' end might be responsible for the locally altered secondary structures of mRNAs and altered protein characteristics. GATA binding protein 4 (GATA4) is a transcription factor with zinc finger binding motifs that binds to the conserved “GATA” motifs within the promoter sequences of multiple genes [56]. It is essential for myocardial differentiation and function. Mutations found within the coding or non-coding regions lead to serious heart malformations, even death. Different mutations were detected within the 3' UTR and predicted to alter the secondary structure of mRNA potentially affecting transport and localisation of mRNA and reduced binding of GATA4 to target DNA sequences. These reduced transactivations by GATA4 are assumed to directly participate in the congenital heart disease phenotype [56]. Congenital adrenal hyperplasia (CAH) is usually caused by the deficiency of the 21-hydroxylase enzyme encoded by the *CYP21A2* gene. However, in the non-classical form of CAH no mutations in *CYP21A2* were identified, suggesting an involvement of non-coding regulatory regions. Indeed, the *13 G→A substitutions within the 3' UTR were identified in this subgroup of patients indicating that a substitution within the *CYP21A2* untranslated region is directly involved in the mild form of the disease probably, due to the altered secondary structure of pre-mRNA [57].

Modified regulatory characteristics of 3' UTRs also result from polymorphisms within the specific *cis*-elements or outside them, both leading to the altered secondary structures of mRNAs that affect the accessibility of target binding sites to interacting *trans*-elements [58]. The A→C polymorphism within the miR-155 binding site within the 3' UTR of *AGTR1* (Angiotensin II type 1 receptor) mRNA is responsible for a reduced binding of inhibitory miR-155 leading to an increased level of AGTR1 protein which is associated with hypertension. More than a weaker complementary sequence binding, the altered mRNA secondary structure and a lower accessibility to the miRNA target site may play role in this process [59]. The SNP rs9818870 (C→U polymorphism) in the *MRAS* 3' UTR is associated with a reduced level of M-ras protein in patients with the coronary artery disease (CAD). A higher accessibility of miR-195 to its target site within the *MRAS* 3' UTR and thus higher repression was shown in “U” allele due to a modified mRNA secondary structure even though the SNP is located within the sequence neighbouring the target site [59].

5. Cancer

Tumour development is generally accepted as a multistep process that involves genetic alterations resulting in the gain of oncogene activity and/or the loss of tumour suppressor gene function. Recent reports show that the alterations within 3' UTRs such as mutations leading to the loss of miRNA complementary sites or changes in the length may significantly influence the expression of many genes [60–62]. In addition, many other regulatory sequences (*cis*-elements) within 3' UTRs have been described to date and their alterations or incorrect interplay with specific binding proteins or RNAs (*trans*-elements) are also known to contribute to the malignant phenotype. Here we describe a few examples how variations in the pre-mRNA 3' end processing step are linked with certain aspects of cancer.

Multiple pA sites are present in human pre-mRNA 3' ends and an alternative processing represents another mechanism of the gene expression control [28,63]. The choice of a polyadenylation signal and

an alternative cleavage and polyadenylation at more proximal or more distal pA site create mRNA isoforms of different 3' UTR lengths that influence the expression levels mainly through the mRNA stability. Shorter 3' UTR isoforms are generally more stable and produce higher protein levels compared to longer isoforms most likely due to the loss of negative regulatory *cis*-elements and the limited binding sites for miRNAs. Therefore, shorter 3' UTR isoforms are more abundantly synthesised in proliferating and non-differentiated cells including cancer cells [31,64]. By analogy, transformed cells showing a high proliferation rate harbour short 3' UTR isoforms synthesised in highly expressed oncogenes and this correlates with a worse prognosis of the disease [28,61,65].

Three different mutation mechanisms (A insertion at position 1176; 3 base deletion at position 942 and partial duplication of A-rich sequence at position 970) were detected to create a new polyadenylation site within the *CCND1* pre-mRNA. This alternative processing produces shorter and more stable cyclin D1 isoforms in mantle cell lymphoma leading to a higher proliferation rate and shorter survival of patients [66]. The most famous tumour suppressor, p53 protein, and its alterations are frequently discussed in relation to malignant transformation [67]. A SNP “AAUAAA”→“AAUACA” (rs78378222) was identified in the polyadenylation signal of *TP53* pre-mRNA disrupting the pre-mRNA processing and leading to decreased levels of p53 mRNA [68,69]. This transversion was found associated with different cancer types: basal cell carcinoma, prostate cancer, glioma, colorectal adenoma [68], oesophageal squamous cell carcinoma (ESCC) [69], diffuse large B-cell lymphoma (DLBCL) [70] or glioma [71]. Although this proposed mechanism potentially affecting p53 expression can contribute to the malignant phenotype, the role of rs78378222 variant and its prognostic value must be further elucidated. Recently, an alternative polyadenylation of *CDC6* (cell division cycle 6) pre-mRNA was shown to be induced by 17 β -oestradiol in oestrogen receptor positive breast cancers cells, revealing new aspects for the posttranscriptional control of gene expression [72].

The poly(A) polymerase (PAP), a member of the multiprotein complex arranging pre-mRNA processing, is mainly responsible for the addition of a poly (A) tail to the cleaved 3' end. This step is crucial in the posttranscriptional regulation of the gene expression. As a higher rate of proteosynthesis is expected in proliferating cells, PAP was also shown to be more active in highly proliferating cells and furthermore more active in cancer cells compared to normal [73]. In this respect, the PAP activity level was shown to reflect the aggressiveness of breast carcinoma or leukaemia and represents an additional prognostic marker in breast carcinoma [74].

The adenylate uridylylate (AU-rich) elements (AREs) are the most common regulatory elements within 3' UTRs influencing the mRNA stability, the translation progress or the alternative RNA processing. AREs mostly destabilise mRNAs and repress the translation through a facilitation of the deadenylation process which results in the accelerated shortening of the poly(A) tail [75]. AREs may be disrupted by a mutation or reduced through the use of alternative PAS in many mRNAs whose elevated expression were related to proto-oncogene activation and a cancer progression. A proto-oncogene, *c-myc*, is constitutively expressed and the *c-myc* mRNA stabilised through a chromosomal translocation disrupting an ARE region in human plasma cell myeloma [76] and Burkitt lymphoma [77]. The enzyme cyclooxygenase 2 (COX-2) encoded by *PTGS2* (prostaglandin-endoperoxide synthase 2) gene is normally expressed at a low level or even absent because of an efficient regulation at the posttranscriptional level through a repressive ARE within

3' UTR. Normally, ARE directs *PTGS2* mRNA to a rapid decay, and only under a pro-inflammatory or a growth-associated stimuli the COX-2 is rapidly elevated for an increased prostaglandin formation. An enhanced stability of a shorter *PTGS2* mRNA isoform resulting from the alternative cleavage of 3' end and from the loss of repressive AREs was shown to be associated with chronic inflammatory diseases and colorectal cancer [78].

Certain types of cancer are known to be related to cellular transformation by oncogenic viruses. Human papilloma virus type 16 (HPV-16) infection belongs among the factors contributing to cervical cancer development and the viral transforming proteins E6 and E7 play key roles in carcinogenesis [79]. DNA of HPV-16 is often found integrated into the host genome in cancer cells [80,81]. Integrated viral DNAs were shown to express higher levels of E6/E7 mRNAs which were more stable than when expressed from extrachromosomal viral genomes. As the disruption of 3' UTR of early viral region was detected in an integrated HPV-16, a disruption of a potential ARE sequence within this AU-rich region is suggested to participate significantly in this phenomenon [82].

6. Future Perspectives

Through this article we aimed to explain the importance of 3' UTRs in the control of gene expression and show a tight relation of the 3' UTR establishment and integrity to a functional proteosynthesis and a healthy cell. As the age of “RNA research” is still at the beginning, many questions about the roles of mRNA sequence/structure relationships and related RNA-binding proteins in the regulation processes remain to be elucidated.

Formerly, the exploration of mechanisms regulating gene expression was mainly oriented towards promoter regions and to related binding transcription factors and the 3' untranslated regions, and their possible regulatory role at the posttranscriptional level were rather omitted. However, as many conserved sequences representing specific regulatory elements were discovered within the 3' end, the 3' UTRs were brought into focus during the last two decades. Different gene expression regulatory mechanisms related to 3' UTRs were described to date, taking effect during cellular proliferation, differentiation, stress conditions and carcinogenesis as well. The disruption of the posttranscriptional events including effector protein complexes mostly represents a primary cause of 3' UTR alteration, touching every step of the processing machinery. Secondly, mutations of regulatory sequences within 3' UTRs and in proteins or ncRNAs (non-coding RNAs) binding to these sequence also occur and lead to a range of disorders in humans. The spectrum of 3' UTR related disorders is wide and based on the results of recent studies may occur across all human diseases.

Acknowledgments

This work was supported by GACR P206/12/G151, GACR P301/11/1678, MH CZ-DRO (MMCI, 00209805) and by the European Regional Development Fund and the State Budget of the Czech Republic (RECAMO, CZ.1.05/2.1.00/03.0101).

Conflict of Interest

The authors declare no conflict of interest.

References

1. Crick, F. Central dogma of molecular biology. *Nature* **1970**, *227*, 561–563.
2. Zhao, J.; Hyman, L.; Moore, C. Formation of mrna 3' ends in eukaryotes: Mechanism, regulation, and interrelationships with other steps in mrna synthesis. *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 405–445.
3. Lemay, J.F.; Lemieux, C.; St-Andre, O.; Bachand, F. Crossing the borders: Poly(a)-binding proteins working on both sides of the fence. *RNA Biol.* **2010**, *7*, 291–295.
4. Proudfoot, N.J.; Brownlee, G.G. 3' non-coding region sequences in eukaryotic messenger rna. *Nature* **1976**, *263*, 211–214.
5. Matoulkova, E.; Michalova, E.; Vojtesek, B.; Hrstka, R. The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. *RNA Biol.* **2012**, *9*, 563–576.
6. Higgs, D.R.; Goodbourn, S.E.; Lamb, J.; Clegg, J.B.; Weatherall, D.J.; Proudfoot, N.J. Alpha-thalassaemia caused by a polyadenylation signal mutation. *Nature* **1983**, *306*, 398–400.
7. Serjeant, G.R.; Serjeant, B.E.; Fraser, R.A.; Hambleton, I.R.; Higgs, D.R.; Kulozik, A.E.; Donaldson, A. Hb s-beta-thalassemia: Molecular, hematological and clinical comparisons. *Hemoglobin* **2011**, *35*, 1–12.
8. Bennett, C.L.; Brunkow, M.E.; Ramsdell, F.; O'Briant, K.C.; Zhu, Q.; Fuleihan, R.L.; Shigeoka, A.O.; Ochs, H.D.; Chance, P.F. A rare polyadenylation signal mutation of the foxp3 gene (aauaaa→aaugaa) leads to the ipex syndrome. *Immunogenetics* **2001**, *53*, 435–439.
9. Bell, D.A.; Badawi, A.F.; Lang, N.P.; Ilett, K.F.; Kadlubar, F.F.; Hirvonen, A. Polymorphism in the n-acetyltransferase 1 (nat1) polyadenylation signal: Association of nat1*10 allele with higher n-acetylation activity in bladder and colon tissue. *Cancer Res.* **1995**, *55*, 5226–5229.
10. de Leon, J.H.; Vatsis, K.P.; Weber, W.W. Characterization of naturally occurring and recombinant human n-acetyltransferase variants encoded by nat1. *Mol. Pharmacol.* **2000**, *58*, 288–299.
11. Harmar, A.J.; Ogilvie, A.D.; Battersby, S.; Smith, C.A.; Blackwood, D.H.; Muir, W.J.; Fink, G.; Goodwin, G.M. The serotonin transporter gene and affective disorder. *Cold Spring Harb. Symp. Quant. Biol.* **1996**, *61*, 791–795.
12. Battersby, S.; Ogilvie, A.D.; Blackwood, D.H.; Shen, S.; Muqit, M.M.; Muir, W.J.; Teague, P.; Goodwin, G.M.; Harmar, A.J. Presence of multiple functional polyadenylation signals and a single nucleotide polymorphism in the 3' untranslated region of the human serotonin transporter gene. *J. Neurochem.* **1999**, *72*, 1384–1388.
13. Bishop, D.F.; Kornreich, R.; Desnick, R.J. Structural organization of the human alpha-galactosidase a gene: Further evidence for the absence of a 3' untranslated region. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 3903–3907.

14. Brady, R.O.; Gal, A.E.; Bradley, R.M.; Martensson, E.; Warshaw, A.L.; Laster, L. Enzymatic defect in fabry's disease. Ceramidetrihexosidase deficiency. *N. Engl. J. Med.* **1967**, *276*, 1163–1167.
15. MacDermot, K.D.; Holmes, A.; Miners, A.H. Natural history of fabry disease in affected males and obligate carrier females. *J. Inherit Metab. Dis.* **2001**, *24*, 13–14.
16. Vedder, A.C.; Linthorst, G.E.; van Breemen, M.J.; Groener, J.E.; Bemelman, F.J.; Strijland, A.; Mannens, M.M.; Aerts, J.M.; Hollak, C.E. The dutch fabry cohort: Diversity of clinical manifestations and gb3 levels. *J. Inherit Metab. Dis.* **2007**, *30*, 68–78.
17. Tian, B.; Manley, J.L. Alternative cleavage and polyadenylation: The long and short of it. *Trends Biochem. Sci.* **2013**, *38*, 312–320.
18. Brown, K.M.; Gilmartin, G.M. A mechanism for the regulation of pre-mrna 3' processing by human cleavage factor im. *Mol. Cell* **2003**, *12*, 1467–1476.
19. de Vries, H.; Ruegsegger, U.; Hubner, W.; Friedlein, A.; Langen, H.; Keller, W. Human pre-mrna cleavage factor ii(m) contains homologs of yeast proteins and bridges two other cleavage factors. *EMBO J.* **2000**, *19*, 5895–5904.
20. Bienroth, S.; Keller, W.; Wahle, E. Assembly of a processive messenger rna polyadenylation complex. *The EMBO J.* **1993**, *12*, 585–594.
21. Gehring, N.H.; Frede, U.; Neu-Yilik, G.; Hundsdoerfer, P.; Vetter, B.; Hentze, M.W.; Kulozik, A.E. Increased efficiency of mrna 3' end formation: A new genetic mechanism contributing to hereditary thrombophilia. *Nature Genet.* **2001**, *28*, 389–392.
22. Chen, F.; MacDonald, C.C.; Wilusz, J. Cleavage site determinants in the mammalian polyadenylation signal. *Nucleic Acids Res.* **1995**, *23*, 2614–2620.
23. Poort, S.R.; Rosendaal, F.R.; Reitsma, P.H.; Bertina, R.M. A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood* **1996**, *88*, 3698–3703.
24. Steinman, R.A. Mrna stability control: A clandestine force in normal and malignant hematopoiesis. *Leukemia* **2007**, *21*, 1158–1171.
25. Kerwitz, Y.; Kuhn, U.; Lilie, H.; Knoth, A.; Scheuermann, T.; Friedrich, H.; Schwarz, E.; Wahle, E. Stimulation of poly(a) polymerase through a direct interaction with the nuclear poly(a) binding protein allosterically regulated by rna. *EMBO J.* **2003**, *22*, 3705–3714.
26. Brais, B.; Bouchard, J.P.; Xie, Y.G.; Rochefort, D.L.; Chretien, N.; Tome, F.M.; Lafreniere, R.G.; Rommens, J.M.; Uyama, E.; Nohira, O.; *et al.* Short gcg expansions in the pabp2 gene cause oculopharyngeal muscular dystrophy. *Nature Genet.* **1998**, *18*, 164–167.
27. Jenal, M.; Elkon, R.; Loayza-Puch, F.; van Haaften, G.; Kuhn, U.; Menzies, F.M.; Oude Vrielink, J.A.; Bos, A.J.; Drost, J.; Rooijers, K.; *et al.* The poly(a)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* **2012**, *149*, 538–553.
28. Fu, Y.; Sun, Y.; Li, Y.; Li, J.; Rao, X.; Chen, C.; Xu, A. Differential genome-wide profiling of tandem 3' utrs among human breast cancer and normal cells by high-throughput sequencing. *Genome Res.* **2011**, *21*, 741–747.
29. Ozsolak, F.; Kapranov, P.; Foissac, S.; Kim, S.W.; Fishilevich, E.; Monaghan, A.P.; John, B.; Milos, P.M. Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell* **2010**, *143*, 1018–1029.

30. Shepard, P.J.; Choi, E.A.; Lu, J.; Flanagan, L.A.; Hertel, K.J.; Shi, Y. Complex and dynamic landscape of rna polyadenylation revealed by pas-seq. *RNA* **2011**, *17*, 761–772.
31. 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.
32. de Klerk, E.; Venema, A.; Anvar, S.Y.; Goeman, J.J.; Hu, O.; Trollet, C.; Dickson, G.; den Dunnen, J.T.; van der Maarel, S.M.; Raz, V.; *et al.* Poly(a) binding protein nuclear 1 levels affect alternative polyadenylation. *Nucleic Acids Res.* **2012**, *40*, 9089–9101.
33. Bhattacharjee, R.B.; Zannat, T.; Bag, J. Expression of the polyalanine expansion mutant of nuclear poly(a)-binding protein induces apoptosis via the p53 pathway. *Cell Biol. Int.* **2012**, *36*, 697–704.
34. 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.
35. 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.
36. Kamsteeg, E.J.; Kress, W.; Catalli, C.; Hertz, J.M.; Witsch-Baumgartner, M.; Buckley, M.F.; van Engelen, B.G.; Schwartz, M.; Scheffer, H. Best practice guidelines and recommendations on the molecular diagnosis of myotonic dystrophy types 1 and 2. *Eur. J. Hum. Genet.* **2012**, *20*, 1203–1208.
37. Salehi, L.B.; Bonifazi, E.; Stasio, E.D.; Gennarelli, M.; Botta, A.; Vallo, L.; Iraci, R.; Massa, R.; Antonini, G.; Angelini, C.; *et al.* Risk prediction for clinical phenotype in myotonic dystrophy type 1: Data from 2,650 patients. *Genet. Test.* **2007**, *11*, 84–90.
38. Savic Pavicevic, D.; Miladinovic, J.; Brkusanin, M.; Svikovic, S.; Djurica, S.; Brajuskovic, G.; Romac, S. Molecular genetics and genetic testing in myotonic dystrophy type 1. *BioMed Res. Int.* **2013**, doi:10.1155/2013/391821.
39. Taneja, K.L.; McCurrach, M.; Schalling, M.; Housman, D.; Singer, R.H. Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J. Cell Biol.* **1995**, *128*, 995–1002.
40. Day, J.W.; Ranum, L.P. Rna pathogenesis of the myotonic dystrophies. *Neuromuscul. Disord.* **2005**, *15*, 5–16.
41. Jiang, H.; Mankodi, A.; Swanson, M.S.; Moxley, R.T.; Thornton, C.A. Myotonic dystrophy type 1 is associated with nuclear foci of mutant rna, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum. Mol. Genet.* **2004**, *13*, 3079–3088.
42. Kuyumcu-Martinez, N.M.; Wang, G.S.; Cooper, T.A. Increased steady-state levels of cugbp1 in myotonic dystrophy 1 are due to pkc-mediated hyperphosphorylation. *Mol. Cell* **2007**, *28*, 68–78.
43. Kanadia, R.N.; Shin, J.; Yuan, Y.; Beattie, S.G.; Wheeler, T.M.; Thornton, C.A.; Swanson, M.S. Reversal of rna missplicing and myotonia after muscleblind overexpression in a mouse poly(cug) model for myotonic dystrophy. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 11748–11753.

44. Mykowska, A.; Sobczak, K.; Wojciechowska, M.; Kozlowski, P.; Krzyzosiak, W.J. Cag repeats mimic cug repeats in the misregulation of alternative splicing. *Nucleic Acids Res.* **2011**, *39*, 8938–8951.
45. Salisbury, E.; Sakai, K.; Schoser, B.; Huichalaf, C.; Schneider-Gold, C.; Nguyen, H.; Wang, G.L.; Albrecht, J.H.; Timchenko, L.T. Ectopic expression of cyclin d3 corrects differentiation of dm1 myoblasts through activation of rna cug-binding protein, cugbp1. *Exp. Cell Res.* **2008**, *314*, 2266–2278.
46. Jones, K.; Wei, C.; Iakova, P.; Bugiardini, E.; Schneider-Gold, C.; Meola, G.; Woodgett, J.; Killian, J.; Timchenko, N.A.; Timchenko, L.T. Gsk3beta mediates muscle pathology in myotonic dystrophy. *J. Clin. Investig.* **2012**, *122*, 4461–4472.
47. Paul, S.; Dansithong, W.; Kim, D.; Rossi, J.; Webster, N.J.; Comai, L.; Reddy, S. Interaction of muscleblind, cug-bp1 and hnnp h proteins in dm1-associated aberrant ir splicing. *EMBO J.* **2006**, *25*, 4271–4283.
48. Krol, J.; Fiszer, A.; Mykowska, A.; Sobczak, K.; de Mezer, M.; Krzyzosiak, W.J. Ribonuclease dicer cleaves triplet repeat hairpins into shorter repeats that silence specific targets. *Mol. Cell* **2007**, *25*, 575–586.
49. Ebralidze, A.; Wang, Y.; Petkova, V.; Ebralidse, K.; Junghans, R.P. Rna leaching of transcription factors disrupts transcription in myotonic dystrophy. *Science* **2004**, *303*, 383–387.
50. Greenstein, P.E.; Vonsattel, J.P.; Margolis, R.L.; Joseph, J.T. Huntington's disease like-2 neuropathology. *Mov. Disord.* **2007**, *22*, 1416–1423.
51. Margolis, R.L.; O'Hearn, E.; Rosenblatt, A.; Willour, V.; Holmes, S.E.; Franz, M.L.; Callahan, C.; Hwang, H.S.; Troncoso, J.C.; Ross, C.A. A disorder similar to huntington's disease is associated with a novel cag repeat expansion. *Ann. Neurol.* **2001**, *50*, 373–380.
52. Paradisi, I.; Ikonomu, V.; Arias, S. Huntington disease-like 2 (hdl2) in venezuela: Frequency and ethnic origin. *J. Hum. Genet.* **2013**, *58*, 3–6.
53. Rudnicki, D.D.; Holmes, S.E.; Lin, M.W.; Thornton, C.A.; Ross, C.A.; Margolis, R.L. Huntington's disease-like 2 is associated with cug repeat-containing rna foci. *Ann. Neurol.* **2007**, *61*, 272–282.
54. Wilburn, B.; Rudnicki, D.D.; Zhao, J.; Weitz, T.M.; Cheng, Y.; Gu, X.; Greiner, E.; Park, C.S.; Wang, N.; Sopher, B.L.; *et al.* An antisense cag repeat transcript at jph3 locus mediates expanded polyglutamine protein toxicity in huntington's disease-like 2 mice. *Neuron* **2011**, *70*, 427–440.
55. Seixas, A.I.; Holmes, S.E.; Takeshima, H.; Pavlovich, A.; Sachs, N.; Pruitt, J.L.; Silveira, I.; Ross, C.A.; Margolis, R.L.; Rudnicki, D.D. Loss of junctophilin-3 contributes to huntington disease-like 2 pathogenesis. *Ann. Neurol.* **2012**, *71*, 245–257.
56. Reamon-Buettner, S.M.; Cho, S.H.; Borlak, J. Mutations in the 3'-untranslated region of gata4 as molecular hotspots for congenital heart disease (chd). *BMC Med.Genet.* **2007**, *8*, 38.
57. Menabo, S.; Balsamo, A.; Baldazzi, L.; Barbaro, M.; Nicoletti, A.; Conti, V.; Pirazzoli, P.; Wedell, A.; Cicognani, A. A sequence variation in 3' utr of cyp21a2 gene correlates with a mild form of congenital adrenal hyperplasia. *J. Endocrinol. Investig.* **2012**, *35*, 298–305.
58. Abrahams, Y.; Laguette, M.J.; Prince, S.; Collins, M. Polymorphisms within the col5a1 3'-utr that alters mrna structure and the mir608 gene are associated with achilles tendinopathy. *Ann. Hum. Genet.* **2013**, *77*, 204–214.

59. Haas, U.; Sczakiel, G.; Laufer, S.D. MicroRNA-mediated regulation of gene expression is affected by disease-associated snps within the 3'-utr via altered rna structure. *RNA Biol.* **2012**, *9*, 924–937.
60. Lee, Y.S.; Dutta, A. The tumor suppressor microRNA let-7 represses the hmga2 oncogene. *Genes Dev.* **2007**, *21*, 1025–1030.
61. 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.
62. Mayr, C.; Hemann, M.T.; Bartel, D.P. Disrupting the pairing between let-7 and hmga2 enhances oncogenic transformation. *Science* **2007**, *315*, 1576–1579.
63. 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.
64. 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.
65. Lembo, A.; Di Cunto, F.; Provero, P. Shortening of 3'utrs correlates with poor prognosis in breast and lung cancer. *PLoS One* **2012**, doi:10.1371/journal.pone.0031129.
66. Wiestner, A.; Tehrani, M.; Chiorazzi, M.; Wright, G.; Gibellini, F.; Nakayama, K.; Liu, H.; Rosenwald, A.; Muller-Hermelink, H.K.; Ott, G.; *et al.* Point mutations and genomic deletions in ccnd1 create stable truncated cyclin d1 mrnas that are associated with increased proliferation rate and shorter survival. *Blood* **2007**, *109*, 4599–4606.
67. Hrstka, R.; Coates, P.J.; Vojtesek, B. Polymorphisms in p53 and the p53 pathway: Roles in cancer susceptibility and response to treatment. *J. Cell. Mol. Med.* **2009**, *13*, 440–453.
68. Stacey, S.N.; Sulem, P.; Jonasdottir, A.; Masson, G.; Gudmundsson, J.; Gudbjartsson, D.F.; Magnusson, O.T.; Gudjonsson, S.A.; Sigurgeirsson, B.; Thorisdottir, K.; *et al.* A germline variant in the tp53 polyadenylation signal confers cancer susceptibility. *Nature Genet.* **2011**, *43*, 1098–1103.
69. Zhou, L.; Yuan, Q.; Yang, M. A functional germline variant in the p53 polyadenylation signal and risk of esophageal squamous cell carcinoma. *Gene* **2012**, *506*, 295–297.
70. Li, Y.; Gordon, M.W.; Xu-Monette, Z.Y.; Visco, C.; Tzankov, A.; Zou, D.; Qiu, L.; Montes-Moreno, S.; Dybkaer, K.; Orazi, A.; *et al.* Single nucleotide variation in the tp53 3' untranslated region in diffuse large b-cell lymphoma treated with rituximab-chop: A report from the international dlbcl rituximab-chop consortium program. *Blood* **2013**, *121*, 4529–4540.
71. Enciso-Mora, V.; Hosking, F.J.; Di Stefano, A.L.; Zelenika, D.; Shete, S.; Broderick, P.; Idbaih, A.; Delattre, J.Y.; Hoang-Xuan, K.; Marie, Y.; *et al.* Low penetrance susceptibility to glioma is caused by the tp53 variant rs78378222. *Br. J. Cancer* **2013**, *108*, 2178–2185.
72. Akman, B.H.; Can, T.; Erson-Bensan, A.E. Estrogen-induced upregulation and 3'-utr shortening of cdc6. *Nucleic Acids Res.* **2012**, *40*, 10679–10688.
73. Kazazoglou, T.; Tsiapalis, C.M.; Havredaki, M. Polyadenylate polymerase activity in stationary and growing cell cultures. *Exp. Cell Biol.* **1987**, *55*, 164–172.
74. Scorilas, A.; Talieri, M.; Ardavanis, A.; Courtis, N.; Dimitriadis, E.; Yotis, J.; Tsiapalis, C.M.; Trangas, T. Polyadenylate polymerase enzymatic activity in mammary tumor cytosols: A new independent prognostic marker in primary breast cancer. *Cancer Res.* **2000**, *60*, 5427–5433.

75. Xu, N.; Chen, C.Y.; Shyu, A.B. Modulation of the fate of cytoplasmic mrna by au-rich elements: Key sequence features controlling mrna deadenylation and decay. *Mol. Cell. Biol.* **1997**, *17*, 4611–4621.
76. Hollis, G.F.; Gazdar, A.F.; Bertness, V.; Kirsch, I.R. Complex translocation disrupts c-myc regulation in a human plasma cell myeloma. *Mol. Cell. Biol.* **1988**, *8*, 124–129.
77. Eick, D.; Piechaczyk, M.; Henglein, B.; Blanchard, J.M.; Traub, B.; Kofler, E.; Wiest, S.; Lenoir, G.M.; Bornkamm, G.W. Aberrant c-myc mnas of burkitt's lymphoma cells have longer half-lives. *EMBO J.* **1985**, *4*, 3717–3725.
78. Young, L.E.; Dixon, D.A. Posttranscriptional regulation of cyclooxygenase 2 expression in colorectal cancer. *Curr. Colorectal Cancer Rep.* **2010**, *6*, 60–67.
79. zur Hausen, H. Human papillomaviruses and their possible role in squamous cell carcinomas. *Curr. Top. Microbiol. Immunol.* **1977**, *78*, 1–30.
80. Schwarz, E.; Freese, U.K.; Gissmann, L.; Mayer, W.; Roggenbuck, B.; Stremlau, A.; zur Hausen, H. Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. *Nature* **1985**, *314*, 111–114.
81. Yee, C.; Krishnan-Hewlett, I.; Baker, C.C.; Schlegel, R.; Howley, P.M. Presence and expression of human papillomavirus sequences in human cervical carcinoma cell lines. *Am. J. Pathol.* **1985**, *119*, 361–366.
82. Jeon, S.; Lambert, P.F. Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of e6 and e7 mnas: Implications for cervical carcinogenesis. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 1654–1658.

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 978-3-03842-011-8
www.mdpi.com

