

GCAT
TACG
GCAT

genes

microRNA Regulation in Health and Disease

Edited by
Clifford J. Steer and Subbaya Subramanian
Printed Edition of the Special Issue Published in *Genes*

microRNA Regulation in Health and Disease

microRNA Regulation in Health and Disease

Special Issue Editors

Clifford J. Steer

Subbaya Subramanian

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade



Special Issue Editors

Clifford J. Steer
University of Minnesota
Medical School
USA

Subbaya Subramanian
University of Minnesota
Medical School
USA

Editorial Office

MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Genes* (ISSN 2073-4425) from 2018 to 2019 (available at: https://www.mdpi.com/journal/genes/special-issues/miRNA_Health_Disease).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. <i>Journal Name</i> Year , Article Number, Page Range.

ISBN 978-3-03921-714-4 (Pbk)

ISBN 978-3-03921-715-1 (PDF)

Cover image courtesy of Subaya Subramanian.

Colon Organoids: MicroRNAs are critical players in organ development. Organoids serve as a bridge in understanding the role of microRNAs in organ development and disease progression.

© 2020 by the authors. Articles in this book are Open Access and distributed under the Creative Commons Attribution (CC BY) license, which allows users to download, copy and build upon published articles, as long as the author and publisher are properly credited, which ensures maximum dissemination and a wider impact of our publications.

The book as a whole is distributed by MDPI under the terms and conditions of the Creative Commons license CC BY-NC-ND.

Contents

About the Special Issue Editors	vii
Subbaya Subramanian and Clifford J. Steer Special Issue: MicroRNA Regulation in Health and Disease Reprinted from: <i>Genes</i> 2019 , <i>10</i> , 457, doi:10.3390/genes10060457	1
Ce Yuan, Clifford J. Steer and Subbaya Subramanian Host–MicroRNA–Microbiota Interactions in Colorectal Cancer Reprinted from: <i>Genes</i> 2019 , <i>10</i> , 270, doi:10.3390/genes10040270	5
Xiao Li, Youbo Zhu, Huifang Zhang, Guangjun Ma, Guofang Wu, Aoqi Xiang, Xin'E. Shi, Gong She Yang and Shiduo Sun MicroRNA-106a-5p Inhibited C2C12 Myogenesis via Targeting PIK3R1 and Modulating the PI3K/AKT Signaling Reprinted from: <i>Genes</i> 2018 , <i>9</i> , 333, doi:10.3390/genes9070333	18
Kristina B. V. Døssing, Christina Kjær, Jonas Vikeså, Tina Binderup, Ulrich Knigge, Michael D. Culler, Andreas Kjær, Birgitte Federspiel and Lennart Friis-Hansen Somatostatin Analogue Treatment Primarily Induce miRNA Expression Changes and Up-Regulates Growth Inhibitory miR-7 and miR-148a in Neuroendocrine Cells Reprinted from: <i>Genes</i> 2018 , <i>9</i> , 337, doi:10.3390/genes9070337	31
Chang Liu, Ze Chen, Yue Hu, Haishuo Ji, Deshui Yu, Wenyuan Shen, Siyu Li, Jishou Ruan, Wenjun Bu and Shan Gao Complemented Palindromic Small RNAs First Discovered from SARS Coronavirus Reprinted from: <i>Genes</i> 2018 , <i>9</i> , 442, doi:10.3390/genes9090442	48
Markus Axmann, Sabine M. Meier, Andreas Karner, Witta Strobl, Herbert Stangl and Birgit Plochberger Serum and Lipoprotein Particle miRNA Profile in Uremia Patients Reprinted from: <i>Genes</i> 2018 , <i>9</i> , 533, doi:10.3390/genes9110533	59
Kadri Rekker, Signe Altmæ, Marina Suhorutshenko, Maire Peters, Juan F. Martinez-Blanch, Francisco M. Codoñer, Felipe Vilella, Carlos Simón, Andres Salumets and Agne Velthut-Meikas A Two-Cohort RNA-seq Study Reveals Changes in Endometrial and Blood miRNome in Fertile and Infertile Women Reprinted from: <i>Genes</i> 2018 , <i>9</i> , 574, doi:10.3390/genes9120574	76
Steven Mullenbrock, Fei Liu, Suzanne Szak, Xiaoping Hronowski, Benbo Gao, Peter Juhasz, Chao Sun, Mei Liu, Helen McLaughlin, Qiurong Xiao, Carol Feghali-Bostwick and Timothy S. Zheng Systems Analysis of Transcriptomic and Proteomic Profiles Identifies Novel Regulation of Fibrotic Programs by miRNAs in Pulmonary Fibrosis Fibroblasts Reprinted from: <i>Genes</i> 2018 , <i>9</i> , 588, doi:10.3390/genes9120588	93
Miguel Hueso, Josep M. Cruzado, Joan Torras and Estanis Navarro An Exonic Switch Regulates Differential Accession of microRNAs to the Cd34 Transcript in Atherosclerosis Progression Reprinted from: <i>Genes</i> 2019 , <i>10</i> , 70, doi:10.3390/genes10010070	108

**Gabrielle Bradshaw, Larisa M. Haupt, Eunise M. Aquino, Rodney A. Lea,
Heidi G. Sutherland and Lyn R. Griffiths**
Single Nucleotide Polymorphisms in *MIR143* Contribute to Protection against Non-Hodgkin
Lymphoma (NHL) in Caucasian Populations
Reprinted from: *Genes* **2019**, *10*, 185, doi:10.3390/genes10030185 **124**

About the Special Issue Editors

Clifford J. Steer is Professor of Medicine and Genetics, Cell Biology, and Development at the University of Minnesota Medical School, Minneapolis, MN. He has been active in the field of biomedical research for over four decades. In that capacity, he has been a long-standing member of several National Institutes of Health Study Sections. He has been Co-Editor of a major scientific journal in liver diseases and presently serves on the Editorial Boards of three journals. He has published his work in top-ranking scientific journals and numerous textbooks. Steer's areas of research over the last decade have included gene therapy, liver regeneration, neurodegenerative disorders, and microRNA regulation of gene function. He has published over 290 scientific articles and has organized and chaired many national and international scientific conferences. In recognition of his work, he has received the Holloman Award in Biotechnology in 1998; the Thorne Stroke Award and Department of Medicine Faculty Award for Outstanding Research, both in 2004; and the Senior Investigator Award from the University of Minnesota Medical School in 2012. He was elected into the American Society for Clinical Investigation in 1991 and, in 2014, was made an inaugural Fellow of the American Association for the Study of Liver Diseases. His work has featured in newspapers around the world, *Time* magazine, and in 1998, was featured in the *Village Voice*. His current major research focus is on the development of human livers in pigs to address the enormous shortage of organs for transplantation—an ultimate example of individualized and precision medicine.

Subbaya Subramanian research is in the fields of cancer biology and novel therapeutic development. At the University of Minnesota, Subramanian has established an internationally recognized research program focused on understanding microRNA regulatory networks in cancer development and progression. His current research focuses on understanding the interactions between the immune system and cancer and how cancer cells manipulate the antitumor immune response. Dr. Subramanian has authored over 100 peer-reviewed publications and has written numerous book chapters and review articles. He is the Section Editor-in-Chief of *Vaccines*, Associate Editor of *Frontiers in Genetics and Molecular Biosciences*, and serves on the Editorial Board of several journals, including *Scientific Reports*.

Editorial

Special Issue: MicroRNA Regulation in Health and Disease

Subbaya Subramanian ^{1,*} and Clifford J. Steer ^{2,3,*}¹ Department of Surgery, University of Minnesota Medical School, Minneapolis, MN 55455, USA² Department of Medicine, University of Minnesota Medical School, Minneapolis, MN 55455, USA³ Department of Genetics, Cell Biology and Development, University of Minnesota Medical School, Minneapolis, MN 55455, USA

* Correspondence: subree@umn.edu (S.S.); steer001@umn.edu (C.J.S.)

Received: 6 June 2019; Accepted: 11 June 2019; Published: 15 June 2019

Our understanding of non-coding RNA has significantly changed based on recent advances in genomics and molecular biology, and their role is recognized to include far more than a link between the sequence of DNA and synthesized proteins. MicroRNAs (miRNAs) are small regulatory RNAs that play a crucial role in posttranscriptional gene regulation. Greater than 2500 miRNAs have been identified and catalogued in humans and many of them are conserved in other species [1]. miRNAs are implicated in almost every facet of fundamental cellular functions including development, senescence and disease. The past decade has experienced a remarkable increase in the understanding of miRNA biogenesis, their target genes, miRNA biomarkers and potential therapeutics for a growing number of disease conditions. RNA-based markers and therapeutics have a potentially significant clinical impact, and many of the miRNA-based therapies are at various stages of application and human clinical trial [2].

As background, non-coding RNAs are divided into (i) transcription RNAs (including both tRNA and rRNA); (ii) small RNAs, which are further subdivided into siRNAs, miRNAs, snoRNAs, and snRNAs; and (iii) most recently, long non-coding RNAs, which are now known to transcribe short peptides [3]. MicroRNAs are single-stranded non-coding RNAs that are typically 18–25 nucleotides (nts) in length and are best known for their role in the post-transcriptional regulation of gene expression. They are the most abundant class of small endogenous non-protein coding RNAs, and make up one of the largest well-conserved gene families found among viruses, plants and animals. The majority of miRNA sequences in humans are typically transcribed by introns of non-coding and coding transcripts, with few transcribed by exonic regions.

MicroRNA genes are typically transcribed by polymerase II or III and generate primary miRNAs (pri-miRNAs), which can contain sequences for multiple miRNAs and be hundreds of nts in length. These structures are then processed and cleaved by Drosha-DGCR8 complex, resulting in the formation of a hairpin-shaped stem-loop structure, which is known as the precursor miRNA (pre-miRNA) and typically around 70 nts in length. The pre-miRNA is exported outside of the nucleus primarily by exportin 5. Further processing takes place in the cytoplasm by Dicer1-TARBP2, which is an RNase III enzyme, resulting in a two-stranded duplex of miRNA-miRNA*. Typically, it is 18–25 nt long, with one strand designated the guide strand and the other as the passenger strand. Finally, the guide strand is incorporated into the RNA-induced silencing complex (RISC), which is a large multiprotein miRNA ribonucleoprotein complex that is the effector compound in modulating target gene transcription. Alternative pathways have been described that are Drosha-DGCR8 independent as well as Dicer-independent, and are likely to greatly advance our understanding of miRNA biogenesis and involvement in conditions of health and disease.

Regulatory interactions between miRNA and other noncoding RNAs, including long noncoding RNAs (lncRNA), and circular RNA (circRNA) are now known to determine the cellular functional

status and phenotype. MicroRNAs use seed sequences (6–8 bases long) to bind microRNA Response Elements (MREs) located on their interacting partners, primarily at the 3'UTR of coding transcripts [4]. However, it is possible that the frequency of MREs in the entire transcriptome of a given cell contributes to the dynamic gene regulatory process by acting as a sponge for mature miRNAs, thus regulating their functional availability. Thus, the coding and noncoding transcripts sharing common MRE sites compete with each other and define the gene expression profile of a given cell. These competing transcripts are collectively called competing endogenous RNAs [5]. Thus, gene expression regulation is a complex process involving the dynamic interactions between miRNA-mRNA-lncRNA-circRNA. This complexity is increased manifold when these interacting partners are exchanged between cells via extracellular vesicles. Recognizing this intricate system will significantly aid in our understanding of the health and disease process. We are beginning to study disease pathogenesis at the cellular, organ, and whole body levels; and the gut microbiota is increasingly recognized as a crucial player [6]. It is time to acknowledge each of these players as they take center stage in maintaining homeostasis and the normal physiological functioning of an organism.

There are still significant gaps in understanding the complex regulatory mechanisms of miRNAs; however, the field is advancing rapidly. Further, it has been shown that a number of nuclear receptors are involved in the transcriptional regulation of miRNA expression, including the small heterodimer partner (SHP) and farnesoid X receptor (FXR). In general, miRNAs are detected as (i) extracellular circulating miRNA bound to different lipoproteins; (ii) part of a non-membrane ribonucleoprotein complex associated with Argonaute proteins; and (iii) contained in exosomes as extracellular vesicles, where they act as nano-sized transporters involved in the communication between neighboring cells. The recent finding that miRNAs are also transported from one cell to another via tunneling nanotubes underscores their importance in maintaining communication among all cell types, including those associated with cancer.

This Special Issue of *Genes*, entitled “MicroRNA Regulation in Health and Disease” consists of a series of articles spanning the clinical realm from colorectal cancer to pulmonary fibrosis. However, we begin with a research article by Liu et al. who reported for the first time the existence of complemented palindromic small RNAs (cpsRNAs) from SARS coronavirus, and propose that cpsRNAs and palindromic small RNAs (psRNAs) constitute a novel class of small RNAs [7]. Such a discovery of cpsRNAs could pave a way to find novel markers for pathogen detection and to reveal the mechanisms underlying infection or pathogenesis from a different perspective. In a study titled, “A Two-Cohort RNA-seq Study Reveals Changes in Endometrial and Blood miRNome in Fertile and Infertile Women”, Rekker et al. compared mid-secretory phase samples between fertile and infertile women [8]. The study revealed 21 differentially expressed miRNAs from the endometrium and one from blood samples. Among the novel miRNAs, chr2_4401 was validated and showed upregulation in the mid-secretory endometrium. In addition to the novel findings, the authors confirmed the involvement of miR-30 and miR-200 family members in mid-secretory endometrial functions. Hueso et al. elegantly showed in their article that an exonic switch regulates the differential accession of microRNAs to the *Cd34* transcript in atherosclerosis progression [9]. Further, they proposed a new mechanism of miRNA action, linked to a cryptic splicing site in the target-host gene, that would regulate the differential accession of miRNAs to their cognate binding sites.

Li et al. studied the role of miRNA-106a-5p in inhibiting C2C12 Myogenesis via targeting PIK3R1 and modulating PI3K/AKT signaling [10]. Their results showed that miR-106a-5p was elevated in aged muscles and dexamethasone (DEX)-treated myotubes. The up-regulation of miR-106a-5p significantly reduced the diameters of myotubes accompanied by increased levels of muscular atrophy genes and decreased PI3K/AKT activities. Finally, miR-106a-5p was demonstrated to directly bind to the 3'-UTR of PIK3R1, thus, repressing PI3K/AKT signaling. The microbiome appears to interact and perhaps influence an unlimited number of metabolic processes in health and disease. In their article, Yuan et al. postulate that the altered nutrient composition and miRNA expression in colorectal cancer (CRC) microenvironment selectively exerts pressure on the surrounding microbiota, leading

to alterations in its composition [11]. Further, the authors present a detailed overview of the current understanding of the role of miRNAs in mediating host-microbiota interactions in CRC. “Single Nucleotide Polymorphisms in *MIR143* Contribute to Protection Against Non-Hodgkin Lymphoma (NHL) in Caucasian Populations” by Bradshaw et al. is first to report a correlation between miRSNPs in *miR-143* and a reduced risk of NHL in Caucasians [12]. Further, it is supported by significant SNPs in high linkage disequilibrium (LD) in a large European NHL genome-wide association study (GWAS) meta-analysis. Axmann et al. compared the miRNA profiles in serum and lipoprotein particles of healthy individuals with those of patients with uremia [13]. They observed a significant increase in levels of cellular miRNA level using reconstituted high-density lipoprotein (HDL) particles artificially loaded with miRNA, whereas incubation with native HDL particles yielded no measurable effect. Based on the results, the authors concluded that there was no relevant effect of lipoprotein-particle-mediated miRNA-transfer under in vivo conditions though the miRNA profile of lipoprotein particles can be used as a diagnostic marker.

Mullenbrock et al., carried out an elegant systems analysis transcriptomic and proteomic study on the potential role of miRNAs in pulmonary fibrosis [14]. They specifically targeted fibroblasts and myofibroblasts as the key effector cells responsible for the excessive extracellular matrix (ECM) deposition and fibrosis progression in both idiopathic pulmonary fibrosis (IPF) and systemic sclerosis (SSc) patient lungs. The comprehensive analyses of mRNA, miRNA, and matrisome proteomic profiles in IPF and SSc lung fibroblasts revealed robust fibrotic signatures at both the gene and protein expression levels and identified novel fibrogenesis-associated miRNAs whose aberrant downregulation in disease fibroblasts likely contributes to their fibrotic and ECM gene expression. Somatostatin (SST) analogues were used to control the proliferation and symptoms of neuroendocrine tumors (NETs) in an article by Døssing et al., entitled “Somatostatin Analogue Treatment Primarily Induce miRNA Expression Changes and Up-Regulates Growth Inhibitory miR-7 and miR-148a in Neuroendocrine Cells” [15]. Two miRNAs which were highly induced by SST analogues, miR-7 and miR-148a, were shown to inhibit the proliferation of NCI-H727 and CNDT2 cells. SST analogues also produced a general up-regulation of the let-7 family members. SST analogues controlled and induced distinct miRNA expression patterns among which miR-7 and miR-148a both have growth inhibitory properties.

As FDA-approved small RNA drugs begin to enter the arena of clinical medicine, it is critical to expand both preclinical and clinical research studies for miRNAs. A growing number of reports suggest a significant utility of miRNAs as biomarkers for pathogenic conditions, modulators of drug resistance, and/or as drugs for medical intervention in almost all human health conditions. The pleiotropic nature of this class of nonprotein-coding RNAs makes them particularly attractive drug targets for diseases with a multifactorial origin and few, if any, available treatments. The landscape of both diagnostic and interventional medicine will arguably continue to evolve as candidate miRNAs pass successfully through phase 2 and 3 clinical trials. In this special issue of *Genes*, we provide a series of articles that highlight microRNAs as diagnostic, predictive and therapeutic agents for human disease. The development of bioinformatics programs to identify miRNA-binding sites in target genes and their corresponding biological pathways, along with an expanding platform of in vitro and in vivo preclinical research models, has propelled miRNAs into clinical medicine. The first siRNA human trial was conducted in 2004 and, in 2018, the first siRNA drug was approved, paving the way for a class of miRNA transcripts whose active investigation began only a little more than 15 years ago. The future of human miRNA clinical trials is absolutely guaranteed and that time has arrived.

The development of miRNA diagnostics and therapeutics is an exciting and potentially new frontier in treating diseases for which few treatment options exist. We believe and hope that this Special Issue of *Genes* will be an important resource for a wide variety of audiences, including students at all levels, and established investigators who are interested in contributing to the remarkable and ever-expanding field of microRNAs in health and disease.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bartel, D.P. Metazoan MicroRNAs. *Cell* **2018**, *173*, 20–51. [[CrossRef](#)] [[PubMed](#)]
2. Hanna, J.; Hossain, G.S.; Kocerha, J. The Potential for microRNA Therapeutics and Clinical Research. *Front. Genet.* **2019**, *10*, 478. [[CrossRef](#)] [[PubMed](#)]
3. Kleaveland, B.; Shi, C.Y.; Stefano, J.; Bartel, D.P. A Network of Noncoding Regulatory RNAs Acts in the Mammalian Brain. *Cell* **2018**, *174*, 350–362. [[CrossRef](#)] [[PubMed](#)]
4. Tay, Y.; Rinn, J.; Pandolfi, P.P. The Multilayered Complexity of ceRNA Crosstalk and Competition. *Nature* **2014**, *505*, 344–352. [[CrossRef](#)] [[PubMed](#)]
5. Karreth, F.A.; Pandolfi, P.P. ceRNA Cross-talk in Cancer: When ce-bling Rivalries Go Awry. *Cancer Discov.* **2013**, *3*, 1113–1121. [[CrossRef](#)] [[PubMed](#)]
6. Yuan, C.; Subramanian, S. microRNA-Mediated Tumor-Microbiota Metabolic Interactions in Colorectal Cancer. *DNA Cell. Biol.* **2019**, *38*, 281–285. [[CrossRef](#)] [[PubMed](#)]
7. Liu, C.; Chen, Z.; Hu, Y.; Ji, H.; Yu, D.; Shen, W.; Li, S.; Ruan, J.; Bu, W.; Gao, S. Complemented Palindromic Small RNAs First Discovered from SARS Coronavirus. *Genes (Basel)* **2018**, *9*, 442. [[CrossRef](#)] [[PubMed](#)]
8. Rekker, K.; Altmae, S.; Suhorutshenko, M.; Peters, M.; Martinez-Blanch, J.F.; Codoner, F.M.; Vilella, F.; Simon, C.; Salumets, A.; Velthut-Meikas, A. A Two-Cohort RNA-seq Study Reveals Changes in Endometrial and Blood miRNome in Fertile and Infertile Women. *Genes (Basel)* **2018**, *9*, 574. [[CrossRef](#)] [[PubMed](#)]
9. Hueso, M.; Cruzado, J.M.; Torras, J.; Navarro, E. An Exonic Switch Regulates Differential Accession of microRNAs to the Cd34 Transcript in Atherosclerosis Progression. *Genes (Basel)* **2019**, *10*, 70. [[CrossRef](#)] [[PubMed](#)]
10. Li, X.; Zhu, Y.; Zhang, H.; Ma, G.; Wu, G.; Xiang, A.; Shi, X.; Yang, G.S.; Sun, S. MicroRNA-106a-5p Inhibited C2C12 Myogenesis via Targeting PIK3R1 and Modulating the PI3K/AKT Signaling. *Genes (Basel)* **2018**, *9*, 333. [[CrossRef](#)] [[PubMed](#)]
11. Yuan, C.; Steer, C.J.; Subramanian, S. Host—MicroRNA—Microbiota Interactions in Colorectal Cancer. *Genes (Basel)* **2019**, *10*, 270. [[CrossRef](#)] [[PubMed](#)]
12. Bradshaw, G.; Haupt, L.M.; Aquino, E.M.; Lea, R.A.; Sutherland, H.G.; Griffiths, L.R. Single Nucleotide Polymorphisms in MIR143 Contribute to Protection Against Non-Hodgkin Lymphoma (NHL) in Caucasian Populations. *Genes (Basel)* **2019**, *10*, 185. [[CrossRef](#)] [[PubMed](#)]
13. Axmann, M.; Meier, S.M.; Karner, A.; Strobl, W.; Stangl, H.; Plochberger, B. Serum and Lipoprotein Particle miRNA Profile in Uremia Patients. *Genes (Basel)* **2018**, *9*, 533. [[CrossRef](#)] [[PubMed](#)]
14. Mullenbrock, S.; Liu, F.; Szak, S.; Hronowski, X.; Gao, B.; Juhasz, P.; Sun, C.; Liu, M.; McLaughlin, H.; Xiao, Q.; et al. Systems Analysis of Transcriptomic and Proteomic Profiles Identifies Novel Regulation of Fibrotic Programs by miRNAs in Pulmonary Fibrosis Fibroblasts. *Genes (Basel)* **2018**, *9*, 588. [[CrossRef](#)] [[PubMed](#)]
15. Døssing, K.B.V.; Kjaer, C.; Vikesa, J.; Binderup, T.; Knigge, U.; Culler, M.D.; Kjaer, A.; Federspiel, B.; Friis-Hansen, L. Somatostatin Analogue Treatment Primarily Induce miRNA Expression Changes and Up-Regulates Growth Inhibitory miR-7 and miR-148a in Neuroendocrine Cells. *Genes (Basel)* **2018**, *9*, 337. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Review

Host–MicroRNA–Microbiota Interactions in Colorectal Cancer

Ce Yuan ^{1,2}, Clifford J. Steer ^{3,4} and Subbaya Subramanian ^{1,2,*}

¹ Bioinformatics and Computational Biology Program, University of Minnesota, Minneapolis, MN 55455, USA; yuanx236@umn.edu

² Department of Surgery, University of Minnesota Medical School, Minneapolis, MN 55455, USA

³ Department of Medicine, University of Minnesota Medical School, Minneapolis, MN 55455, USA; steer001@umn.edu

⁴ Department of Genetics, Cell Biology, and Development, University of Minnesota Medical School, Minneapolis, MN 55455, USA

* Correspondence: subree@umn.edu; Tel.: +612-626-4330

Received: 29 January 2019; Accepted: 27 March 2019; Published: 2 April 2019

Abstract: Changes in gut microbiota composition have consistently been observed in patients with colorectal cancer (CRC). Yet, it is not entirely clear how the gut microbiota interacts with tumor cells. We know that tumor cells undergo a drastic change in energy metabolism, mediated by microRNAs (miRNAs), and that tumor-derived miRNAs affect the stromal and immune cell fractions of the tumor microenvironment. Recent studies suggest that host intestinal miRNAs can also affect the growth and composition of the gut microbiota. Our previous CRC studies showed a high-level of interconnectedness between host miRNAs and their microbiota. Considering all the evidence to date, we postulate that the altered nutrient composition and miRNA expression in the CRC microenvironment selectively exerts pressure on the surrounding microbiota, leading to alterations in its composition. In this review article, we present our current understanding of the role of miRNAs in mediating host–microbiota interactions in CRC.

Keywords: colorectal cancer; microRNAs; gut microbiota; metabolic interactions

1. Introduction

An average human intestine contains more than 100 trillion bacteria (collectively known as the gut microbiota) [1]. In recent decades, a number of studies have suggested that the gut microbiota is crucial to human health and to the development of diseases, including colorectal cancer (CRC) [2–7]. Those studies determined that altered microbiota composition and function (dysbiosis) is a common signature of CRC. Bacterial candidates such as *Fusobacterium nucleatum* and *Bacteroides fragilis* are consistently enriched in tumor tissues, and included in that signature. Specific factors in those bacteria, including FadA and Fap2 protein from *F. nucleatum* and *B. fragilis* toxins, that play a role in CRC pathobiology have been identified [8–16]. However, our knowledge of the vast majority of other bacteria associated with the CRC microenvironment is limited. Moreover, we are just beginning to understand the complex interactions between host and microbiota in CRC, as well as other clinical disorders including neurodegenerative diseases [17].

In healthy humans, a key factor associated with microbiota variations is host genetics [18–21]. In a study of healthy twins, Goodrich et al. found that host genetics drive microbiota composition and can also affect the host metabolic phenotype [19]. Several other studies have found an association between the abundance of *Bifidobacterium* species and the presence of single-nucleotide polymorphisms (SNPs) in close proximity to the host lactase gene locus [18,22]. This association suggests that the *Bifidobacterium* species conceivably assists the host in metabolizing lactose.

A recent CRC study found that loss-of-function mutations in the mitogen-activated protein kinase (MAPK) and Wnt signaling pathways are associated with specific sets of microbiota profiles [2]. Furthermore, mutations in the tumor-suppressor adenomatous polyposis coli (APC) gene are also associated with a distinct inter-microbiota association network [2]. These findings suggest that a common genetic factor might orchestrate the dynamic host–microbiota interaction(s) and functional relationship(s). Indeed, other recent studies have provided experimental evidence that microRNAs (miRNAs) can influence the survival and composition of gut bacteria [3,23,24]. Moreover, miRNAs have important intermediate roles in regulating CRC transformation and progression via the action of signaling pathways, including MAPK, Wnt, and APC [25,26]. MicroRNAs are small noncoding RNAs (about 22 nt) that play an important role in regulating and fine-tuning gene expression [27]. In mammalian cells, miRNAs regulate gene expression through posttranscriptional modifications in two distinct, albeit paired, mechanisms. First, if the miRNA has an extensive complementary binding site in the messenger RNA (mRNA) target, then it will guide the RNA-induced silencing complex (RISC) to cleave the mRNA, thus inhibiting translation. Second, if the miRNA only partially binds to the 3′ untranslated region (3′UTR) of the mRNA, then the miRNA-RISC will act to repress mRNA translation [28]. Both mechanisms lead to the decreased translation of mRNAs, which alters their respective downstream functions. Because miRNAs can act upon mRNA targets with limited complementarity, each miRNA can target a wide range of mRNAs in mammalian cells and each mRNA can be targeted by numerous miRNAs. More than 30% of human genes are estimated to have conserved binding sites in the 3′UTR [29]. Clearly, given this vast and enormously complex regulatory network, miRNAs are immensely important in regulating critical cellular processes. We are only now beginning to understand the sophisticated cross-talk of miRNAs, not only with each other, but with the myriad of target mRNAs.

Based on mounting evidence, we postulate that the altered nutrient composition and miRNA expression in the CRC microenvironment selectively influences the surrounding microbiota, leading to alterations in its composition. In this review, we present our current understanding of the role of miRNAs in mediating host–microbiota interactions in CRC (Figure 1). After highlighting the evidence pointing to their central role, we reflect on the future direction of this rapidly evolving field.

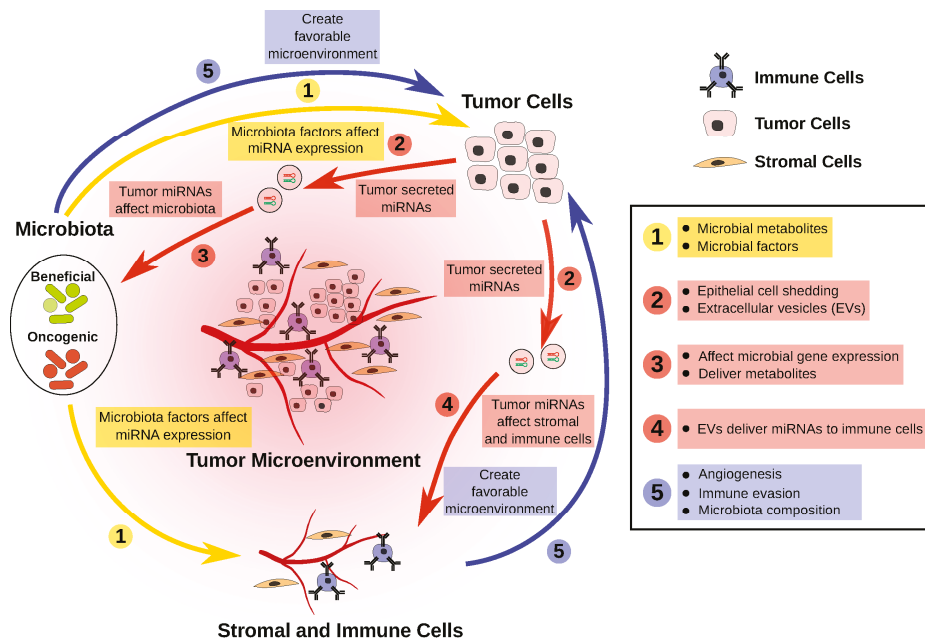


Figure 1. Host–microRNA–microbiota interactions in colorectal cancer. Microbiota composition has a functional effect on the cancer cells, via stromal and tumor infiltrating immune cells by regulating various cellular process (1). Microbial-metabolites and other secreted factors affect miRNA/gene expression profiles in cells present in the tumor microenvironment. In turn, tumor cells affect the microbiota composition of the stromal and tumor infiltrating immune cells through shedding of epithelial cells and/or secreting extracellular vesicles (EVs) containing miRNAs (2). The tumor-miRNAs alter the microbiota composition by affecting the gene expression of the microbiota and by delivering cancer-secreted metabolites (3). The tumor-derived miRNAs also have a role in regulating stromal and tumor infiltrating immune cells by affecting gene expression through miRNAs delivered in EVs (4). Such interactions will finally create a favorable microenvironment for tumor cells that include angiogenesis, immune evasion, and microbiota composition (5).

2. Microbiota and Colorectal Cancer

In the healthy intestine, the microbiota maintains a stable structure and actively participates in energy harvesting and nutrient production from undigested food [30,31]. However, this balance is disrupted in patients with CRC. Current evidence suggests that the microbiota regulates host functions via both metabolites and secreted factors.

2.1. Microbial Metabolites

In a normal colon, the microbiota produces a vast number of metabolites. Some of them, including vitamin K, biotin, and short-chain fatty acids (SCFAs), are essential for maintaining homeostasis in the colon microenvironment [31]. In fact, the major energy source (~70%) required by colon epithelium is butyrate, which is produced by the microbiota through fermentation of complex carbohydrates. Without the microbiota, the colon epithelium undergoes autophagy and fails to maintain its normal structure and function [32]. Similarly, mice lacking a microbiota (i.e., germ-free mice or those treated by broad-spectrum antibiotics) develop significantly fewer tumors in the colon [33–35]. However, in humans, using broad-spectrum antibiotics to treat CRC is not feasible, because of the risk of introducing harmful and highly resistant secondary infections such as *Clostridium difficile*.

In our current understanding, a few main classes of bacterial metabolites play a key role in the pathogenesis of CRC and the immune microenvironment. These metabolites include SCFAs, polyamines, secondary bile acids, and phytochemicals. Their role in CRC has been extensively reviewed and documented [31,36–38]. We have recently also explored the role of miRNAs in mediating the effect of microbial metabolites on CRC and its microenvironment [39].

2.2. Microbial Factors

Early studies consistently found a greater population of *F. nucleatum* in microbiota samples in patients with CRC than in healthy controls [40,41]. This bacterium is commonly found in the human oral microbiota and is frequently associated with gum diseases; it is, however, not commonly present in the gut microbiota. Through the Fap2 virulence factor, it uniquely binds with the D-galactose- β (1-3)-N-acetyl-D-galactosamine (Gal-GalNAc) carbohydrate moiety expressed on the tumor surface of CRCs [14]. Once it localizes to the CRC microenvironment, it targets the Wnt/ β -catenin signaling pathway by binding, via association with the FadA virulence factor, to the E-cadherin protein on the cell surface [16]. The Wnt/ β -catenin signaling pathway is critical during tumor initiation, tumor migration, and metabolic reprogramming [42–45]. The role of the Wnt/ β -catenin signaling pathway in CRC has been previously reviewed [45].

Another bacterial protein targeting the same Wnt/ β -catenin signaling pathway is the *Bacteroides fragilis* toxin (bft) produced by *B. fragilis* [46]. The bft virulence factor is able to bind to the E-cadherin protein, similar to that of FadA, but additionally cleaves the protein, which can alter the intestinal tight-junction function [47]. The Wnt/ β -catenin pathway is a major signaling pathway that controls the expression of many important tumor-related genes, including MYC. The transcription factor MYC, transactivates miRNAs, such as the miR-17-92 cluster, that are highly expressed in CRC [48–51].

Additionally, *F. nucleatum* can also induce CRC cell proliferation by upregulating miR-21, via activation of the nuclear factor kappa-light-chain-enhancer of activated B cells (NF- κ B) pathway via toll-like receptor 4 (TLR4) signaling [52]. The *Escherichia coli* bacterium harboring the pks genomic island also plays an important role in CRC. When CRC cells come in contact with the colibactin genotoxin produced by *E. coli*, the cells undergo cellular senescence [53,54]. This process is mediated by the cellular upregulation of miR-20a-5p, which results in the downregulation of sentrin-specific protease 1 (SENP1). This process then alters p53 small ubiquitin-like modifier (SUMO)ylation, which has been shown to affect the growth and metastasis of tumor cells [55].

In addition to factors that are virulent, many bacteria also produce beneficial factors that can reduce inflammation and modulate the immune system. In germ-free mice, early studies found impaired intestinal immune systems, which were amenable to treatment [56]. Specifically, the *B. fragilis* polysaccharide A (PSA) is one such immunomodulatory factor that maintains the proper function of CD4+ T cells [57]. Several other polysaccharides produced by *B. fragilis* are also beneficial in maintaining proper immune function. Immunization with *B. fragilis* polysaccharides, or the adoptive transfer of T cells specific to *B. fragilis*, can even boost the treatment effect of anti-cytotoxic T-lymphocyte antigen 4 (CTLA-4) immunotherapy [58]. The seemingly conflicting role of *B. fragilis* within gut bacteria is only the tip of the iceberg in current microbiota research and the fine and highly complex balance between functions.

In light of all this evidence, we created the first system-level map of interactions between host miRNAs and the microbiota [3]. Our comprehensive map helped us analyze correlations between host miRNA expression levels and mucosa-associated microbiota profiles, specifically in patients with CRC.

3. MicroRNAs and Colorectal Cancer

Previous studies have identified numerous aberrant miRNA expression patterns in CRC [25,59–62]. Specifically, the miR-17-92 cluster, miR-21, miR-182, and miR-503 are consistently overexpressed in tumor (vs. normal) tissues [3,26,48,49,59,63–71]. Any alteration(s) in expression levels of these miRNAs could, in turn, affect a wide array of downstream gene targets. Together, these miRNAs regulate all

aspects of tumor pathobiology, including (i) altering tumor metabolism; (ii) promoting cell proliferation; (iii) stimulating angiogenesis; (iv) down-regulating tumor-suppressor genes; (v) promoting evasion of immune surveillance; and (vi) creating a favorable tumor microenvironment that promotes invasion and metastasis.

Our laboratory previously reported that, during the adenoma to adenocarcinoma transition, miR-182 and miR-503 were sequentially overexpressed and targeted the tumor-suppressor *FBXW7* gene [69]. Other researchers have observed, during CRC transformation, an increased expression of the miR-17-92 cluster and miR-21 [48,72]. In CRC adenocarcinoma, members of the miR-17-92 cluster target transforming growth factor-beta (TGF- β), which in turn stimulates angiogenesis in the tumor microenvironment, thus promoting tumor growth [70]. Additionally, miR-19, a member of the miR-17-92 cluster, downregulates expression of the tumor-suppressor phosphatase and tensin homolog (PTEN), thereby activating the protein kinase B (AKT)/mammalian target of rapamycin (mTOR) pathway in tumor cells [73]. The AKT/mTOR pathway is the main metabolic sensing pathway, responsible for regulating glucose transport into cells [74]. Since glucose is the main fuel source of CRC cells, an activated AKT/mTOR pathway promotes tumor cell proliferation [75].

The tumor-suppressor *PDCD4* gene, which is commonly downregulated in CRC, is a target of miR-21 [67]. Inhibiting the *PDCD4* gene can lead to an increase in the metastasis potential of tumor cells. Another important pathway commonly altered in CRC tumors is the Wnt/ β -catenin pathway [25,26]. Dozens of miRNAs have been shown to extensively regulate the genes involved in the Wnt/ β -catenin pathway [25].

The complex microenvironment of the CRC tumor also involves stromal cell and immune cell fractions, which can be regulated by cancer-derived miRNAs [76–78]. Studies have found that the miR-17-92 cluster, commonly overexpressed in CRC cells, is also upregulated in CRC stromal cells [68,72,79]. Strikingly, these miRNAs are not only endogenously produced by stromal cells, but also packaged in the microvesicles of tumor cells, and then delivered to stromal cells [80,81]. Similar intracellular regulation mediated by miRNAs is also found in immune cell fractions [82]. Additionally, endogenous miRNA dysregulation is prevalent in CRC immune cell fractions, usually as a downstream effect of tumor-secreted factors such as cytokines and chemokines [83–85]. Collectively, this evidence suggests that miRNAs are important in regulating tumor cells, in addition to maintaining the tumor microenvironment. It is clear that the relationship between miRNAs and CRC is multifaceted, interrelated, and highly complex.

4. Host Regulation of Microbiota Mediated by MicroRNAs

In reestablishing germ-free mice with a normal microbiota, studies have found altered intestinal miRNA profiles, suggesting that the microbiota regulates host miRNA expression [86,87]. Moreover, the responses of intestinal cells to facilitating the microbiota process depends on the cell type, and intestinal epithelial stem cells are especially sensitive to microbiota reestablishment [87].

Because miRNAs are highly stable, several studies in the clinical arena were able to detect higher levels of miR-21 and miR-92a, among other miRNAs, in the fecal samples of patients with CRC [65,88,89]. This finding facilitated in developing a noninvasive CRC screening method and delineating the potential role of miRNAs in interacting with the trillions of microbes in the human gut.

Intestinal miRNAs develop from two main sources, including the host and the food [23,24]. The intestinal epithelial cells are the main contributors of host-derived miRNAs, either via shedding of cells or excretion of exosomes. Evidence has shown that miRNAs from food can be absorbed by the host and can affect host gene expression [90–92]. But certain food-sourced miRNAs remain stable in the digestive tract and reach the intestines [93,94]. This evidence suggests that miRNAs can mediate cross-species regulation. The idea remains nascent, so insight into how miRNAs mediate host–microbiota interactions is still limited. Liu et al. first demonstrated such regulation, showing that miRNAs present in the feces can regulate gene expression and growth of bacteria [23]. Specifically, they found that mice lacking the *Dicer* gene, which enables mature miRNA processing, had different

microbiota profiles than wild-type mice. More importantly, the study reported that hsa-miR-515-5p promoted the growth of *F. nucleatum* in vitro by targeting the 16S ribosomal RNA (rRNA) gene.

Notably, however, hsa-miR-515-5p shows very low expression levels in CRC tumors, so they are not significantly different from normal tissue. Thus, interactions between hsa-miR-515-5p and *F. nucleatum* might not be significant in CRC pathogenesis. However, more importantly, this study found that fecal miRNA transplantation restores fecal microbiota composition in mice with Dicer gene knockout. Several recent studies found that fecal microbiota transplantation (FMT) offers a potential therapeutic benefit that enables an immunotherapeutic response [35,95–98]. Based on growing evidence, it is plausible that fecal miRNAs play an important role in modulating the CRC microbiota as well as immunotherapy responses.

Recently, Teng et al. demonstrated that miRNAs encapsulated in plant-derived exosome-like nanoparticles (ELNs) can enter bacteria and alter bacterial genes [24]. The process for bacterial uptake of ELNs is determined primarily by the lipid composition of the outer membrane. They found that ELNs enriched with phosphatidylcholine were preferentially taken up by the *Ruminococcus* species, whereas ELNs enriched with phosphatidic acid (PA) were primarily taken up by *Lactobacillus rhamnosus*. After the ELNs are taken up by specific bacteria, the miRNA contents are released into bacterial cells. Teng et al. also found that mdo-miR7267-3p encapsulated in the PA-enriched ELNs targets the *Lactobacillus* monoxygenase *ycnE*, which then increases its production of indole-3-carboxaldehyde (I3A). The I3A metabolite then promotes interleukin-22 (IL-22) production and helps repair damaged colon mucosa [99].

There is developing evidence to support the notion that host or exogenous miRNAs might be biologically active in bacteria, thereby affecting bacterial gene expression. Although small RNAs similar to miRNAs exist in bacteria and function similarly to miRNAs, it remains unknown as to how miRNAs function in bacteria [100]. Several studies have reported that exogenous miRNAs from plant or animal sources can be taken up by human cells and exert biological functions [90–94,101–103]. Additional studies are required to ascertain whether miRNAs can indeed affect bacteria and to delineate the precise mechanism(s).

5. Metabolic Changes in Colorectal Cancer and Microbiota Mediated by MicroRNAs

The prevailing “driver-passenger” model suggests that dysbiosis in the CRC microbiota is initially caused by colonization of driver bacteria. This is followed by a gradual change in the tumor microenvironment, an increase in the number of driver bacteria, and secondary colonization of passenger bacteria that benefit from the changed environment [104]. That model, together with other studies, suggest that a gradual metabolic change in the tumor microenvironment during cancer progression could be the cause of dysbiosis [31]. Again, we explored that issue in our recent review of the role of miRNAs in mediating the effect of microbial metabolites on CRC and its microenvironment [39].

One of the hallmarks of tumor growth is their increased use of glycolysis as a main energy source, known universally as the Warburg effect [105]. Because the normal colon uses butyrate as its major energy source, any change in that source preferred by proliferating tumor cells will undoubtedly profoundly alter the nutrient composition of the tumor microenvironment [106,107]. Indeed, several studies have found altered metabolite levels in CRC tissues and stools [107–110]. A significantly lower glucose level and higher levels of lactate and fatty acids have been found in CRC tumor tissues, as compared with adjacent normal tissues. In stool samples from patients with CRC, a higher level of amino acids and a lower level of fatty acids have also been observed [108]. Interestingly, the CRC microbiota has shown reduced carbohydrate metabolism and an increase in the biosynthesis of amino acids and fatty acids [41]. In CRC, the switch in the nutrient source preferred by proliferating tumor cells appears to alter the nutrient composition in the tumor microenvironment. At the same time, the nutrient metabolism of the tumor microbiota seems to complement the nutrient needs of the tumor. This could be due to factors associated with the tumor nutrient microenvironment, and by the miRNAs excreted by tumor cells, on the surrounding microbiota [3]. Given the role of miRNAs

in mediating such metabolic changes, we believe that miRNAs play a central, if not critical role, in mediating host–microbiota metabolic interactions in CRC.

6. Conclusions and Perspectives

With thousands of bacterial species living in the human digestive tract, it is becoming quite evident that they profoundly affect human health. Our review of the recent literature regarding CRC underscores a complex metabolic interplay between the host and its microbiota, mediated in part by miRNAs. Based on the current literature, we offer five major points in host–microbiota interactions mediated by miRNAs (Figure 1):

1. The CRC microbiota has reduced representation of beneficial bacteria. These bacteria produce metabolites and other factors that can potentially slow CRC progression, in part via the modulation of miRNAs that regulate tumor cells.
2. Dysregulation of miRNAs in tumor cells can affect the survival, or the gene expression, of certain bacteria in the microbiota.
3. Dysregulated miRNAs in tumor cells can be packaged and delivered to both stromal and immune cell fractions, creating a more favorable microenvironment for tumor cells.
4. Overrepresentation of oncogenic bacteria in the CRC microbiota can modulate tumor cells, as well as the tumor microenvironment, through miRNA modulation, thereby resulting in a more favorable condition for tumor growth.
5. This negative feedback loop perpetuates CRC progression.

Potential methods to break such a negative feedback loop include:

1. Interfering with host-mediated microbiota modulation by designing strategies to deliver anti-miRNAs to block the effect of host-miRNAs on the microbiota.
2. Modulating the microbiota through miRNAs that promote the growth of beneficial bacteria while suppressing the growth of oncogenic bacteria, in conjunction with chemotherapy or immunotherapy.

Based on both experimental and computational data, we conclude that miRNAs mediate and critically influence host–microbiota interactions. Clearly, miRNAs are a major part of a complex web of highly dynamic interactions. Other factors, such as nutrient availability in the CRC microenvironment, could also play an important role. In the future, it will be imperative to use a combination of approaches to comprehensively survey the CRC microenvironment, in order to discover all potential players in mediating such interactions.

Acknowledgments: S.S. is supported by research grants funded by the NIH R03CA219129 and Chainbreaker Grant; C.Y., by the MnDrive—University of Minnesota Informatics Institute graduate fellowship. Due to space restrictions, we cannot cite many other significant contributions made by numerous researchers and laboratories in this potentially important and rapidly progressing field. We thank Mary Knatterud for assisting in manuscript editing.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Turnbaugh, P.J.; Ley, R.E.; Hamady, M.; Fraser-Liggett, C.M.; Knight, R.; Gordon, J.I. The human microbiome project. *Nature* **2007**, *449*, 804–810. [[CrossRef](#)]
2. Burns, M.B.; Montassier, E.; Abrahante, J.; Priya, S.; Niccum, D.E.; Khoruts, A.; Starr, T.K.; Knights, D.; Blekhman, R. Colorectal cancer mutational profiles correlate with defined microbial communities in the tumor microenvironment. *PLoS Genet.* **2018**, *14*, e1007376. [[CrossRef](#)]
3. Yuan, C.; Burns, M.B.; Subramanian, S.; Blekhman, R. Interaction between host microRNAs and the gut microbiota in colorectal cancer. *mSystems* **2018**, *3*, e00205-17. [[CrossRef](#)] [[PubMed](#)]

4. Dejea, C.M.; Wick, E.C.; Hechenbleikner, E.M.; White, J.R.; Mark Welch, J.L.; Rossetti, B.J.; Peterson, S.N.; Snesrud, E.C.; Borisy, G.G.; Lazarev, M.; et al. Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 18321–18326. [[CrossRef](#)]
5. Kostic, A.D.; Gevers, D.; Pedamallu, C.S.; Michaud, M.; Duke, F.; Earl, A.M.; Ojesina, A.I.; Jung, J.; Bass, A.J.; Taberero, J.; et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res.* **2012**, *22*, 292–298. [[CrossRef](#)]
6. Shah, M.S.; DeSantis, T.; Yamal, J.-M.; Weir, T.; Ryan, E.P.; Cope, J.L.; Hollister, E.B. Re-purposing 16S rRNA gene sequence data from within case paired tumor biopsy and tumor-adjacent biopsy or fecal samples to identify microbial markers for colorectal cancer. *PLoS ONE* **2018**, *13*, e0207002. [[CrossRef](#)]
7. García-Castillo, V.; Sanhueza, E.; McNeerney, E.; Onate, S.A.; García, A. Microbiota dysbiosis: A new piece in the understanding of the carcinogenesis puzzle. *J. Med. Microbiol.* **2016**, *65*, 1347–1362. [[CrossRef](#)] [[PubMed](#)]
8. Boleij, A.; Hechenbleikner, E.M.; Goodwin, A.C.; Badani, R.; Stein, E.M.; Lazarev, M.G.; Ellis, B.; Carroll, K.C.; Albesiano, E.; Wick, E.C.; et al. The *Bacteroides fragilis* toxin gene is prevalent in the colon mucosa of colorectal cancer patients. *Clin. Infect. Dis.* **2015**, *60*, 208–215. [[CrossRef](#)] [[PubMed](#)]
9. Toprak, N.U.; Yagci, A.; Gulluoglu, B.M.; Akin, M.L.; Demirkalem, P.; Celenk, T.; Soyletir, G. A possible role of *Bacteroides fragilis* enterotoxin in the aetiology of colorectal cancer. *Clin. Microbiol. Infect.* **2006**, *12*, 782–786. [[CrossRef](#)]
10. Wu, S.; Shin, J.; Zhang, G.; Cohen, M.; Franco, A.; Sears, C.L. The *Bacteroides fragilis* toxin binds to a specific intestinal epithelial cell receptor. *Infect. Immun.* **2006**, *74*, 5382–5390. [[CrossRef](#)]
11. Kostic, A.D.; Chun, E.; Robertson, L.; Glickman, J.N.; Gallini, C.A.; Michaud, M.; Clancy, T.E.; Chung, D.C.; Lochhead, P.; Hold, G.L.; et al. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe* **2013**, *14*, 207–215. [[CrossRef](#)]
12. Mima, K.; Cao, Y.; Chan, A.T.; Qian, Z.R.; Nowak, J.A.; Masugi, Y.; Shi, Y.; Song, M.; da Silva, A.; Gu, M.; et al. *Fusobacterium nucleatum* in colorectal carcinoma tissue according to tumor location. *Clin. Transl. Gastroenterol.* **2016**, *7*, e200. [[CrossRef](#)]
13. Mima, K.; Nishihara, R.; Qian, Z.R.; Cao, Y.; Sukawa, Y.; Nowak, J.A.; Yang, J.; Dou, R.; Masugi, Y.; Song, M.; et al. *Fusobacterium nucleatum* in colorectal carcinoma tissue and patient prognosis. *Gut* **2016**, *65*, 1973–1980. [[CrossRef](#)]
14. Abed, J.; Emgård, J.E.M.; Zamir, G.; Faroja, M.; Almog, G.; Grenov, A.; Sol, A.; Naor, R.; Pikarsky, E.; Atlan, K.A.; et al. Fap2 mediates *Fusobacterium nucleatum* colorectal adenocarcinoma enrichment by binding to tumor-expressed Gal-GalNAc. *Cell Host Microbe* **2016**, *20*, 215–225. [[CrossRef](#)]
15. Rubinstein, M.R.; Baik, J.E.; Lagana, S.M.; Han, R.P.; Raab, W.J.; Sahoo, D.; Dalerba, P.; Wang, T.C.; Han, Y.W. *Fusobacterium nucleatum* promotes colorectal cancer by inducing Wnt/ β -catenin modulator Annexin A1. *EMBO Rep.* **2019**. [[CrossRef](#)]
16. Rubinstein, M.R.; Wang, X.; Liu, W.; Hao, Y.; Cai, G.; Han, Y.W. *Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/ β -catenin signaling via its FadA adhesin. *Cell Host Microbe* **2013**, *14*, 195–206. [[CrossRef](#)]
17. Schroeder, B.O.; Bäckhed, F. Signals from the gut microbiota to distant organs in physiology and disease. *Nat. Med.* **2016**, *22*, 1079–1089. [[CrossRef](#)]
18. Blekhman, R.; Goodrich, J.K.; Huang, K.; Sun, Q.; Bukowski, R.; Bell, J.T.; Spector, T.D.; Keinan, A.; Ley, R.E.; Gevers, D.; et al. Host genetic variation impacts microbiome composition across human body sites. *Genome Biol.* **2015**, *16*, 191. [[CrossRef](#)] [[PubMed](#)]
19. Goodrich, J.K.; Waters, J.L.; Poole, A.C.; Sutter, J.L.; Koren, O.; Blekhman, R.; Beaumont, M.; Van Treuren, W.; Knight, R.; Bell, J.T.; et al. Human genetics shape the gut microbiome. *Cell* **2014**, *159*, 789–799. [[CrossRef](#)]
20. Bongers, G.; Pacer, M.E.; Geraldino, T.H.; Chen, L.; He, Z.; Hashimoto, D.; Furtado, G.C.; Ochando, J.; Kelley, K.A.; Clemente, J.C.; et al. Interplay of host microbiota, genetic perturbations, and inflammation promotes local development of intestinal neoplasms in mice. *J. Exp. Med.* **2014**, *211*, 457–472. [[CrossRef](#)]
21. Davenport, E.R.; Cusanovich, D.A.; Michelini, K.; Barreiro, L.B.; Ober, C.; Gilad, Y. Genome-wide association studies of the human gut microbiota. *PLoS ONE* **2015**, *10*, e0140301. [[CrossRef](#)]
22. Goodrich, J.K.; Davenport, E.R.; Beaumont, M.; Jackson, M.A.; Knight, R.; Ober, C.; Spector, T.D.; Bell, J.T.; Clark, A.G.; Ley, R.E. Genetic determinants of the gut microbiome in UK twins. *Cell Host Microbe* **2016**, *19*, 731–743. [[CrossRef](#)] [[PubMed](#)]

23. Liu, S.; da Cunha, A.P.; Rezende, R.M.; Cialic, R.; Wei, Z.; Bry, L.; Comstock, L.E.; Gandhi, R.; Weiner, H.L. The host shapes the gut microbiota via fecal microRNA. *Cell Host Microbe* **2016**, *19*, 32–43. [[CrossRef](#)] [[PubMed](#)]
24. Teng, Y.; Ren, Y.; Sayed, M.; Hu, X.; Lei, C.; Kumar, A.; Hutchins, E.; Mu, J.; Deng, Z.; Luo, C.; et al. Plant-derived exosomal microRNAs shape the gut microbiota. *Cell Host Microbe* **2018**, *24*, 637–652.e8. [[CrossRef](#)] [[PubMed](#)]
25. Slattery, M.L.; Mullany, L.E.; Sakoda, L.C.; Samowitz, W.S.; Wolff, R.K.; Stevens, J.R.; Herrick, J.S. Expression of Wnt-signaling pathway genes and their associations with miRNAs in colorectal cancer. *Oncotarget* **2018**, *9*, 6075–6085. [[CrossRef](#)] [[PubMed](#)]
26. Li, Y.; Lauriola, M.; Kim, D.; Francesconi, M.; D’Uva, G.; Shibata, D.; Malafa, M.P.; Yeatman, T.J.; Coppola, D.; Solmi, R.; et al. Adenomatous polyposis coli (APC) regulates miR17-92 cluster through β -catenin pathway in colorectal cancer. *Oncogene* **2016**, *35*, 4558–4568. [[CrossRef](#)] [[PubMed](#)]
27. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297. [[CrossRef](#)]
28. Fabian, M.R.; Sonenberg, N. The mechanics of miRNA-mediated gene silencing: A look under the hood of miRISC. *Nat. Struct. Mol. Biol.* **2012**, *19*, 586–593. [[CrossRef](#)]
29. 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. [[CrossRef](#)]
30. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **2012**, *486*, 207–214. [[CrossRef](#)] [[PubMed](#)]
31. Louis, P.; Hold, G.L.; Flint, H.J. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat. Rev. Microbiol.* **2014**, *12*, 661–672. [[CrossRef](#)]
32. Donohoe, D.R.; Garge, N.; Zhang, X.; Sun, W.; O’Connell, T.M.; Bunker, M.K.; Bultman, S.J. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab.* **2011**, *13*, 517–526. [[CrossRef](#)]
33. Dove, W.F.; Clipson, L.; Gould, K.A.; Luongo, C.; Marshall, D.J.; Moser, A.R.; Newton, M.A.; Jacoby, R.F. Intestinal neoplasia in the ApcMin mouse: Independence from the microbial and natural killer (beige locus) status. *Cancer Res.* **1997**, *57*, 812–814. [[PubMed](#)]
34. Zackular, J.P.; Baxter, N.T.; Chen, G.Y.; Schloss, P.D. Manipulation of the gut microbiota reveals role in colon tumorigenesis. *mSphere* **2016**, *1*. [[CrossRef](#)] [[PubMed](#)]
35. Vivarelli, S.; Salemi, R.; Candido, S.; Falzone, L.; Santagati, M.; Stefani, S.; Torino, F.; Banna, G.L.; Tonini, G.; Libra, M. Gut microbiota and cancer: From pathogenesis to therapy. *Cancers* **2019**, *11*, 38. [[CrossRef](#)]
36. Blacher, E.; Levy, M.; Tatrovsky, E.; Elinav, E. Microbiome-modulated metabolites at the interface of host immunity. *J. Immunol.* **2017**, *198*, 572–580. [[CrossRef](#)]
37. Johnson, C.H.; Spilker, M.E.; Goetz, L.; Peterson, S.N.; Siuzdak, G. Metabolite and microbiome interplay in cancer immunotherapy. *Cancer Res.* **2016**, *76*, 6146–6152. [[CrossRef](#)]
38. O’Keefe, S.J.D. Diet, microorganisms and their metabolites, and colon cancer. *Nat. Rev. Gastroenterol. Hepatol.* **2016**, *13*, 691–706. [[CrossRef](#)]
39. Yuan, C.; Subramanian, S. MicroRNA mediated tumor-microbiota metabolic interactions in colorectal cancer. *DNA Cell Biol.* **2019**. [[CrossRef](#)] [[PubMed](#)]
40. Castellarin, M.; Warren, R.L.; Freeman, J.D.; Dreolini, L.; Krzywinski, M.; Strauss, J.; Barnes, R.; Watson, P.; Allen-Vercoe, E.; Moore, R.A.; et al. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res.* **2012**, *22*, 299–306. [[CrossRef](#)]
41. Burns, M.B.; Lynch, J.; Starr, T.K.; Knights, D.; Blekhan, R. Virulence genes are a signature of the microbiome in the colorectal tumor microenvironment. *Genome Med.* **2015**, *7*, 55. [[CrossRef](#)] [[PubMed](#)]
42. Pate, K.T.; Stringari, C.; Sprowl-Tanio, S.; Wang, K.; TeSlaa, T.; Hoverter, N.P.; McQuade, M.M.; Garner, C.; Digman, M.A.; Teitell, M.A.; et al. Wnt signaling directs a metabolic program of glycolysis and angiogenesis in colon cancer. *EMBO J.* **2014**, *33*, 1454–1473. [[CrossRef](#)]
43. Konsavage, W.M.; Kyler, S.L.; Rennoll, S.A.; Jin, G.; Yochum, G.S. Wnt/ β -catenin signaling regulates Yes-associated protein (YAP) gene expression in colorectal carcinoma cells. *J. Biol. Chem.* **2012**, *287*, 11730–11739. [[CrossRef](#)]
44. Strillacci, A.; Valerii, M.C.; Sansone, P.; Caggiano, C.; Sgromo, A.; Vittori, L.; Fiorentino, M.; Poggioli, G.; Rizzello, F.; Campieri, M.; et al. Loss of miR-101 expression promotes Wnt/ β -catenin signalling pathway activation and malignancy in colon cancer cells. *J. Pathol.* **2013**, *229*, 379–389. [[CrossRef](#)]
45. Bienz, M.; Clevers, H. Linking colorectal cancer to Wnt signaling. *Cell* **2000**, *103*, 311–320. [[CrossRef](#)]

46. Wu, S.; Morin, P.J.; Maouyo, D.; Sears, C.L. Bacteroides fragilis enterotoxin induces c-Myc expression and cellular proliferation. *Gastroenterology* **2003**, *124*, 392–400. [[CrossRef](#)] [[PubMed](#)]
47. Wu, S.; Lim, K.C.; Huang, J.; Saidi, R.F.; Sears, C.L. Bacteroides fragilis enterotoxin cleaves the zonula adherens protein, E-cadherin. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14979–14984. [[CrossRef](#)]
48. Diosdado, B.; van de Wiel, M.A.; Terhaar Sive Droste, J.S.; Mongera, S.; Postma, C.; Meijerink, W.J.H.J.; Carvalho, B.; Meijer, G.A. MiR-17-92 cluster is associated with 13q gain and c-Myc expression during colorectal adenoma to adenocarcinoma progression. *Br. J. Cancer* **2009**, *101*, 707–714. [[CrossRef](#)]
49. Mogilyansky, E.; Rigoutsos, I. The miR-17/92 cluster: A comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. *Cell Death Differ.* **2013**, *20*, 1603–1614. [[CrossRef](#)]
50. O'Donnell, K.A.; Wentzel, E.A.; Zeller, K.I.; Dang, C.V.; Mendell, J.T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **2005**, *435*, 839–843. [[CrossRef](#)]
51. Dang, C.V. c-Myc target genes involved in cell growth, apoptosis, and metabolism. *Mol. Cell Biol.* **1999**, *19*, 1–11. [[CrossRef](#)] [[PubMed](#)]
52. Yang, Y.; Weng, W.; Peng, J.; Hong, L.; Yang, L.; Toyama, Y.; Gao, R.; Liu, M.; Yin, M.; Pan, C.; et al. *Fusobacterium nucleatum* increases proliferation of colorectal cancer cells and tumor development in mice by activating Toll-Like receptor 4 signaling to nuclear factor- κ B, and up-regulating expression of microRNA-21. *Gastroenterology* **2017**, *152*, 851–866.e24. [[CrossRef](#)]
53. Dalmaso, G.; Cougnoux, A.; Delmas, J.; Darfeuille-Michaud, A.; Bonnet, R. The bacterial genotoxin colibactin promotes colon tumor growth by modifying the tumor microenvironment. *Gut Microbes* **2014**, *5*, 675–680. [[CrossRef](#)] [[PubMed](#)]
54. Cougnoux, A.; Dalmaso, G.; Martinez, R.; Buc, E.; Delmas, J.; Gibold, L.; Sauvanet, P.; Darcha, C.; Déchelotte, P.; Bonnet, M.; et al. Bacterial genotoxin colibactin promotes colon tumour growth by inducing a senescence-associated secretory phenotype. *Gut* **2014**, *63*, 1932–1942. [[CrossRef](#)] [[PubMed](#)]
55. Baek, S.H. A novel link between SUMO modification and cancer metastasis. *Cell Cycle* **2006**, *5*, 1492–1495. [[CrossRef](#)]
56. Gordon, H.A. Morphological and physiological characterization of germfree life. *Ann. N.Y. Acad. Sci.* **1959**, *78*, 208–220. [[CrossRef](#)] [[PubMed](#)]
57. Mazmanian, S.K.; Liu, C.H.; Tzianabos, A.O.; Kasper, D.L. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* **2005**, *122*, 107–118. [[CrossRef](#)]
58. Vétizou, M.; Pitt, J.M.; Daillère, R.; Lepage, P.; Waldschmitt, N.; Flament, C.; Rusakiewicz, S.; Routy, B.; Roberti, M.P.; Duong, C.P.M.; et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* **2015**, *350*, 1079–1084. [[CrossRef](#)] [[PubMed](#)]
59. Sarver, A.L.; French, A.J.; Borralho, P.M.; Thayanyithy, V.; Oberg, A.L.; Silverstein, K.A.T.; Morlan, B.W.; Riska, S.M.; Boardman, L.A.; Cunningham, J.M.; et al. Human colon cancer profiles show differential microRNA expression depending on mismatch repair status and are characteristic of undifferentiated proliferative states. *BMC Cancer* **2009**, *9*, 401. [[CrossRef](#)]
60. Sarver, A.L.; Sarver, A.E.; Yuan, C.; Subramanian, S. OMCD: Oncomir cancer database. *BMC Cancer* **2018**, *18*, 1223. [[CrossRef](#)] [[PubMed](#)]
61. Wong, N.W.; Chen, Y.; Chen, S.; Wang, X. OncomiR: An online resource for exploring pan-cancer microRNA dysregulation. *Bioinformatics* **2018**, *34*, 713–715. [[CrossRef](#)]
62. Falzone, L.; Scola, L.; Zanghi, A.; Biondi, A.; Di Cataldo, A.; Libra, M.; Candido, S. Integrated analysis of colorectal cancer microRNA datasets: Identification of microRNAs associated with tumor development. *Aging* **2018**, *10*, 1000–1014. [[CrossRef](#)] [[PubMed](#)]
63. Koga, Y.; Yasunaga, M.; Takahashi, A.; Kuroda, J.; Moriya, Y.; Akasu, T.; Fujita, S.; Yamamoto, S.; Baba, H.; Matsumura, Y. MicroRNA expression profiling of exfoliated colonocytes isolated from feces for colorectal cancer screening. *Cancer Prev. Res. (Phila Pa)* **2010**, *3*, 1435–1442. [[CrossRef](#)] [[PubMed](#)]
64. Fang, L.; Li, H.; Wang, L.; Hu, J.; Jin, T.; Wang, J.; Yang, B.B. MicroRNA-17-5p promotes chemotherapeutic drug resistance and tumour metastasis of colorectal cancer by repressing PTEN expression. *Oncotarget* **2014**, *5*, 2974–2987. [[CrossRef](#)] [[PubMed](#)]
65. Wu, C.W.; Ng, S.S.M.; Dong, Y.J.; Ng, S.C.; Leung, W.W.; Lee, C.W.; Wong, Y.N.; Chan, F.K.L.; Yu, J.; Sung, J.J.Y. Detection of miR-92a and miR-21 in stool samples as potential screening biomarkers for colorectal cancer and polyps. *Gut* **2012**, *61*, 739–745. [[CrossRef](#)] [[PubMed](#)]

66. Kulda, V.; Pesta, M.; Topolcan, O.; Liska, V.; Treska, V.; Sutnar, A.; Rupert, K.; Ludvikova, M.; Babuska, V.; Holubec, L.; et al. Relevance of miR-21 and miR-143 expression in tissue samples of colorectal carcinoma and its liver metastases. *Cancer Genet. Cytogenet.* **2010**, *200*, 154–160. [[CrossRef](#)]
67. Asangani, I.A.; Rasheed, S.A.K.; Nikolova, D.A.; Leupold, J.H.; Colburn, N.H.; Post, S.; Allgayer, H. MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdc4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene* **2008**, *27*, 2128–2136. [[CrossRef](#)] [[PubMed](#)]
68. Bullock, M.D.; Pickard, K.M.; Nielsen, B.S.; Sayan, A.E.; Jenei, V.; Mellone, M.; Mitter, R.; Primrose, J.N.; Thomas, G.J.; Packham, G.K.; et al. Pleiotropic actions of miR-21 highlight the critical role of deregulated stromal microRNAs during colorectal cancer progression. *Cell Death Dis.* **2013**, *4*, e684. [[CrossRef](#)]
69. Li, L.; Sarver, A.L.; Khatri, R.; Hajeri, P.B.; Kamenev, I.; French, A.J.; Thibodeau, S.N.; Steer, C.J.; Subramanian, S. Sequential expression of miR-182 and miR-503 cooperatively targets FBXW7, contributing to the malignant transformation of colon adenoma to adenocarcinoma. *J. Pathol.* **2014**, *234*, 488–501. [[CrossRef](#)]
70. Dews, M.; Fox, J.L.; Hultine, S.; Sundaram, P.; Wang, W.; Liu, Y.Y.; Furth, E.; Enders, G.H.; El-Deiry, W.; Schelter, J.M.; et al. The myc-miR-17~92 axis blunts TGF β signaling and production of multiple TGF β -dependent antiangiogenic factors. *Cancer Res.* **2010**, *70*, 8233–8246. [[CrossRef](#)]
71. Hu, S.; Liu, L.; Chang, E.B.; Wang, J.-Y.; Raufman, J.-P. Butyrate inhibits pro-proliferative miR-92a by diminishing c-Myc-induced miR-17-92a cluster transcription in human colon cancer cells. *Mol. Cancer* **2015**, *14*, 180. [[CrossRef](#)] [[PubMed](#)]
72. Yamamichi, N.; Shimomura, R.; Inada, K.; Sakurai, K.; Haraguchi, T.; Ozaki, Y.; Fujita, S.; Mizutani, T.; Furukawa, C.; Fujishiro, M.; et al. Locked nucleic acid in situ hybridization analysis of miR-21 expression during colorectal cancer development. *Clin. Cancer Res.* **2009**, *15*, 4009–4016. [[CrossRef](#)] [[PubMed](#)]
73. Grillari, J.; Hackl, M.; Grillari-Voglauer, R. miR-17-92 cluster: Ups and downs in cancer and aging. *Biogerontology* **2010**, *11*, 501–506. [[CrossRef](#)] [[PubMed](#)]
74. Makinoshima, H.; Takita, M.; Saruwatari, K.; Umemura, S.; Obata, Y.; Ishii, G.; Matsumoto, S.; Sugiyama, E.; Ochiai, A.; Abe, R.; et al. Signaling through the phosphatidylinositol 3-kinase (PI3K)/mammalian target of rapamycin (mTOR) axis is responsible for aerobic glycolysis mediated by glucose transporter in epidermal growth factor receptor (EGFR)-mutated lung adenocarcinoma. *J. Biol. Chem.* **2015**, *290*, 17495–17504. [[CrossRef](#)] [[PubMed](#)]
75. Weijenberg, M.P.; Hughes, L.A.E.; Bours, M.J.L.; Simons, C.C.J.M.; van Engeland, M.; van den Brandt, P.A. The mTOR pathway and the role of energy balance throughout life in colorectal cancer etiology and prognosis: Unravelling mechanisms through a multidimensional molecular epidemiologic approach. *Curr. Nutr. Rep.* **2013**, *2*, 19–26. [[CrossRef](#)] [[PubMed](#)]
76. Kohlhaas, F.J.; Mitra, A.K.; Lengyel, E.; Peter, M.E. MicroRNAs as mediators and communicators between cancer cells and the tumor microenvironment. *Oncogene* **2015**, *34*, 5857–5868. [[CrossRef](#)]
77. Kosaka, N.; Yoshioka, Y.; Hagiwara, K.; Tominaga, N.; Katsuda, T.; Ochiya, T. Trash or Treasure: Extracellular microRNAs and cell-to-cell communication. *Front. Genet.* **2013**, *4*, 173. [[CrossRef](#)]
78. Runtsch, M.C.; Round, J.L.; O'Connell, R.M. MicroRNAs and the regulation of intestinal homeostasis. *Front. Genet.* **2014**, *5*, 347. [[CrossRef](#)]
79. Nielsen, B.S.; Jørgensen, S.; Fog, J.U.; Søkilde, R.; Christensen, I.J.; Hansen, U.; Brünner, N.; Baker, A.; Møller, S.; Nielsen, H.J. High levels of microRNA-21 in the stroma of colorectal cancers predict short disease-free survival in stage II colon cancer patients. *Clin. Exp. Metastasis* **2011**, *28*, 27–38. [[CrossRef](#)]
80. Dews, M.; Homayouni, A.; Yu, D.; Murphy, D.; Seignani, C.; Wentzel, E.; Furth, E.E.; Lee, W.M.; Enders, G.H.; Mendell, J.T.; et al. Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster. *Nat. Genet.* **2006**, *38*, 1060–1065. [[CrossRef](#)]
81. Zhuang, G.; Wu, X.; Jiang, Z.; Kasman, I.; Yao, J.; Guan, Y.; Oeh, J.; Modrusan, Z.; Bais, C.; Sampath, D.; et al. Tumour-secreted miR-9 promotes endothelial cell migration and angiogenesis by activating the JAK-STAT pathway. *EMBO J.* **2012**, *31*, 3513–3523. [[CrossRef](#)]
82. Fanini, F.; Fabbri, M. Cancer-derived exosomal microRNAs shape the immune system within the tumor microenvironment: State of the art. *Semin. Cell Dev. Biol.* **2017**, *67*, 23–28. [[CrossRef](#)] [[PubMed](#)]
83. Sonda, N.; Simonato, F.; Peranzoni, E.; Cali, B.; Bortoluzzi, S.; Bisognin, A.; Wang, E.; Marincola, F.M.; Naldini, L.; Gentner, B.; et al. miR-142-3p prevents macrophage differentiation during cancer-induced myelopoiesis. *Immunity* **2013**, *38*, 1236–1249. [[CrossRef](#)] [[PubMed](#)]

84. Zhang, M.; Liu, Q.; Mi, S.; Liang, X.; Zhang, Z.; Su, X.; Liu, J.; Chen, Y.; Wang, M.; Zhang, Y.; et al. Both miR-17-5p and miR-20a alleviate suppressive potential of myeloid-derived suppressor cells by modulating STAT3 expression. *J. Immunol.* **2011**, *186*, 4716–4724. [[CrossRef](#)]
85. Mei, S.; Xin, J.; Liu, Y.; Zhang, Y.; Liang, X.; Su, X.; Yan, H.; Huang, Y.; Yang, R. MicroRNA-200c promotes suppressive potential of myeloid-derived suppressor cells by modulating PTEN and FOG2 expression. *PLoS ONE* **2015**, *10*, e0135867. [[CrossRef](#)]
86. Dalmaso, G.; Nguyen, H.T.T.; Yan, Y.; Laroui, H.; Charania, M.A.; Ayyadurai, S.; Sitaraman, S.V.; Merlin, D. Microbiota modulate host gene expression via microRNAs. *PLoS ONE* **2011**, *6*, e19293. [[CrossRef](#)] [[PubMed](#)]
87. Peck, B.C.E.; Mah, A.T.; Pitman, W.A.; Ding, S.; Lund, P.K.; Sethupathy, P. Functional transcriptomics in diverse intestinal epithelial cell types reveals robust microRNA sensitivity in intestinal stem cells to microbial status. *J. Biol. Chem.* **2017**, *292*, 2586–2600. [[CrossRef](#)]
88. Phua, L.C.; Chue, X.P.; Koh, P.K.; Cheah, P.Y.; Chan, E.C.Y.; Ho, H.K. Global fecal microRNA profiling in the identification of biomarkers for colorectal cancer screening among Asians. *Oncol. Rep.* **2014**, *32*, 97–104. [[CrossRef](#)]
89. Rotelli, M.T.; Di Lena, M.; Cavallini, A.; Lippolis, C.; Bonfrate, L.; Chetta, N.; Portincasa, P.; Altomare, D.F. Fecal microRNA profile in patients with colorectal carcinoma before and after curative surgery. *Int. J. Colorectal Dis.* **2015**, *30*, 891–898. [[CrossRef](#)]
90. Baier, S.R.; Nguyen, C.; Xie, F.; Wood, J.R.; Zempeni, J. MicroRNAs are absorbed in biologically meaningful amounts from nutritionally relevant doses of cow milk and affect gene expression in peripheral blood mononuclear cells, HEK-293 kidney cell cultures, and mouse livers. *J. Nutr.* **2014**, *144*, 1495–1500. [[CrossRef](#)] [[PubMed](#)]
91. Izumi, H.; Tsuda, M.; Sato, Y.; Kosaka, N.; Ochiya, T.; Iwamoto, H.; Namba, K.; Takeda, Y. Bovine milk exosomes contain microRNA and mRNA and are taken up by human macrophages. *J. Dairy Sci.* **2015**, *98*, 2920–2933. [[CrossRef](#)]
92. Zhang, L.; Hou, D.; Chen, X.; Li, D.; Zhu, L.; Zhang, Y.; Li, J.; Bian, Z.; Liang, X.; Cai, X.; et al. Exogenous plant MIR168a specifically targets mammalian LDLRAP1: Evidence of cross-kingdom regulation by microRNA. *Cell Res.* **2012**, *22*, 107–126. [[CrossRef](#)]
93. Mu, J.; Zhuang, X.; Wang, Q.; Jiang, H.; Deng, Z.-B.; Wang, B.; Zhang, L.; Kakar, S.; Jun, Y.; Miller, D.; et al. Interspecies communication between plant and mouse gut host cells through edible plant derived exosome-like nanoparticles. *Mol. Nutr. Food Res.* **2014**, *58*, 1561–1573. [[CrossRef](#)]
94. Wolf, T.; Baier, S.R.; Zempeni, J. The intestinal transport of bovine milk exosomes is mediated by endocytosis in human colon carcinoma Caco-2 cells and rat small intestinal IEC-6 cells. *J. Nutr.* **2015**, *145*, 2201–2206. [[CrossRef](#)]
95. Gopalakrishnan, V.; Spencer, C.N.; Nezi, L.; Reuben, A.; Andrews, M.C.; Karpnits, T.V.; Prieto, P.A.; Vicente, D.; Hoffman, K.; Wei, S.C.; et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* **2018**, *359*, 97–103. [[CrossRef](#)]
96. Matson, V.; Fessler, J.; Bao, R.; Chongsuwat, T.; Zha, Y.; Alegre, M.-L.; Luke, J.J.; Gajewski, T.F. The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. *Science* **2018**, *359*, 104–108. [[CrossRef](#)]
97. Sivan, A.; Corrales, L.; Hubert, N.; Williams, J.B.; Aquino-Michaels, K.; Earley, Z.M.; Benyamin, F.W.; Lei, Y.M.; Jabri, B.; Alegre, M.-L.; et al. Commensal *Bifidobacterium* promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science* **2015**, *350*, 1084–1089. [[CrossRef](#)]
98. Gopalakrishnan, V.; Helmink, B.A.; Spencer, C.N.; Reuben, A.; Wargo, J.A. The influence of the gut microbiome on cancer, immunity, and cancer immunotherapy. *Cancer Cell* **2018**, *33*, 570–580. [[CrossRef](#)]
99. Zelante, T.; Iannitti, R.G.; Cunha, C.; De Luca, A.; Giovannini, G.; Pieraccini, G.; Zecchi, R.; D’Angelo, C.; Massi-Benedetti, C.; Fallarino, F.; et al. Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. *Immunity* **2013**, *39*, 372–385. [[CrossRef](#)]
100. Gottesman, S.; Storz, G. Bacterial small RNA regulators: Versatile roles and rapidly evolving variations. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*. [[CrossRef](#)]
101. Lässer, C.; Alikhani, V.S.; Ekström, K.; Eldh, M.; Paredes, P.T.; Bossios, A.; Sjöstrand, M.; Gabrielsson, S.; Lötval, J.; Valadi, H. Human saliva, plasma and breast milk exosomes contain RNA: Uptake by macrophages. *J. Transl. Med.* **2011**, *9*, 9. [[CrossRef](#)] [[PubMed](#)]

102. Deng, Z.; Rong, Y.; Teng, Y.; Mu, J.; Zhuang, X.; Tseng, M.; Samykutty, A.; Zhang, L.; Yan, J.; Miller, D.; et al. Broccoli-derived nanoparticle inhibits mouse colitis by activating dendritic cell AMP-activated protein kinase. *Mol. Ther.* **2017**, *25*, 1641–1654. [[CrossRef](#)]
103. Ju, S.; Mu, J.; Dokland, T.; Zhuang, X.; Wang, Q.; Jiang, H.; Xiang, X.; Deng, Z.-B.; Wang, B.; Zhang, L.; et al. Grape exosome-like nanoparticles induce intestinal stem cells and protect mice from DSS-induced colitis. *Mol. Ther.* **2013**, *21*, 1345–1357. [[CrossRef](#)]
104. Tjalsma, H.; Boleij, A.; Marchesi, J.R.; Dutilh, B.E. A bacterial driver-passenger model for colorectal cancer: Beyond the usual suspects. *Nat. Rev. Microbiol.* **2012**, *10*, 575–582. [[CrossRef](#)]
105. Warburg, O. On the origin of cancer cells. *Science* **1956**, *123*, 309–314. [[CrossRef](#)]
106. Bishop, K.S.; Xu, H.; Marlow, G. Epigenetic regulation of gene expression induced by butyrate in colorectal cancer: Involvement of microRNA. *Genet. Epigenet.* **2017**, *9*, 1179237X17729900. [[CrossRef](#)]
107. Hu, S.; Dong, T.S.; Dalal, S.R.; Wu, F.; Bissonnette, M.; Kwon, J.H.; Chang, E.B. The microbe-derived short chain fatty acid butyrate targets miRNA-dependent p21 gene expression in human colon cancer. *PLoS ONE* **2011**, *6*, e16221. [[CrossRef](#)]
108. Weir, T.L.; Manter, D.K.; Sheflin, A.M.; Barnett, B.A.; Heuberger, A.L.; Ryan, E.P. Stool microbiome and metabolome differences between colorectal cancer patients and healthy adults. *PLoS ONE* **2013**, *8*, e70803. [[CrossRef](#)]
109. Hirayama, A.; Kami, K.; Sugimoto, M.; Sugawara, M.; Toki, N.; Onozuka, H.; Kinoshita, T.; Saito, N.; Ochiai, A.; Tomita, M.; et al. Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. *Cancer Res.* **2009**, *69*, 4918–4925. [[CrossRef](#)]
110. Brown, D.G.; Rao, S.; Weir, T.L.; O'Malia, J.; Bazan, M.; Brown, R.J.; Ryan, E.P. Metabolomics and metabolic pathway networks from human colorectal cancers, adjacent mucosa, and stool. *Cancer Metab.* **2016**, *4*, 11. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

MicroRNA-106a-5p Inhibited C2C12 Myogenesis via Targeting PIK3R1 and Modulating the PI3K/AKT Signaling

Xiao Li ^{1,†}, Youbo Zhu ^{1,†}, Huifang Zhang ¹, Guangjun Ma ¹, Guofang Wu ², Aoqi Xiang ¹, Xin'E. Shi ¹, Gong She Yang ¹ and Shiduo Sun ^{1,*}

¹ Laboratory of Animal Fat Deposition and Muscle Development, College of Animal Science and Technology, Northwest A and F University, Yangling 712100, China; nicelixiao@nwsuaf.edu.cn (X.L.); zhuyoubo_2015@nwsuaf.edu.cn (Y.Z.); 13135587517@163.com (H.Z.); Guangjunma0904@126.com (G.M.); aoqixiang0305@126.com (A.X.); xineshi@nwsuaf.edu.cn (X.E.S.); gsyang999@hotmail.com (G.S.Y.)

² Stake Key Laboratory of Plateau Ecology and Agriculture, Qinghai Academy of Animal Science and Veterinary Medicine, Qinghai University, Qinghai 810000, China; letitbe521@163.com

* Correspondence: sssm@tom.com; Tel.: +86-029-8709-1017

† These authors contributed equally to this work.

Received: 17 June 2018; Accepted: 28 June 2018; Published: 2 July 2018

Abstract: The microRNA (miR)-17 family is widely expressed in mammalian tissues and play important roles in various physiological and pathological processes. Here, the functions of miR-106a-5p, a member of miR-17 family, were explored during myogenic differentiation in C2C12 cell line. First, miR-106a-5p was found to be relatively lower expressed in two-month skeletal muscle tissues and gradually decreased upon myogenic stimuli. Forced expression of miR-106a-5p significantly reduced the differentiation index, fusion index as well as the expression of myogenic markers (MyoD, MyoG, MyHC, Myomixer, Myomarker). Meanwhile, the levels of phosphorylated AKT were reduced by overexpression of miR-106a-5p, and administration of insulin-like growth factor 1 (IGF1), a booster of myogenic differentiation, could recover all the inhibitory effects above of miR-106a-5p. Furthermore, miR-106a-5p was elevated in aged muscles and dexamethasone (DEX)-treated myotubes, and up-regulation of miR-106a-5p significantly reduced the diameters of myotubes accompanied with increased levels of muscular atrophy genes and decreased PI3K/AKT activities. Finally, miR-106a-5p was demonstrated to directly bind to the 3'-UTR of PIK3R1, thus, repress the PI3K/AKT signaling.

Keywords: miR-106a-5p; myogenic differentiation; muscle atrophy; PIK3R1; C2C12 cell line

1. Introduction

Skeletal muscle comprises approximately 50% of the body's weight and is an endocrine and paracrine organ that plays a key role in the maintenance of the internal environment and homeostasis [1]. Myogenesis is an ordered and extremely complicated process, including myoblasts proliferation, withdrawal from the cell cycle, fusion into multinucleated myotubes, and hypotrophy/atrophy [2]. During muscle development, Myf5, MyoD and MRF4 are myogenic determination factors, and Myogenin is a downstream effector of MyoD and MRF4, and activate the myogenic differentiation program [3,4]. It has been reported that many pathways are involved in myogenesis, such as Wnt/ β -Catenin Signaling [5], the TGF β signaling pathway [6,7], JAK/STAT signaling pathway [8], and PI3K/AKT signaling pathway [9,10]. Recently, more and more microRNAs (miRs) have been reported to participate in myogenesis, such as miR-127 [11], miR-133 [12], miR-1/206 [13], and miR-432 [14]. However, there are still many miRs that remain to be discovered in myoblasts differentiation and muscle formation.

MiR is a class of evolutionarily conserved, regulatory noncoding RNAs of 18–24 nucleotides that participate in the fine-tuning of many, if not all, fundamental biological processes, including skeletal muscle development and muscle-related diseases [15,16]. MiRs have been exquisitely described to control gene expression by binding to mRNA targets using the seed sequences, and miRs with the same or similar seed sequences and origin are grouped into one family or cluster [15]. The miR-17 family consists of the miR-17-92 cluster, miR-106a-363 cluster and miR-106b-25 cluster [17], and the miR-17 family has been well documented to influence the survival, differentiation and functions of various kinds of mammalian cells [18,19]. However, the effects of miRs associated with the miR-17 family in myogenesis are controversial. MiR-17, miR-20a and miR-92a are reported to strongly repress myoblasts differentiation by targeting Enigma homolog 1 (*ENH1*)/Inhibitor of differentiation 1 (*Id1*) [20]. Over the same period, Luo reported that miR-20a/b promoted quail muscle clone 7 (QM-7) myoblasts differentiation by negatively regulating *E2F1* [21]. Intriguingly, miR-106a-5p, a member of the miR-106a-363 cluster, was reported to down-regulate during myogenic differentiation [22,23], and its role in myogenic differentiation deserves to be analyzed.

PI3K (p85 α), encoded by *PIK3R1* gene, is a key protein involved in the PI3K/AKT signaling pathway [24], which is essential for myogenic differentiation [25,26]. It has been recently shown that PI3K/AKT signaling controls muscle-abundant miRs (myomiR) maturation during C2C12 myoblasts differentiation [27]. Insulin and insulin-like growth factor 1 (IGF1) are known as physiological activators of PI3K/AKT signaling in different cell types, including C2C12 myoblasts [28]. Administration of IGF1 promotes myoblast proliferation, differentiation [25,29], and induces myotube hypertrophy [30] by activating the PI3K/AKT signaling pathway. However, regulation of *PIK3R1* by miR-106a-5p and how miR-106a-5p responds to IGF1 stimuli to regulate myogenesis are still poorly understood.

In this study, we analyzed the expression profiles of miR-106a-5p and determined its role and mechanism on myogenesis. Our study identified miR-106a-5p as a novel negative regulator for myogenesis, and miR-106a-5p could repress differentiation and promote atrophy by blocking the PI3K/AKT signaling pathway through targeting *PIK3R1*.

2. Materials and Methods

2.1. Ethics Statement

All experimental mice were operated on in accordance with approved guidelines of the Animal Care and Use Committee of the Northwest A and F University, Yangling, China (NWAFU-314020038) and the guidelines of the Animal Use Committee of the Chinese Ministry of Agriculture (Beijing, China). Furthermore, C2C12 and HEK293T cell lines were obtained from American Type Culture Collection (ATCC, Manassas, VA, USA).

2.2. Cell Culture

The C2C12 myoblasts (ATCC, USA) were used to determine the function of miR-106a-5p during myogenesis. HEK293T (ATCC, USA) was employed for the luciferase reporter analysis. Cells were cultured in Dulbecco's Modified Eagle Media (DMEM, Gibco, ThermoFisher, Waltham, MA, USA) supplemented with 10% FBS (Gibco) and 100 IU/mL penicillin-streptomycin at 37 °C with humidified 5% CO₂ atmosphere. Upon shifting to a 2% horse serum containing medium, C2C12 myoblasts were induced to fusion and differentiation. The medium was changed every day.

2.3. Animals

The C57BL/6 male mice were purchased from the Fourth Military Medical University Animal Center (Xi'an, China) and raised in a controlled temperature (25 \pm 1 °C) with a 12 h light/12 h dark cycle. Tissues were collected from two-month and six-month old mice. All procedures with mice were

in accordance with approved guidelines of the Animal Care and Use Committee of the Northwest A and F University.

2.4. Transfections and Treatment of Myoblasts and Myotubes

To test the effects of miR-106a-5p on differentiating cells, myoblasts were seeded in 6-well or 12-well plates and transfected with 50 nM FAM-labeled miR-106a-5p agomir or negative control (GenePharm, Shanghai, China) using Lipo Plus (Sagecreation, Beijing, China) in Opti-EME (Gibco) according to the manufacturer's instructions. Furthermore, 75 nM IGF1 recombinant protein (Sino Biological, Beijing, China), a PI3K-AKT signaling activator, was used to recover the effects of miR-106a-5p.

To test the effects of miR-106a-5p on well-differentiated myotubes (5 days post-differentiation), myotubes were incubated with 50 μ M Dexamethasone (DEX) for 36 h, then cells were harvested for further analysis. Furthermore, myotube transfection (50 nM miR-106a-5p agomir or negative control) was performed with Lipo Plus (Sagecreation) according to the manufacturer's instructions, and cells were harvested 36 h after transfection.

2.5. Real-Time Quantitative PCR

The total RNA was extracted with Trizol reagent (TakaRa, Otsu, Japan) as recommended by the manufacturer, and the concentration and quality were analyzed by the NanoDrop 2000 (ThermoFisher). Then complementary DNA (cDNA) was synthesized by reverse transcription kit (TakaRa). Real-Time quantitative PCR (RT-qPCR) was performed using Applied Biosystems qPCR instrument (ThermoFisher) and SYBR green PCR Master Mix (Vazyme, Nanjing, China). The expressions of all coding genes were normalized to β -actin, and U6 small RNA was the internal reference when testing the level of miR-106a-5p. miRs quantification was determined by using Bulge-loopTM miRNA qRT-PCR Primer Set (one RT primer and a pair of qPCR primers in each set) specific for miR-106a-5p, designed and synthesized by RiboBio (Guangzhou, China) and other primers which were synthesized by Invitrogen (Shanghai, China). The sequences were shown in Table 1.

Table 1. The primer sequences used for real-time quantitative PCR.

Name	Forward	Reverse
MyoD	CCACTCCGGGACATAGACTTG	AAAAGCGCAGGTCTGGTGAG
MyoG	GAGACATCCCCCTATTCTACCA	GCTCAGTCCGCTCATAGCC
MyHC	GCGAATCGAGGCTCAGAACAA	GTAGTTCGGCCTTCGGTCTTG
Myomarker	CCTGCTGTCTCTCCCAAG	AGAACCAGTGGGTCCCCTAA
Myomixer	GTTAGAAGTGGTGAGCAGGAG	CCATCGGGAGCAATGGAA
MAFbx	CAGCTTCGTGAGCGACCTC	GGCAGTCGAGAAGTCCAGTC
MuRF1	GTGTGAGGTGCTACTTGCTC	GCTCAGTCTTCTGTCCTTGGA
β -actin	GCCATGTACGTAGCCATCCA	ACGCTCGGTCAGGATCTTCA

2.6. Western Blotting Analysis

Cells were harvested in radioimmunoprecipitation assay (RIPA) lysis buffer (Applygen Technologies Inc., Beijing, China) supplemented with protease and phosphatase inhibitor cocktail (Cwbiotech, Jiangsu, China). Protein concentration was determined by BCA Protein Assay Kit (Cwbiotech) and 25 μ g proteins per sample were loaded and separated using a 5% stacking gel and a 10% separating gel. Separated proteins were transferred to PVDF membrane (CST, Boston, MA, USA), and then the membrane was blocked in 5% BSA buffer for 2 h at room temperature, and incubated with primary antibodies against MyoD (1:500, #NB100-56511SS, Novus Biologicals, Littleton, CO, USA), MyoG (1:500, #NB100-56510SS, Novus Biologicals), MyHC (1:1000, #MAB4470, R and D Systems, Minneapolis, MN, USA), p-PI3K (1:500, #4228S, CST, Danvers, MA, USA) and PI3K (p85 α) (1:500, #4257S, CST), p-AKT (1:1000, #4257S, CST) and AKT (1:1000, #9272S, CST), p-mTOR (1:1000,

#5536S, CST) and mTOR (1:1000, #2983S, CST), MAFbx (1:500, #sc-166806, Santa Cruz, Dallas, TX, USA), and MuRF1 (1:500, #C-11, Santa Cruz) at 4 °C overnight. After washing three times (10 min once) in TBST, membranes were incubated with HRP-conjugated goat anti-mouse IgG (1:3000, #BA1050, BosterBio, Wuhan, China) or goat anti-Rabbit IgG (1:3000, #BA1054, BosterBio) for 1.5 h at 4 °C. Imaging and quantification of the bands were carried out by Gel Doc XR system (Bio-Rad, Hercules, CA, USA) and Image Lab software (Bio-Rad).

2.7. Immunofluorescence Analysis

Differentiated C2C12 myotubes were fixed in 4% paraformaldehyde, permeabilized with 0.5% Triton X-100, and then blocked in 5% BSA for 30 min. Later, myotubes were sequentially incubated with anti-myosin heavy chain monoclonal antibody (1:200, #MAB4470, R and D Systems) overnight and an Alexa Fluor 594-conjugated anti-mouse IgG (1:1000, #SA00006-3, Proteintech, Chicago, IL, USA) for 1 h at 4 °C. Finally, nuclei were stained with DAPI. Images were captured with a fluorescence microscope (Nikon, Tokyo, Japan). The myotubes with 1–3, and > 4 nuclei were counted, respectively. The differentiation index was determined as the percentage of MyHC-positive nuclei among total nuclei, and the myotube fusion index was determined as the distribution of the nucleus number in total myotubes according to a previous report [31].

2.8. Luciferase Reporter Assays

The 3'-UTR of PIK3R1 including miR-106a-5p complementary sequences were synthesized by GENERABIOL (Chuzhou, Anhui, China), and inserted into psiCHECKTM-2 Vector (Promega, Madison, WI, USA). Wild-type or mutated constructs and miR-106a-5p agomir or negative control (NC) were co-transfected. Transfected cells were analyzed 48 h post transfection with Dual Luciferase Reporter Assay System (Promega) based on the instructions.

2.9. Statistical Analyses

Values are presented as the mean ± standard error of the mean (SEM), and statistical significance of differences was determined by Student's *t*-test or one-way analysis of variance (ANOVA) by IBM SPSS Statistics 22.0 (Armonk, NY, USA). A *p*-value of < 0.05 was considered statistically significant.

3. Results

3.1. The Expression Profiles of miR-106a-5p in Mice

MiR-106a-5p was highly conserved among detected species based on the nucleotide sequences (Figure 1A). MiR-106a-5p was widely expressed in various tissues in two-month mice, especially in heart and kidney, but a small amount of expression was detected in skeletal muscle (Figure 1B). Further, a significantly higher expression was found in Extensor Digitorum Longus (EDL) and Soleus muscle (SOL) of six-month old mice than that of two-month (Figure 1C). Furthermore, the expression of miR-106a-5p gradually decreased during myogenic differentiation in C2C12 cell line (Figure 1D), while the expression levels of myogenic markers MyoD, MyoG, and MyHC were significantly up-regulated (Figure 1E–G).

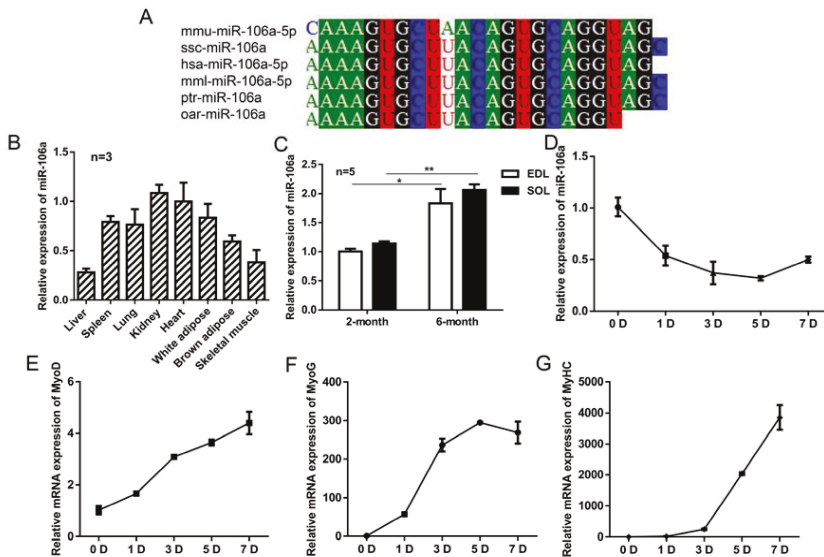


Figure 1. The profiles of miR-106a-5p in mice. (A) The homology comparison of microRNA (miR)-106a-5p between different species. mmu: *mus musculus*; ssc: *sus scrofa*; has: *homo sapiens*; mml: *macaca mulatta*; ptr: *pan troglodytes*; oar: *ovis aries*. Nucleotides with the same shadow were conserved across species. (B) The expression level of miR-106a-5p in different tissues of two-month-old mice. White adipose: subcutaneous white adipose; Skeletal muscle: tibialis anterior muscle. (C) The expression profiles of miR-106a-5p in skeletal muscles of two-month-old and six-month-old mice. EDL: extensor digitorum longus; SOL: soleus. (D–G) Real-time quantitative PCR (RT-qPCR) analysis of miR-106a-5p, MyoD, MyoG, MyHC expression during myoblast differentiation ($n = 3$ per group). D: days. Data were presented as mean \pm standard error of the mean (SEM). * $p < 0.05$, ** $p < 0.01$.

3.2. MiR-106a-5p Suppresses Myoblast Differentiation by Inhibiting the PI3K-AKT Signaling Pathway

Cells were transfected with FAM-labeled miR-106a-5p agomir when reaching 80–90% confluence and then induced to myogenic differentiation at full confluence. Transfection increased miR-106a-5p expression level by around 100-fold (Figure 2A), and almost all cells were FAM positive (Figure 2B). Overexpression of miR-106a-5p significantly decreased the quantity of MyHC positive cells (Figure 2C). In addition, the percentage of multinucleated myotubes, myotube size and myotube fusion index were significantly decreased in cells transfected with miR-106a-5p (Figure 2D–F). Moreover, myogenic regulatory factors (MyoD, MyoG, and MyHC) were down-regulated by miR-106a-5p agomir (Figure 2G–L). Consistently, the expression of fusion genes, Myomixer and Myomarker, were inhibited by enforced miR-106a-5p (Figure 2I,J). Furthermore, the phosphorylations of AKT (ser473) was significantly inhibited by miR-106a-5p agomir, although the phosphorylations of mTOR (ser2448) and PI3K (p85 α) were not significantly changed (Figure 2M,N). Collectively, these results suggested that miR-106a-5p could interfere with C2C12 myoblast differentiation and fusion by blocking PI3K-AKT signaling.

In addition, treatment of 75 nM IGF1 recombinant proteins during myogenic differentiation significantly increased the number of myotubes, differentiation index, and multinucleated myotube fusion index (Figure 3A–D), and dramatically reduced the expression of miR-106a-5p (Figure 3E). Furthermore, IGF1 up-regulated the expression of MyoD, MyoG, and MyHC (Figure 3F–H), triggered the activation of PI3K/AKT signaling pathway, stimulated the phosphorylation of AKT (ser473) (Figure 3G,I). Notably, IGF1 fully restored miR-106a-5p-induced inhibitory effects on

myogenic differentiation, suggested by the increased MyHC positive cells, differentiation index, and multinucleated myotube fusion index (Figure 3J,M) and up-regulated expression of MyoD, MyoG, MyHC (Figure 3N–P). Furthermore, the reduced AKT (ser473) phosphorylation induced by miR-106a-5p was also relieved by recombinant IGF1 protein (Figure 3O,P).

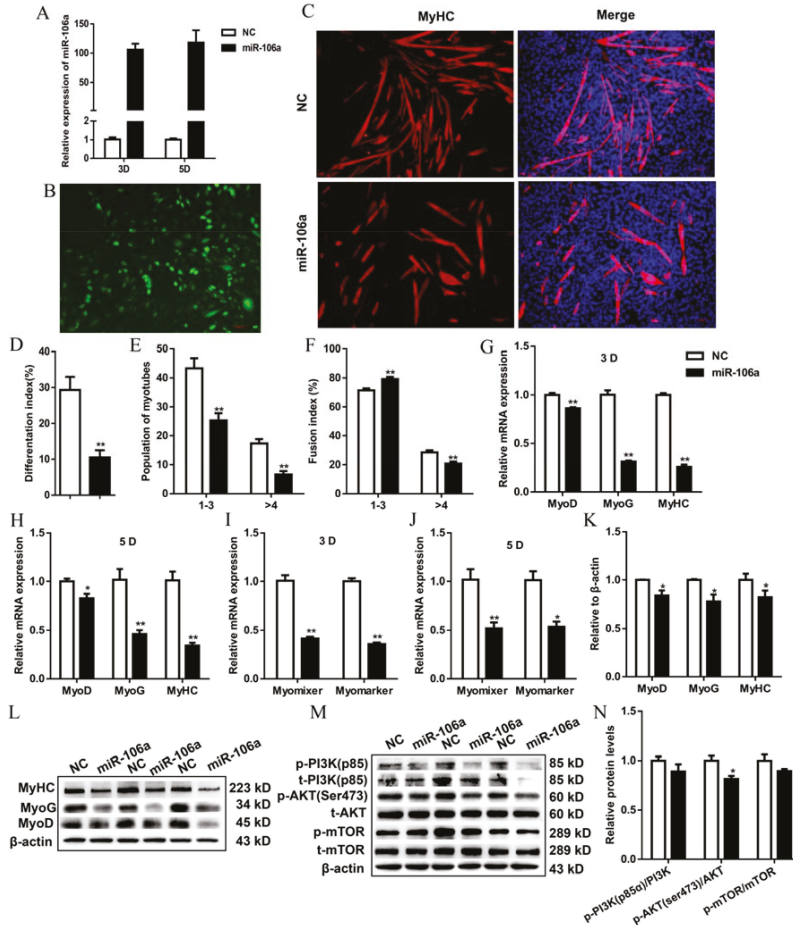


Figure 2. MiR-106a-5p inhibited the myogenic differentiation of C2C12 myoblasts. (A) Overexpression efficiency of miR-106a-5p 3 days (d) and 5 d post differentiation. NC: negative control; (B) The fluorescent microscopy images of C2C12 cells transfected with FAM-labeled miR-106a-5p agomir ($\times 10$). Scale bars = 500 μm ; (C) Immunostaining for MyHC (red) and DAPI (blue) on 5 d post differentiation ($\times 20$). Scale bars = 100 μm ; (D–F) The statistical results of differentiation index, fusion index and the populations of myotubes, respectively; 1–3 indicates myotubes with 1, 2 or 3 nucleus, >4 indicates myotubes with 4 more nucleus; (G,H) The mRNA expression of MyoD, MyoG, MyHC on 3 d and 5 d post differentiation; (I,J) The mRNA expression of Myomarker and Myomixer 3 d and 5 d post differentiation; (K) The statistical results of MyoD, MyoG, MyHC proteins in Figure 2L; (L) Western blot analyzed for MyoD, MyoG, MyHC proteins 5 d post differentiation; (M) Protein levels of key molecules in PI3K–AKT pathway in C2C12 cells transfected with miR-106a-5p agomir or NC on 5 d post differentiation; (N) The statistical analysis of phosphorylated PI3K (p85 α), AKT (ser473) and mTOR (ser2448). Data were presented as mean \pm SEM. $n = 3$ per group. * $p < 0.05$, ** $p < 0.01$.

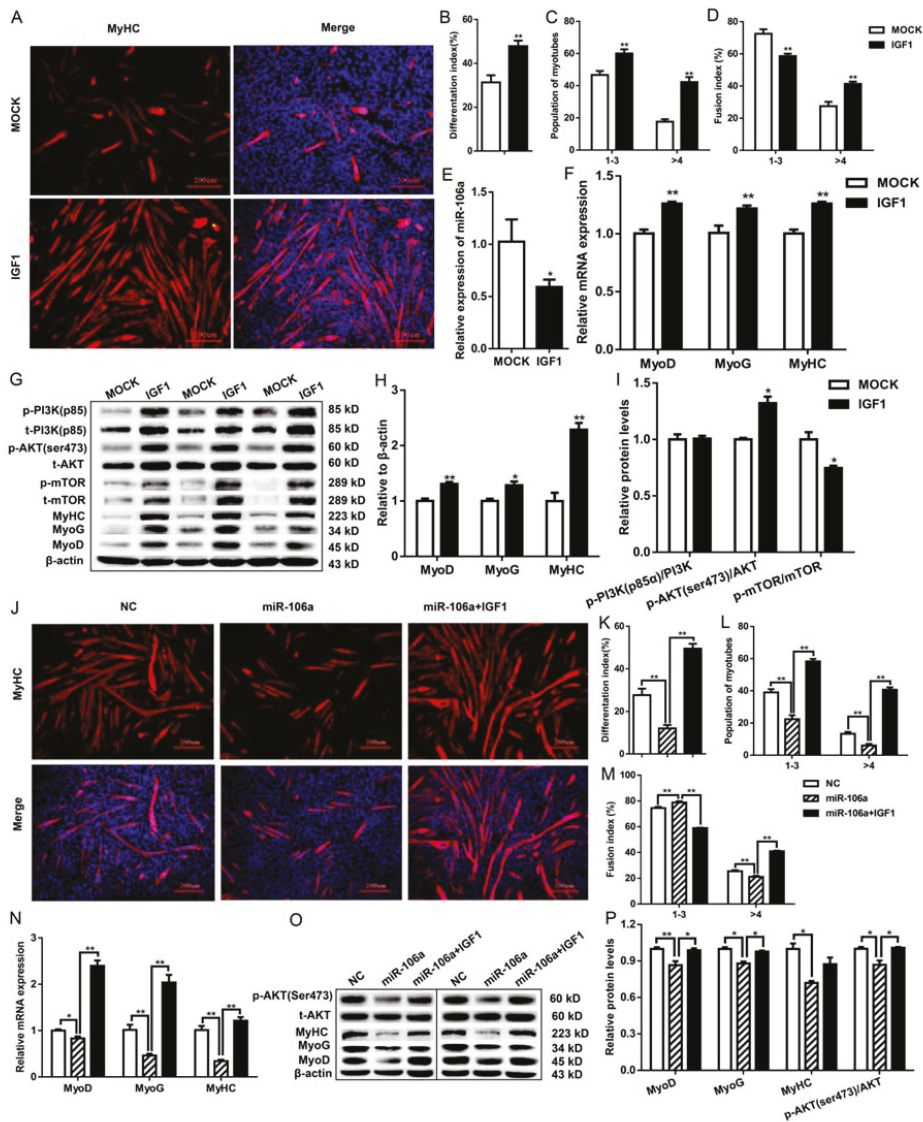


Figure 3. Insuline-like growth factor (IGF1) antagonized the effects of miR-106a-5p on myogenic differentiation in C2C12. All data were collected from C2C12 myotubes 5 d post differentiation. (A) MyHC immuno-staining of cells incubated in differentiation medium with or without 75 μ M IGF1 recombinant protein. Scale bars = 200 μ m; (B–D) The statistical results of differentiation index, fusion index and the populations of myotubes, respectively; (E) The expression of miR-106a-5p upon IGF1 stimuli; (F) The mRNA levels of MyoD, MyoG, MyHC in myoblasts cultured in differentiation medium containing 75 μ M IGF1 recombinant protein or not; (G) Western-blot analysis of myogenic regulatory factors (MyoD, MyoG, MyHC) and key molecules in PI3K-AKT pathway in cells; (H,I) The statistical results of Figure 3G; (J) Immunostaining of MyHC (red) and DAPI (blue). Scale bars = 200 μ m; (K–M) The statistical results of differentiation index, fusion index and the populations of myotubes, respectively; (N) The mRNA levels of MyoD, MyoG, MyHC; (O) Western blot analysis for MyoD, MyoG, MyHC and p-AKT/AKT upon IGF1 challenge; (P) Statistical analysis of MyoD, MyoG, MyHC and p-AKT/AKT levels in Figure 3O. Data were presented as mean \pm SEM. $n = 3$ per group. * $p < 0.05$, ** $p < 0.01$.

3.3. MiR-106a-5p Contributes to C2C12 Myotubes Atrophy by Suppressing PI3K-AKT Signaling Pathway

Expressions of miR-106a-5p, as well as MAFbx and MuRF1 in C2C12 myotubes, were significantly elevated with the treatment of 50 μ M DEX (Figure 4A). DEX significantly reduced the diameter of C2C12 myotubes (Figure 4B,C), indicating muscle atrophy. In addition, expressions of miR-106a-5p together with MAFbx, not MuRF1 were much higher in tibialis anterior (TA) muscles of six-month mice than that of two-month mice (Figure 4D).

In well-differentiated C2C12 myotubes, miR-106a-5p agomir was transfected to confirm its function in regulating myotubes atrophy. The overexpression efficiency of miR-106a-5p agomir was 30,000-fold higher than NC (Figure 4E), and the FAM-labeled miR-106a-5p agomir could be observed in almost all myotubes (Figure 4F). Moreover, enforced miR-106a-5p expression significantly decreased the diameter of C2C12 myotubes (Figure 4G,H) and increased the expression of MAFbx both at mRNA and protein levels (Figure 4I,J,L). The protein levels of another atrophy marker, MuRF-1 were also significantly enhanced by miR-106a-5p agomir (Figure 4J,L). Again, enhanced miR-106a-5p expression significantly repressed the phosphorylation of AKT (ser473) (Figure 4M). Together, these results indicated that miR-106a-5p promote C2C12 myotubes atrophy by repressing PI3K/AKT signaling pathway.

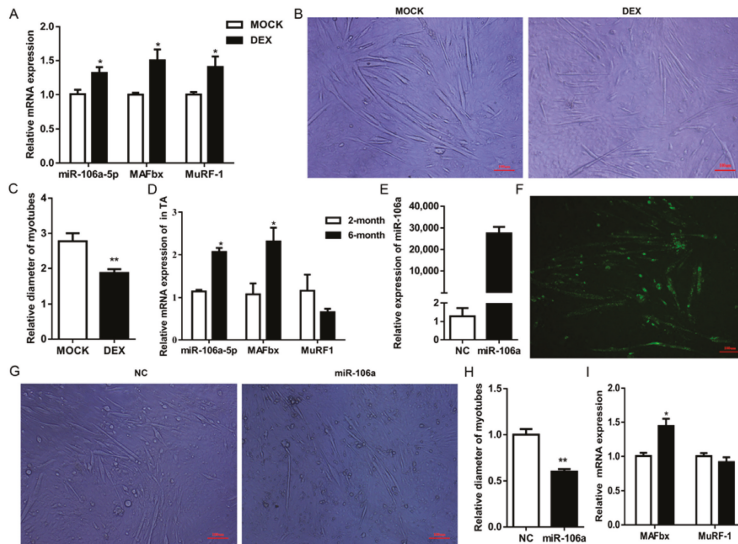


Figure 4. Cont.

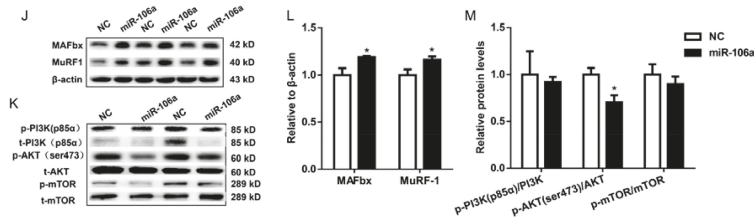


Figure 4. MiR-106a-5p contributed to myotubes atrophy. (A) Expression level of miR-106a-5p, MAFbx and MuRF1 mRNA after incubated with Dexamethasone (DEX) 36 h; (B) Morphological change of myotubes after DEX treatment 36 h (×20). Scale bars = 100 μm; (C) Relative diameter of myotube in Figure 4B; (D) RT-qPCR determined the expression of miR-106a-5p, MAFbx and MuRF1 in tibialis anterior (TA) muscle from two-month and six-month old mice (n = 5 per group); (E) Expression level of miR-106a-5p in C2C12 myotubes transfected with miR-106a-5p agomir or negative control for 36 h; (F) The fluorescent microscopy image of C2C12 cells transfected with FAM-labeled miR-106a-5p agomir. Scale bars= 100 μm; (G) The myotubes white light images after being transfected 36 h (×20); (H) Relative diameter of myotube from Figure 4G; (I) RT-qPCR analysis showed up-regulated MAFbx expression in C2C12 myotubes transfected with miR-106a-5p agomir 36 h; (J,L) Western blot showed that miR-106a-5p positively regulated analysis MAFbx and MuRF1 in C2C12 myotubes; (K,M) miR-106a-5p suppressed PI3K/AKT pathway in C2C12 differentiated myotubes. Results are presented as the mean ± SEM. n = 3 per group. * p < 0.05, ** p < 0.01.

3.4. PI3K (p85α) Is a Target Gene of miR-106a-5p in Differentiating and Well-Differentiated C2C12 Cells

PIK3R1, encoding PI3K (p85α) protein, is predicted to be a putative target of miR-106a-5p by TargetScan (<http://www.targetscan.org/>). Therefore, the 3' untranslated region (3'UTR) of PIK3R1 (W: wild-type; M: mutant type) was cloned into psi-CHECKTM-2 backbone (Figure 5A). In addition, transfection of miR-106a-5p mimics significantly repressed the expression levels of PI3K (p85α) protein (Figure 5B,C) both in differentiating and well-differentiated C2C12 cells. Furthermore, luciferase assays confirmed that miR-106a-5p could bind to the wild-type of 3'UTR of PIK3R1, instead of the mutated one (Figure 5D).

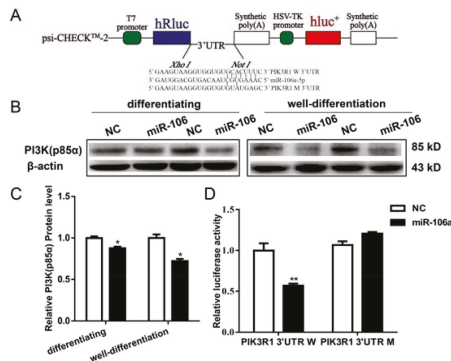


Figure 5. PI3KR1 was demonstrated to be a target of miR-106a-5p in C2C12 cells. (A) The schematic map of psiCHECK-PIK3R1 3'UTR luciferase reporter constructs (hRluc: synthetic *Renilla* luciferase gene; HSV-TK: herpes simplex virus type 1 thymidine kinase promoter; hLuc+: synthetic firefly luciferase gene; W: wild-type; M: mutant type); (B,C) Western blot analysis the expression of PI3K (p85α) protein (n = 3 per group); differentiating, 3 d post differentiation; well-differentiated, 36 h post transfection of myotubes; (D) The psiCHECK-PIK3R1 3'UTR plasmid were co-transfected into 293 T-cells with miR-106a-5p agomir or NC. Luciferase activities were measured 48 h after transfection (n = 5 per group). Data were presented as mean ± SEM. * p < 0.05, ** p < 0.01.

4. Discussion

Previous studies reveal a critical requirement for miR-106a at the early mammalian development. MiR-106a is differentially expressed in developing mice embryos and functions to control differentiation of stem cells [32]. In addition, miR-106a is down-regulated in myoblasts differentiation [22,23]. In our study, miR-106a-5p was observed to decrease in a time-dependent manner during myogenic differentiation in C2C12 cells and is relatively lower expressed in adult skeletal muscle. These data indicate that miR-106a-5p might be a negative regulator for myogenesis. However, there was no significant difference in the expression of miR-106a-5p in fast EDL and slow SOL muscle, suggesting miR-106a-5p might have limited effects on muscle fiber type transition.

Given that miR-106a-5p was down-regulated in C2C12 myogenic differentiation, miR-106a-5p agomir was used to explore the effects of miR-106a-5p on myogenesis. Here, transfection of miR-106a-5p agomir significantly reduced the differentiation index, fusion index as well as the expression levels of myogenic markers, suggesting miR-106a-5p could dramatically repress myogenic differentiation. Meanwhile, enforced expression of miR-106a-5p significantly reduced the level of p-AKT (ser473) in C2C12 myoblasts. It has been well documented that PI3K/AKT is a crucial signaling pathway to promote myogenesis and induce muscle hypertrophy [33,34]. IGF1 is a stimulator of PI3K/AKT signaling during muscle differentiation [35]. In our study, administration of 75 nM IGF1 recombinant protein totally reversed the inhibitory effects of miR-106a-5p. Taken together, miR-106a-5p might interrupt C2C12 myoblasts differentiation by disrupting AKT activity.

Furthermore, miR-106a-5p is previously reported to be upregulated in limb-girdle muscular dystrophies types 2A and 2B and is involved in muscular disorders [36]. In the present study, miR-106a-5p accompanied with atrophic factors were up-regulated in aged muscles and DEX-treated myotubes, and overexpression of miR-106a-5p was sufficient to reduce the diameters of well-differentiated myotubes in vitro, indicating miR-106a-5p might be involved in muscle atrophy. During miR-106a-5p-induced muscle atrophy, decreased phosphorylated AKT (ser473) was observed. Similarly, miR-18a, a member of miR-17-92 cluster, is demonstrated to decrease during myogenic differentiation and promotes muscle atrophy by targeting IGF1 [37]. Collectively, our data indicate that miR-106a-5p might induce myotube atrophy through blocking the AKT signaling pathway.

In our study, pan-PI3K (p85 α) protein levels were repressed by enforced miR-106a-5p agomir in both differentiating and well-differentiated myotubes. PI3K (p85 α), encoded by *PIK3R1* gene, is a regulatory subunit of PI3Ks and essential for myoblasts proliferation and differentiation [24,38]. Germline deletion of the *PIK3R1* gene leads to impaired muscle growth, and a significant reduction in muscle weight and fiber size [39]. Moreover, *PI3KR1* was identified as the target of miR-128a [40] and miR-29b in muscle cells [41]. Here, the luciferase reporter assay showed that miR-106a-5p directly bound to the *PIK3R1* 3' UTR, and *PI3KR1* is a novel target of miR-106a-5p in C2C12 cells.

In conclusion, miR-106a-5p is identified as a novel repressor of myogenesis, and miR-106a-5p represses differentiation and promotes atrophy by blocking the PI3K-AKT signaling pathway through targeting *PIK3R1*.

Author Contributions: Y.Z. performed experiments, analyzed data and drafted the manuscript, H.Z., G.M. and A.X. gave lots of help in cell culture and transfection, X.E.S. and G.S.Y. gave critical comments on this work, X.L. and G.W. edited and revised manuscript, S.S. approved final version of manuscript.

Funding: This work was supported by the National Key Technology R and D Program of China (2015BAD03B01-10), the Scientific and Technological Innovation Project of Shaanxi Province (2013KTCL02-04), National Natural Science Foundation (31501925), and Natural Science Foundation of Qinghai Province (2015-ZJ-920Q).

Acknowledgments: Xin'E Shi and Guofang Wu are greatly thanked for helpful comments on the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Cardinet, G.H. Skeletal muscle function. In *Clinical Biochemistry of Domestic Animals*, 5th ed.; Elsevier: Amsterdam, The Netherlands, 1997; pp. 407–440.
- Le Grand, F.; Rudnicki, M.A. Skeletal muscle satellite cells and adult myogenesis. *Curr. Opin. Cell Biol.* **2007**, *19*, 628–633. [[CrossRef](#)] [[PubMed](#)]
- Comai, G.; Tajbakhsh, S. Molecular and cellular regulation of skeletal myogenesis. In *Current Topics in Developmental Biology*; Elsevier: Amsterdam, The Netherlands, 2014; Volume 110, pp. 1–73.
- Zammit, P.S. *Function of the Myogenic Regulatory Factors Myf5, Myod, Myogenin and Mrf4 in Skeletal Muscle, Satellite Cells and Regenerative Myogenesis*; Seminars in Cell & Developmental Biology; Elsevier: Amsterdam, The Netherlands, 2017.
- Von Maltzahn, J.; Chang, N.C.; Bentzinger, C.F.; Rudnicki, M.A. Wnt signaling in myogenesis. *Trends Cell Biol.* **2012**, *22*, 602–609. [[CrossRef](#)] [[PubMed](#)]
- Zhang, P.; Li, W.; Wang, L.; Liu, H.; Gong, J.; Wang, F.; Chen, X. Salidroside inhibits myogenesis by modulating p-Smad3-induced Myf5 transcription. *Front. Pharmacol.* **2018**, *9*, 209. [[CrossRef](#)] [[PubMed](#)]
- Bhushan, R.; Grünhagen, J.; Becker, J.; Robinson, P.N.; Ott, C.-E.; Knaus, P. MiR-181a promotes osteoblastic differentiation through repression of TGF- β signaling molecules. *Int. J. Biochem. Cell Biol.* **2013**, *45*, 696–705. [[CrossRef](#)] [[PubMed](#)]
- Jang, Y.-N.; Baik, E.J. JAK-STAT pathway and myogenic differentiation. *Jak-Stat* **2013**, *2*, e23282. [[CrossRef](#)] [[PubMed](#)]
- Chen, X.; Wan, J.; Yu, B.; Diao, Y.; Zhang, W. Pip5k1 α promotes myogenic differentiation via AKT activation and calcium release. *Stem Cell Res. Ther.* **2018**, *9*, 33. [[CrossRef](#)] [[PubMed](#)]
- Xu, Q.; Wu, Z. The insulin-like growth factor-phosphatidylinositol 3-kinase-AKT signaling pathway regulates myogenin expression in normal myogenic cells but not in rhabdomyosarcoma-derived RD cells. *J. Biol. Chem.* **2000**, *275*, 36750–36757. [[CrossRef](#)] [[PubMed](#)]
- Zhai, L.; Wu, R.; Han, W.; Zhang, Y.; Zhu, D. MiR-127 enhances myogenic cell differentiation by targeting S1pPR3. *Cell Death Dis.* **2017**, *8*, e2707. [[CrossRef](#)] [[PubMed](#)]
- Feng, Y.; Niu, L.; Wei, W.; Zhang, W.; Li, X.; Cao, J.; Zhao, S. A feedback circuit between miR-133 and the ERK1/2 pathway involving an exquisite mechanism for regulating myoblast proliferation and differentiation. *Cell Death Dis.* **2013**, *4*, e934. [[CrossRef](#)] [[PubMed](#)]
- Goljanek-Whysall, K.; Mok, G.; Alrefaei, A.; Kennerley, N.; Wheeler, G.; Munsterberg, A. The coordinated regulation of BAF60 variants by miR-1/206 and miR-133 clusters stabilises myogenic differentiation during embryogenesis. *Int. J. Exp. Pathol.* **2015**, *96*, A8–A9.
- Ma, M.; Wang, X.; Chen, X.; Cai, R.; Chen, F.; Dong, W.; Yang, G.; Pang, W. MicroRNA-432 targeting E2F3 and P55PIK inhibits myogenesis through PI3K/AKT/mTOR signaling pathway. *RNA Biol.* **2017**, *14*, 347–360. [[CrossRef](#)] [[PubMed](#)]
- Jebessa, E.; Ouyang, H.; Abdalla, B.A.; Li, Z.; Abdullahi, A.Y.; Liu, Q.; Nie, Q.; Zhang, X. Characterization of miRNA and their target gene during chicken embryo skeletal muscle development. *Oncotarget* **2018**, *9*, 17309. [[CrossRef](#)] [[PubMed](#)]
- Zhang, P.; Xu, H.; Li, R.; Wu, W.; Chao, Z.; Li, C.; Xia, W.; Wang, L.; Yang, J.; Xu, Y. Assessment of myoblast circular RNA dynamics and its correlation with miRNA during myogenic differentiation. *Int. J. Biochem. Cell Biol.* **2018**, *99*, 211–218. [[CrossRef](#)] [[PubMed](#)]
- Khuu, C.; Jevnaker, A.-M.; Bryne, M.; Osmundsen, H. An investigation into anti-proliferative effects of microRNAs encoded by the miR-106a-363 cluster on human carcinoma cells and keratinocytes using microarray profiling of miRNA transcriptomes. *Front. Genet.* **2014**, *5*, 246. [[CrossRef](#)] [[PubMed](#)]
- Khuu, C.; Utheim, T.P.; Sehic, A. The three paralogous microRNA clusters in development and disease, miR-17-92, miR-106a-363, and miR-106b-25. *Scientifica* **2016**, *2016*, 1379643. [[CrossRef](#)] [[PubMed](#)]
- Sengupta, D.; Govindaraj, V.; Kar, S. Alteration in microRNA-17-92 dynamics accounts for differential nature of cellular proliferation. *FEBS Lett.* **2018**, *592*, 446–458. [[CrossRef](#)] [[PubMed](#)]
- Qiu, H.; Liu, N.; Luo, L.; Zhong, J.; Tang, Z.; Kang, K.; Qu, J.; Peng, W.; Liu, L.; Li, L. MicroRNA-17-92 regulates myoblast proliferation and differentiation by targeting the ENH1/ID1 signaling axis. *Cell Death Differ.* **2016**, *23*, 1658. [[CrossRef](#)] [[PubMed](#)]

21. Luo, W.; Li, G.; Yi, Z.; Nie, Q.; Zhang, X. E2f1-miR-20a-5p/20b-5p auto-regulatory feedback loop involved in myoblast proliferation and differentiation. *Sci. Rep.* **2016**, *6*, 27904. [[CrossRef](#)] [[PubMed](#)]
22. Tang, Z.; Liu, N.; Luo, L.; Kang, K.; Li, L.; Ni, R.; Qiu, H.; Gou, D. MicroRNA-17-92 regulates the transcription factor E2F3b during myogenesis in vitro and in vivo. *Int. J. Mol. Sci.* **2017**, *18*, 727. [[CrossRef](#)] [[PubMed](#)]
23. Imig, J.; Brunschweiler, A.; Brümmer, A.; Guennewig, B.; Mittal, N.; Kishore, S.; Tsirikka, P.; Gerber, A.P.; Zavolan, M.; Hall, J. MiR-clip capture of a miRNA targetome uncovers a lincRNA h19–miR-106a interaction. *Nat. Chem. Biol.* **2015**, *11*, 107. [[CrossRef](#)] [[PubMed](#)]
24. Mellor, P.; Furber, L.A.; Nyarko, J.N.; Anderson, D.H. Multiple roles for the p85 α isoform in the regulation and function of PI3K signalling and receptor trafficking. *Biochem. J.* **2012**, *441*, 23–37. [[CrossRef](#)] [[PubMed](#)]
25. Madhala-Levy, D.; Williams, V.; Hughes, S.; Reshef, R.; Halevy, O. Cooperation between Shh and IGF-I in promoting myogenic proliferation and differentiation via the MAPK/ERK and PI3K/Akt pathways requires smo activity. *J. Cell. Physiol.* **2012**, *227*, 1455–1464. [[CrossRef](#)] [[PubMed](#)]
26. Schiaffino, S.; Mammucari, C. Regulation of skeletal muscle growth by the IGF1-Akt/PKB pathway: Insights from genetic models. *Skelet. Muscle* **2011**, *1*, 4. [[CrossRef](#)] [[PubMed](#)]
27. Briata, P.; Lin, W.-J.; Giovarelli, M.; Pasero, M.; Chou, C.-F.; Trabucchi, M.; Rosenfeld, M.G.; Chen, C.-Y.; Gherzi, R. PI3K/AKT signaling determines a dynamic switch between distinct KSRP functions favoring skeletal myogenesis. *Cell Death Differ.* **2012**, *19*, 478. [[CrossRef](#)] [[PubMed](#)]
28. Conejo, R.; Lorenzo, M. Insulin signaling leading to proliferation, survival, and membrane ruffling in C2C12 myoblasts. *J. Cell. Physiol.* **2001**, *187*, 96–108. [[CrossRef](#)]
29. Yu, M.; Wang, H.; Xu, Y.; Yu, D.; Li, D.; Liu, X.; Du, W. Insulin-like growth factor-1 (IGF-1) promotes myoblast proliferation and skeletal muscle growth of embryonic chickens via the PI3K/Akt signalling pathway. *Cell Biol. Int.* **2015**, *39*, 910–922. [[CrossRef](#)] [[PubMed](#)]
30. Rommel, C.; Bodine, S.C.; Clarke, B.A.; Rossmann, R.; Nunez, L.; Stitt, T.N.; Yancopoulos, G.D.; Glass, D.J. Mediation of IGF-1-induced skeletal myotube hypertrophy by PI(3)K/Akt/mTOR and PI(3)K/Akt/GSK3 pathways. *Nat. Cell Biol.* **2001**, *3*, 1009. [[CrossRef](#)] [[PubMed](#)]
31. Zhang, Q.; Shi, X.-E.; Song, C.; Sun, S.; Yang, G.; Li, X. Bambi promotes C2C12 myogenic differentiation by enhancing Wnt/ β -catenin signaling. *Int. J. Mol. Sci.* **2015**, *16*, 17734–17745. [[CrossRef](#)] [[PubMed](#)]
32. Foshay, K.M.; Gallicano, G.I. MiR-17 family miRNAs are expressed during early mammalian development and regulate stem cell differentiation. *Dev. Biol.* **2009**, *326*, 431–443. [[CrossRef](#)] [[PubMed](#)]
33. Glass, D.J. PI3 kinase regulation of skeletal muscle hypertrophy and atrophy. In *Phosphoinositide 3-Kinase in Health and Disease*; Springer: Berlin, Germany, 2010; pp. 267–278.
34. Serra, C.; Palacios, D.; Mozzetta, C.; Forcales, S.V.; Morante, I.; Ripani, M.; Jones, D.R.; Du, K.; Jhala, U.S.; Simone, C. Functional interdependence at the chromatin level between the MKK6/p38 and IGF1/PI3K/AKT pathways during muscle differentiation. *Mol. Cell* **2007**, *28*, 200–213. [[CrossRef](#)] [[PubMed](#)]
35. Miyata, S.; Yada, T.; Ishikawa, N.; Taheruzzaman, K.; Hara, R.; Matsuzaki, T.; Nishikawa, A. Insulin-like growth factor 1 regulation of proliferation and differentiation of *Xenopus laevis* myogenic cells in vitro. *In Vitro Cell. Dev. Biol. Anim.* **2017**, *53*, 231–247. [[CrossRef](#)] [[PubMed](#)]
36. Eisenberg, I.; Eran, A.; Nishino, I.; Moggio, M.; Lamperti, C.; Amato, A.A.; Lidov, H.G.; Kang, P.B.; North, K.N.; Mitrani-Rosenbaum, S. Distinctive patterns of microRNA expression in primary muscular disorders. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 17016–17021. [[CrossRef](#)] [[PubMed](#)]
37. Liu, C.; Wang, M.; Chen, M.; Zhang, K.; Gu, L.; Li, Q.; Yu, Z.; Li, N.; Meng, Q. MiR-18a induces myotubes atrophy by down-regulating IGF1. *Int. J. Biochem. Cell Biol.* **2017**, *90*, 145–154. [[CrossRef](#)] [[PubMed](#)]
38. Ito, Y.; Vogt, P.K.; Hart, J.R. Domain analysis reveals striking functional differences between the regulatory subunits of phosphatidylinositol 3-kinase (PI3K), p85 α and p85 β . *Oncotarget* **2017**, *8*, 55863. [[CrossRef](#)] [[PubMed](#)]
39. Luo, J.; Sobkiw, C.L.; Hirshman, M.F.; Logsdon, M.N.; Li, T.Q.; Goodyear, L.J.; Cantley, L.C. Loss of class I α PI3K signaling in muscle leads to impaired muscle growth, insulin response, and hyperlipidemia. *Cell Metab.* **2006**, *3*, 355–366. [[CrossRef](#)] [[PubMed](#)]

40. Motohashi, N.; Alexander, M.S.; Shimizu-Motohashi, Y.; Myers, J.A.; Kawahara, G.; Kunkel, L.M. Regulation of IRS1/AKT insulin signaling by microRNA-128a during myogenesis. *J. Cell Sci.* **2013**, *126*, 2678–2691. [[CrossRef](#)] [[PubMed](#)]
41. Li, J.; Chan, M.C.; Yu, Y.; Bei, Y.; Chen, P.; Zhou, Q.; Cheng, L.; Chen, L.; Ziegler, O.; Rowe, G.C. MiR-29b contributes to multiple types of muscle atrophy. *Nat. Commun.* **2017**, *8*, 15201. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

Somatostatin Analogue Treatment Primarily Induce miRNA Expression Changes and Up-Regulates Growth Inhibitory miR-7 and miR-148a in Neuroendocrine Cells

Kristina B. V. Døssing ^{1,2,3,†}, Christina Kjær ^{4,†}, Jonas Vikeså ¹, Tina Binderup ^{2,3}, Ulrich Knigge ^{5,6}, Michael D. Culler ⁷, Andreas Kjær ^{2,3}, Birgitte Federspiel ⁸ and Lennart Friis-Hansen ^{1,9,*}

¹ Center for Genomic Medicine, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark; kbdoessing@sund.ku.dk (K.B.V.D.); jonas.vikesaa@roche.com (J.V.)

² Department of Clinical Physiology, Nuclear Medicine and PET, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark; tina.binderup@rh.regionh.dk (T.B.); andreas.kjaer@rh.regionh.dk (A.K.)

³ Cluster for Molecular Imaging, Faculty of Health Sciences University of Copenhagen, Blegdamsvej 3B, 2100 Copenhagen, Denmark

⁴ University College Copenhagen, Sigurdsgade 26, 2200 Copenhagen, Denmark; chkj@phmetropol.dk

⁵ Department of Surgical Gastroenterology C, Rigshospitalet, University of Copenhagen, Blegdamsvej 9, 2100 Copenhagen, Denmark; ulrich.knigge@rh.regionh.dk

⁶ Department of Clinical Endocrinology PE, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark

⁷ Biomeasure Incorporated/IPSEN, 650 W Kendall St, Cambridge, MA 02142, USA; mdculler@comcast.net

⁸ Department of Pathology, Rigshospitalet, University of Copenhagen, Blegdamsvej 9, 2100 Copenhagen, Denmark; birgitte.federspiel@rh.regionh.dk

⁹ Department of Clinical Biochemistry, Hillerød Hospital, Dyrhavevej 29, 3400 Hillerød, Denmark

* Correspondence: Lennart.jan.friis-hansen@regionh.dk; Tel.: +45-215437777

† The two authors contributed equally.

Received: 22 May 2018; Accepted: 2 July 2018; Published: 4 July 2018

Abstract: Somatostatin (SST) analogues are used to control the proliferation and symptoms of neuroendocrine tumors (NETs). MicroRNAs (miRNA) are small non-coding RNAs that modulate posttranscriptional gene expression. We wanted to characterize the miRNAs operating under the control of SST to elucidate to what extent they mediate SST actions. NCI-H727 carcinoid cell line was treated with either a chimeric SST/dopamine analogue; a SST or dopamine analogue for proliferation assays and for identifying differentially expressed miRNAs using miRNA microarray. The miRNAs induced by SST analogue treatment are investigated in carcinoid cell lines NCI-H727 and CNDT2 using in situ hybridization, qPCR and proliferation assays. SST analogues inhibited the growth of carcinoid cells more potently compared to the dopamine analogue. Principal Component Analysis (PCA) of the samples based on miRNA expression clearly separated the samples based on treatment. Two miRNAs which were highly induced by SST analogues, miR-7 and miR-148a, were shown to inhibit the proliferation of NCI-H727 and CNDT2 cells. SST analogues also produced a general up-regulation of the let-7 family members. SST analogues control and induce distinct miRNA expression patterns among which miR-7 and miR-148a both have growth inhibitory properties.

Keywords: somatostatin analogues; neuroendocrine tumors; cancer; miRNAs; miR-148a; miR-7; let-7

1. Introduction

Gastro-Entero-Pancreatic neuroendocrine neoplasms (GEP-NEN) are generally slow growing tumors originating from neuroendocrine cells in the gastro-intestinal tract, the diffuse neuroendocrine

system, pancreas and the bronco-pulmonary system [1]. According to the World Health Organization (WHO) 2010 classification [2] GEP-NEN can be classified according to Ki67 index and/or mitotic count into neuroendocrine tumors (NETs, NET-G1 and NET-G2) and neuroendocrine carcinomas (NEC G3) [2,3]. The bronco-pulmonary NENs are classified according to mitoses pr. 2 mm² and the presence or absence of necrosis according to the 2015 WHO classification [4]. This separates small-cell and large-cell, poorly differentiated based on mitosis only, and the typical carcinoids (TCs) and atypical carcinoids (ACs) by <2 mitoses per 2 mm²/no necrosis and 2–10 mitoses per 2 mm²/with focal necrosis. In general, NETs of the lung comprise <20% of all lung cancers, and TCs and ACs comprise 1–2% of these [4,5]. The yearly incidence of GEP-NENs and bronco-pulmonary NENs is 2–3 and 0.2–2 per 1,000,000 inhabitants respectively, but this has been increasing during the last decades [1,6–8]. TCs share homologies with G1 GEP-NENs and ACs with G2 GEP-NENs coinciding with the fact that ACs are more malignant in nature and more prone to metastasize compared to TCs [8,9]. The clinical picture depends on the site of the primary tumor and its ability to secrete neuroamines and/or peptides at supra-physiological levels (functioning tumors) as the clinical syndrome depends on the transmitter/hormone secreted [10]. However, most of the GEP-NENs and bronco-pulmonary NENs are silent or ‘non-functional’, as they either do not secrete any transmitters/hormones or the one they secrete does not cause clinical symptoms [10,11]. The lack of endocrine symptoms in these patients often delays the diagnosis until the presence of symptoms caused by the mass effect and/or the presence of metastases, mainly hepatic metastases [12]. In patients with localized low grade GEP-NETs, the 5-year survival rate is 60–100% whereas patients with regional disease or distant metastases have 40% and 29% 5-year survival rates respectively [13]. In bronco-pulmonary NENs the overall survival is higher in TCs versus ACs with a 5-year survival rate of 90% and 60% respectively and increases with resectable tumors [14].

A common feature for endocrine cells is that both their secretion and growth can be inhibited by somatostatin (SST) [15]. In the nervous tissue, SST acts like a neurotransmitter and a neuromodulator. In the gastrointestinal tract, SST is expressed by specialized endocrine cells and acts as a paracrine factor [16]. The somatostatin receptors (SSTRs) belong to the G-protein coupled receptor family and there are five different SSTR subtypes (SSTR1–5), all differently expressed by neuroendocrine cells [17]. Seventy to ninety percent of all GEP-NET express SSTRs and tumors expressing SSTRs often contain one or more receptor subtypes, most often SSTR1, 2, 3 and 5, whereas SSTR4 is less commonly seen in endocrine tumors [18]. This has made SST useful for both visualization and treatment [19,20]. The expression is also related to the tumor grade, since the receptors are preferably expressed in low grade tumors, whereas the expression of some receptor subtypes is reduced in the more dedifferentiated tumors [21]. The different SSTR subtypes all bind the endogenous ligands; SST14, SST 28 and cortistatin; with equally high affinity in the nanomolar (nM) range. The widespread expression of SSTRs by a large number of human tumors and the biological actions of SSTRs form the basis for in vivo tumor targeting. However, the short half-life in circulation (1–3 min) of the endogenous SST peptides makes them unsuited for therapy [22]. Therefore a number of synthetic analogues with longer half-lives, such as somatostatin SMS201-995 (Octreotide), RC-160 (Vapreotide), BIM-23014 (Lanreotide), MK678 (Seglitide) and SOM-230 (Pasireotide LAR), were developed and are currently in clinical use [23,24]. Surgical removal is the optimal treatment for GEP-NENs and bronco-pulmonary NENs, and somatostatin analogue (SSA) treatment constitutes the gold standard for symptomatic and anti-proliferative control [8,25].

MicroRNAs (miRNAs) are small non-coding RNA molecules (on average 20–23 nucleotide (nt) long) that modulate gene expression by binding to complementary sequences on target messenger RNA (mRNAs). This predominantly results in a decrease of target mRNA levels [26]. An increasing number of studies indicate that miRNAs are involved in many important biological processes, including proliferation, apoptosis, differentiation, angiogenesis and immune response. miRNA deregulation leads to aberrant gene expression in various diseases and dysregulation of miRNA expression have

been shown to be involved in cancer development [27]. We therefore wanted to examine which miRNA operate under the control of SST and if they contribute to the growth- inhibitory effects of SST.

2. Materials and Methods

2.1. Cell Lines and Tissue Culture

Four human carcinoid cell lines, two intestinal (CNDT2 and HC45) and two of pulmonary origin (NCI-H720 and NCI-H727) were used for experiments.

CNDT2 is a human midgut carcinoid cell line kindly provided by Lee M. Ellis M.D. Anderson Center Texas USA [28] and grown in DMEM/F12 with 15 mM HEPES (Life Technologies, Carlsbad, CA, USA) supplemented with 10% FBS (Th. Geyer GmbH, Stuttgart, Germany), penicillin 100 U/mL and streptomycin 100 µg/mL (Life Technologies), 5 mL Sodium pyruvate 100 mM (Sigma-Aldrich, St. Louis, MO, USA), 5 mL MEM NEAA 100x (Life Technologies), 5 mL L-Glutamine 200 mM 100x (Life) and 10 ng/mL NGF (Life Technologies) and kept at 37 °C/5% CO₂. HC45 is a human ileal carcinoid cell line kindly provided by Ricardo V. Lloyd Mayo Clinic [29] and kept in RPMI 1640/Glutamax (GIBCO, Waltham, MA, USA) supplemented with 10% FBS, 1% P/S and 10 ng/mL Insulin (Invitrogen, Carlsbad, CA, USA) at 37 °C and 5% CO₂.

NCI-H720 is an atypical and NCI-H727 is a typical human pulmonary carcinoid cell line obtained from ATCC (Boras, Sweden). Both cell lines were cultured in RPMI 1640 Glutamax supplemented with 10% FBS, penicillin 100 U/mL and streptomycin 100 µg/mL, 1 mM Sodium Pyruvate (Life Technologies) at 37 °C and 5% CO₂.

For experiments involving seeding cells into new plates, cells were always allowed to adhere overnight.

2.2. Tumor Tissue

Five formalin-fixed and paraffin-embedded (FFPE) tissue samples of carcinoid tumors (see Table 1), obtained from the Department of Pathology (Rigshospitalet, Copenhagen, Denmark), were used for Laser Capture Microdissection, qPCR and in situ hybridization, see descriptions for the procedures below.

Table 1. Tissue used for immunohistochemistry, in situ hybridization and qPCR.

Patient	Age/Sex	Location	Ki67 Index	Treatment
1	65/F	Colon	1%	Surgery
2	69/F	Small intestine	3%	Surgery/SSA
3	62/M	Ileum	2%	Surgery
4	55/F	Small intestine	7–8%	Surgery
5	64/F	Ileocecal	1%	Surgery/SSA

2.3. Somatostatin Analogues

Two SST analogues (BIM-23014 (Lanreotide) and BIM-23023), a combined SST-dopamine 2 receptor analogue (BIM-23A760) and a pure dopamine receptor D2 (DRD2) analogue (BIM-53097) were used. The compounds' binding affinities for the SST and dopamine receptors are shown in (Table 2). The compounds were initially dissolved in 100 µL 99% EtOH then in 0.1 N acetic acid/0.1% BSA to a final stock concentration of 10⁻³ M and used for experiments at suitable concentrations in complete growth medium.

Table 2. Human Somatostatin Receptor Subtype Specificity (IC50-nM) [30,31].

Compound	Somatostatin Receptor Subtype					Dopamine
	1	2	3	4	5	
Somatostatin 14	1.95	0.25	1.2	1.77	1.4	
Somatostatin 28	1.86	0.31	1.3	5.4	0.4	
BIM-23014 (Lanreotide)	>1000	0.75	98	>1000	12.7	
BIM-23023	>1000	0.42	87	2.7	4.2	
BIM-53097						22.1
BIM-23A760	622	0.03	160	>1000	42.0	15

2.4. Transfection Studies and Cell Growth Analyses

A total of 4×10^6 NCI-H727 or 1.5×10^6 CNDT2 cells were seeded and used for each transfection. To 965 μ L Opti-MEM (Invitrogen) 10 μ L Negative control 1 or 2 (Thermo Scientific, Waltham, MA, USA) or mature miRNA miR-7/miR-148a (Applied Biosystems, Carlsbad, CA, USA) or inhibitor—miR-7 LNA/miR-148a LNA (Exiqon, Vedbæk, Denmark) was added to a final concentration of 50 nM together with 25 μ L Turbofect transfection reagent (Fermentas, Leon-Rot, Germany) in 5 mL complete growth medium and left to incubate for 15–20 min at room temperature (RT) before being added drop-wise to the cells.

For growth analyses, 4×10^4 cells (NCI-H727) or 1.5×10^4 cells (CNDT2) were seeded in each well into E-plates for use in the xCELLigence system (Roche/ACEA, San Diego, CA, USA) for proliferation studies. The xCELLigence analyzer, which is an electronic cell sensor array technology, allows label-free and real-time monitoring of cell proliferation. The presence of the cells on top of the electrodes will affect the local ionic environment at the electrode/solution interface, leading to an increase in the electrode impedance. The more cells that are attached to the electrodes, the larger the increases in electrode impedance. For further details, see references [32,33]. The difference in cell number seeding for growth assays is due to differences in size and proliferative rate between the two cell lines. A series of analogue concentrations were used to find the optimal concentration for the actual growth experiments. The analogues were added daily directly to the wells of the E-plates in complete growth medium used for the normal passage of cells, without changing it for the duration of the entire growth experiment. The concentrations of the analogues for growth experiments were BIM-23a760 10^{-7} M, BIM-23023 10^{-7} M, BIM-23014 10^{-9} M and finally BIM-53097 10^{-9} M.

2.5. RNA Extraction

Total RNA was extracted using Trizol reagent (Life Technologies) according to the manufactures specifications. The RNA concentration was measured on the NanoDrop (Thermo Fisher Scientific, Wilmington, DE, USA) and the integrity determined using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA).

2.6. Somatostatin Receptors mRNA Quantification

The expressions of the SSTRs were quantitated as previously described [34] and normalized to the expression of β -actin.

2.7. Laser Capture Microdissection

Tumor and normal cells from FFPE tissue, see tissue specifications above, were Laser Capture Microdissected (LCM) using an Arcturus LCM system (ThermoFisher Scientific, Waltham, MA, USA). 10 μ m sections of tissue were mounted on Pen Membrane slides (Applied Biosystems, Foster City, CA, USA). Sections were stained with cresyl violet, using the LCM staining kit (Ambion, Foster City, CA, USA/Life Technologies, Carlsbad, CA, USA) with cresyl violet according to the manufacturer's instructions. After tissue sections had been collected and transferred to the collection

cap, the cap was immediately transferred and clicked on to an Eppendorf tube containing the 100 μ L of the Lysis Solution used in the first step of the RNAqueous[®]-Micro Kit (ThermoFisher Scientific). The tube was inverted to ensure complete coverage of the dissected cells in the buffer. RNA isolation was performed according to the manufacturer's instructions.

2.8. qPCR of microRNA Expression

The expression of miR-7 and miR-148a was quantitated using TaqMan miRNA assay (Applied Biosystems). Briefly, 100 ng of the total RNA was reversely transcribed into cDNA using gene specific primers and the TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems) according to the manufacturer's protocol. Samples were run on the ABI PRISM 7900 HT Sequence Detection System (Applied Biosystems). For normalization of the miRNA expression data the geometric mean of hsa-miR-191 and RNU-44 were used [35]. Primer sequences are listed in Table 3.

Table 3. Primer sequences.

MiRNA	Mature Sequence	Product ID
miR-7	UGGAAGACUAGUGAUUUUGUUGU	000268
miR-148a	UCAGUGCACUACAGAACUUUGU	000470
MiR-191	CAACGGAAUCCAAAAGCAGCUG	002299
RNU-44	CCTGGATGATGATAGCAAATGCTG-ACTGAACATGAAGGTCTT AATTAGCTCTAACTGACT	001094

2.9. In Situ Hybridization

Formalin-fixed and paraffin-embedded (FFPE) tissue samples of carcinoid tumors were obtained from the Department of Pathology (Rigshospitalet, Copenhagen, Denmark). A double-DIG-labeled LNA-modified oligos miR-7 (Exiqon, Munich, Germany), probe sequence 5'-ACAACAAAATCACTAGTCTTCCA-3', RNA-Tm 80 °C and double-DIG-labeled LNA-modified oligos miR-148a (Exiqon), probe sequence 5'-ACAAAGTTCTGTAGTGCCTGA-3', RNA-Tm 80 °C were used for detection as described [36]. Probe concentration was 100 nM and slides were hybridized at 50 °C. Sections were counterstained with Nuclear Fast Red. Pictures of representative areas of the slides were taken with a Zeiss Axio Imager (Zeiss, Jena, Germany), original magnification $\times 20/10$. Cells with intense blue nuclear stain were scored as positive. The level of expression within a positive cell was not scored. A LNA probe against snRNA U6 (Exiqon) was used as positive control and a scramble probe (Exiqon) as negative control.

2.10. MicroRNA Microarray

For microarray analysis 1 μ g of total RNA was labeled using the Flashtag RNA labeling kit for Affymetrix (Genisphere LLC., Hatfield, PA, USA) according to the manufacturer's instructions. The labeled samples were hybridized to GeneChip miRNA Array (Affymetrix, Santa Clara, CA, USA). The Affymetrix miRNA array assay miRNAs includes small nucleolar RNAs (snoRNAs) and small Cajal Body specific RNAs (scaRNAs) in human. The 847 human miRNAs on the array are derived from Sanger miRbase miRNA database V11. Four copies of each miRNA probe are distributed on the array.

2.11. RNA Profiling

Arrays were washed and stained with phycoerythrin conjugated streptavidin (SAPE) using the Affymetrix Fluidics Station[®] 450, and the arrays were scanned in the Affymetrix GeneArray[®] 3000 scanner to generate fluorescent images, as described in the Affymetrix Gene Chip[®] protocol. To minimize batch variation, equal numbers of treatment groups were included in each batch.

2.12. Data Analysis

Raw data files were imported into Affymetrix's miRNA QC Tool (Affymetrix) and normalized using the quantiles normalization and median Polish summarization following a background correction that corrects for the GC content of the each particular probe. Log₂ intensities of the 847 human miRNAs were imported into the Data Analysis software package Qlucore Omics Explorer v2.1. Principal Component Analysis (PCA) visualization of the clustering of samples using the genes selected in the class comparison was performed using the build-in PCA tool in Qlucore Omics Explorer v2.1 (Qlucore AB, Lund, Sweden).

Class comparison analysis was performed using Students *t*-test. A multiple comparison test was used for all three tissue-types and a two-group comparison was used for comparing AP versus NF. The miRNA was defined as being differentially expressed between the compared groups if the *p* value was less than 0.05 and the fold change above 1.5.

2.13. Statistical Analyses

Students' unpaired *t*-test was used and differences with a $p \leq 0.05$ were considered significant and indicated by *. Unless otherwise stated results are given as median \pm standard deviation (SD).

3. Results

3.1. Somatostatin and Dopamine Analogues Inhibit the Growth of a Carcinoid Cell Line NCI-H727

We first examined the expression of the SSTRs in four different carcinoid cell lines, HC45 and CNDT2 (both intestinal) and the atypical NCI-H720 and typical NCI-H727 pulmonary cell lines, in order to choose an optimal cell line as a model system for examining the effect of SSAs on miRNA expression in NETs. All the carcinoid cell lines expressed SSTR subtypes 2 and 5 Figure 1A, and for further analyses we selected two of the cell lines with highest SSTR2 mRNA level and each representing common NET origins in lung (NCI-H727) and intestinal (CNDT2). Furthermore, the HC45 proved very difficult to grow even after having been immortalized by retroviral transfection with a constitutive active human Telomerase Reverse Transcriptase TERT expression vector and we discontinued using this cell line. We have also previously shown that the CNDT2 and NCI-H727 cell lines are good model systems when examining NETs [37] and proceed with these two cell lines as our model.

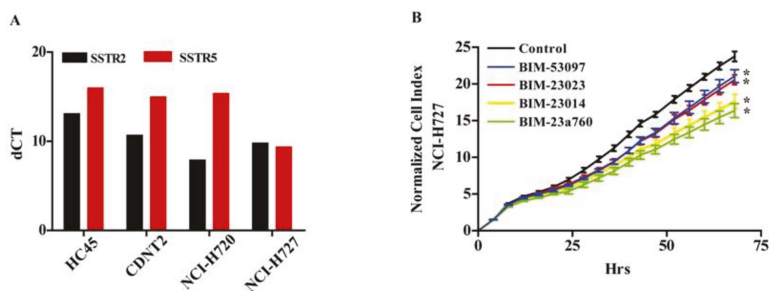


Figure 1. Carcinoid cell lines express all somatostatin receptors (SSRTs), and activation of both SSTRs and DRD2 gives the most potent growth inhibition of a carcinoid cell line (A) All the four examined carcinoid cell lines express SSTR subtype 2 and 5 when analyzed by qPCR (B) The growth of NCI-H727 is inhibited by somatostatin, dopamine and the chimeric somatostatin-dopamine agonists. The dopamine analogue BIM-53097 (blue) and the somatostatin analogues (SSAs) BIM—23023 (red) were the weaker inhibitors of cell growth. The somatostatin agonist BIM-23014 (yellow) and the chimeric somatostatin-dopamine compound BIM-23A760 (green) were the stronger inhibitors of carcinoid cell growth. The control (black) is vehicle without agonist.

Having shown that NCI-H727 cells expressed SSTR2- and five subtypes we examined the effect of the SST and dopamine analogues on their growth. We found that all compounds inhibited the growth of the carcinoid cell line (Figure 1B). The most potent inhibitor of proliferation was the combined SSTR and dopamine receptor DR agonist BIM-23A760 (green) followed by the SSTR agonist BIM-23014 (yellow) and BIM-23023 (red). The least effective of the SSAs was the DR agonist BIM-53097 (blue) (Figure 1B).

3.2. Somatostatin Receptor Activation Primarily Induces microRNA Expression Changes

After having seen that SST analogues inhibit growth we treated NCI-H727 cells with the most potent inhibitor—BIM-23A760—and compared the mRNA and miRNA profiles of treated cells with that of untreated cells (Figure 2). PCA plots based on either mRNAs or miRNAs of cells treated with SST analogue showed that PCA based miRNA expression profiles clearly separated controls from treated cells. In contrast, PCA based mRNA profiles did not (Figure 2). Furthermore, when calculating the number of transcripts changed and regulated by SST 24 h after treatment it is clear that treating with SST regulates miRNAs to a higher degree than mRNAs (Figure 2).

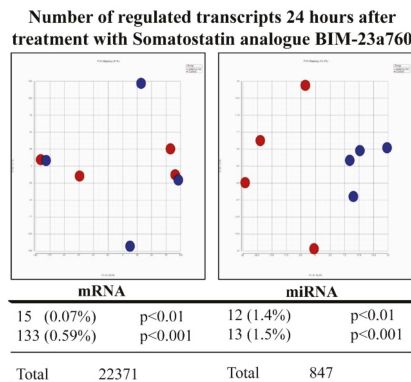


Figure 2. The dual somatostatin-dopamine agonist BIM-23a760 induces greater changes in microRNA (miRNA) than in messenger RNA (mRNA). Carcinoid NCI-H727 cells were treated with the dual somatostatin-dopamine agonist BIM-23a760 for 24 h and the changes in mRNA and miRNA were examined using microarray. Principal component analysis (PCA) of the changes in mRNA and miRNA demonstrated that mRNA based PCA did not clearly separate the BIM-23a760 treated cells (shown in red) from the control cells treated with vehicle without BIM-23a76 (shown in blue). In contrast a PCA based on the changes in miRNA expression clearly separated the treated cells (shown in red) from the untreated cells (shown in blue). This suggests that the miRNA changes induced by BIM-23a760 are more specific than the changes in mRNA expression. This was also supported by the fact, that a higher proportion of miRNA transcripts than of mRNA are changed by BIM-23a760 treatment.

3.3. Somatostatin Induces Distinct Receptor Based/Activated microRNA Expression Profiles and Particularly Up-Regulates miR-7 and miR-148a

Having demonstrated that SST and dopamine analogues inhibited the growth of the carcinoid cell line NCI-H727 and that activation of these receptors primarily affected miRNA expression, we hypothesized that at least some of their growth inhibitory effects depended on receptor activation and miRNA expression regulation. Also, after identifying the most potent inhibitor of growth among the different SST analogues we expanded the study and examined all the different SST analogues' effect on miRNA expression by miRNA analysis of NCI-H727 cells treated with BIM-23A760, BIM-23014, BIM-23023, BIM-53097 and with a control. A PCA of the miRNA expression profiles showed that, depending on which analogue the cells were treated with, they could be separated into groups

according to the analogue used and receptor(s) activated. In the first dimension the samples treated with BIM-23a760 (SSTR/DRD2) and BIM-23014 (SSTR-only) analogues separated from the rest and in the third dimension there was a difference in miRNA expression depending on receptor type activation (Figure 3A).

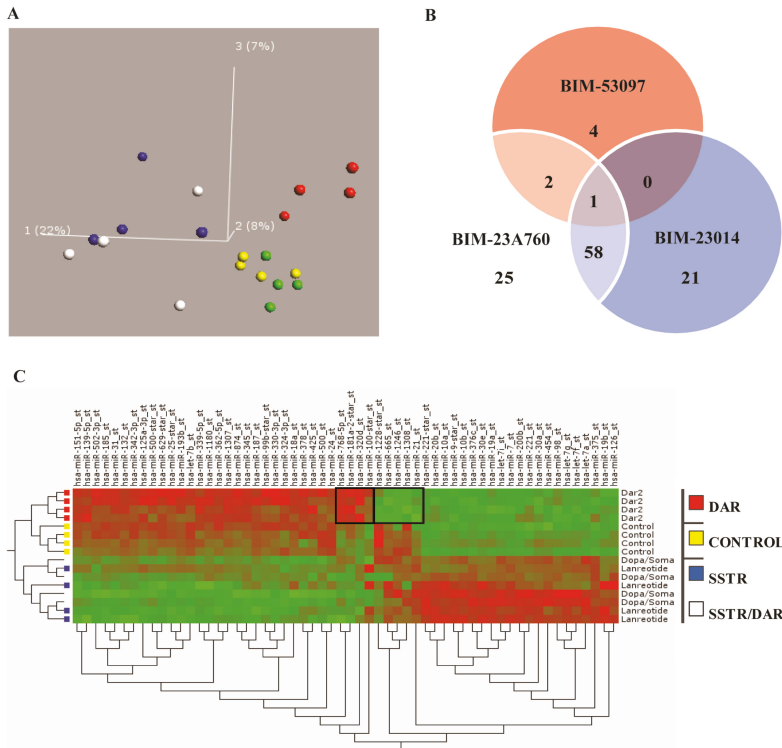


Figure 3. PCA shows specific miRNA expression depending on receptor activation. (A) The unbiased PCA separates the samples based on variation in the expression of miRNA in each sample. In the first dimension samples treated with BIM-23014 (blue—SSTR agonist) and BIM- 23A760 (white—SSTR and dopamine receptor D2 (DRD2) chimeric agonist) separated from the others. BIM-23023 (green—SSTR) did not separate from the control (yellow). Also in the first dimension BIM-53097 (red—DRD2) did not separate from the controls. However, in the third dimension the BIM-53097 (red—DRD2) treated samples clearly separated from the controls. The percentage indicates how much of the total variation is present in each dimension. In this experiment 22% of the total variation of miRNA expression is found in the first dimension, 8% in the second and 7% in the third. The two most potent inhibitors of growth BIM-23A760 and BIM-23014 group together, the DRD2 agonist and the least effective growth inhibitor separates by itself. (B) The Venn diagram shows the miRNAs shared between the different analogue treatments and that the number of miRNAs shared between BIM-23A760 (SSTR-DRD2) and BIM-23014 (SSTR) are greater than BIM-53097 (DRD2). (C) On the right side of the map is the clustering of treatment where DRD2 and controls separate from chimeric compound BIM-23A760 (SSTR-DRD2) and BIM-23014 (SSTR). Furthermore it seems to separate controls and DRD2. Notice in the middle of the control/DRD2 treatment, there is a small fraction of miRNA which seems to be controlled by dopamine alone and different from the control group.

A Venn diagram shows the overlap between miRNA expression shared between the analogues and analogue treatment. The results show that a higher number of miRNAs are affected and changed after

treatments with BIM-23a760 and BIM-23014, which target SSTR either alone or in combination with DRD2, compared to BIM-52097 which only targets DRD2. The highest degree of miRNA expressional changes was induced by SSTR activation compared to selective DRD2 activation (Figure 3B). However, a more specific miRNA expression change was observed in the small fraction of miRNAs changed only in response to BIM-53097 and DRD2 receptor activation. Since BIM-23023 (SSTR) did not separate from the control group in the PCA plot, it was excluded (Figure 3B). Based on the miRNA array results we created a heat map to better visualize the changes in miRNA expression between the treatments with the different compounds, again data with BIM-23023 (SSTR) being excluded (Figure 3C). We also created a list of the most significantly down- or up regulated miRNAs from the two compounds which showed the biggest inhibitory effect on cell proliferation BIM-23A760 (SSTR/DRD2), BIM-23014 (SSTR) and the least effective growth inhibitor BIM-53097 (DRD2) (Table 4). From this list we chose to focus on miR-7 and miR-148a.

Table 4. Up- and down-regulated miRNAs based on Somatostatin analogue (SSA) treatment.

miRNA	Change	BIM-23014		BIM-23A760		BIM-53097	
		FC	p-Value	FC	p-Value	FC	p-Value
miR-769-3p		−2.8	0.001	−2.6	0.002	−1.1	NS
miR-663b		−2.0	0.009	−2.5	0.001	−2.9	0.00
miR-663	↓	−2.0	0.007	−2.3	0.002	−3.3	0.00
miR-30b-star		−2.0	0.001	−2.2	0.000	−1.3	NS
miR-297		−1.7	0.008	−2.0	0.001	−1.7	0.01
miR-483-5p		−1.9	0.003	−2.0	0.002	−1.0	NS
miR-376c		2.1	0.000	2.0	0.000	−1.2	NS
Let-7f		1.8	0.001	2.1	0.000	−1.2	NS
miR-10a-star	↑	1.9	0.001	2.1	0.000	1.1	NS
miR-495		2.0	0.001	2.2	0.000	1.6	0.01
miR-7		2.3	0.000	2.2	0.000	−1.1	NS
miR-9-star		2.3	0.000	2.3	0.000	−1.4	NS
miR-454		1.9	0.001	2.3	0.000	−1.0	NS
miR-26b		2.2	0.010	2.6	0.003	1.0	NS
miR-429		2.9	0.004	2.8	0.004	−1.4	NS
miR-9		1.5	NS	2.9	0.002	−1.5	NS
miR-148a		2.5	0.006	2.9	0.002	2.0	0.03
miR-30e-star		2.8	0.000	3.1	0.000	1.5	NS

The up and down regulated miRNAs in NIH-H727 cells after 24 h treatment with somatostatin analogue BIM-23014, dopamine analogue BIM-53097 and the chimeric somatostatin-dopamine analogue BIM-23a760. Cut-off $p \leq 0.001$, N = 4. NS—p-value not statistically significant; Fold Change—FC.

3.4. In Situ Hybridization and qPCR on Neuroendocrine Tumors and Neuroendocrine Tumor Laser Capture Microdissected Tissue Confirms the Presence of miR-7 and miR-148a

To characterize the localization and to visualize the expression of miR-7 and miR-148a in NETs we performed in situ hybridization on five NETs and found miR-7 expressed exclusively by the endocrine cells, both tumor and normal cells; miR-148a was also expressed primarily by the endocrine cells also both tumor and normal cells but to a lesser extent (Figure 4A). Expression analysis of LCM NET tissue confirmed the up-regulation of particularly miR-7, but also shows the presence of miR-148a (Figure 4B).

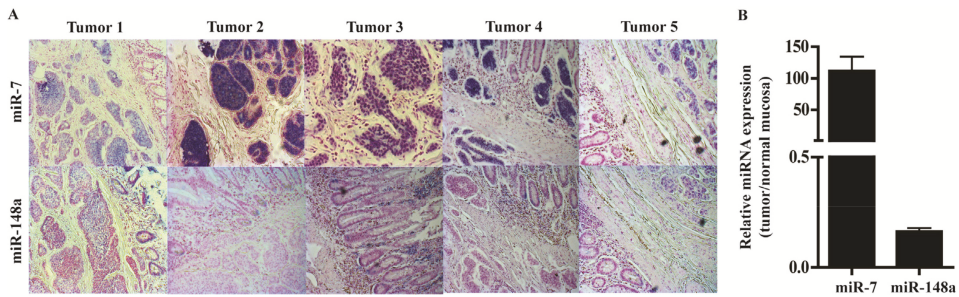


Figure 4. miR-7 and miR-148a are primarily expressed in neuroendocrine cells and inhibit the growth of a carcinoid cell line. (A) In situ hybridization of miR-7 show robust expression of miR-7 specifically located to the neuroendocrine cells. The expression of miR-148a is also predominantly seen in the neuroendocrine cells, although to a lesser extent than miR-7. (B) qPCR of laser capture micro dissected cells confirmed that miR-7 is robustly expressed in NETs compared to normal mucosa. In contrast, the expression of miR-148a is lower in NETs than in the normal mucosa that contains endocrine cells.

3.5. miR-7 and miR-148a Modulate the Growth of NCI-H727 and CNDT2 Carcinoid Cell Lines

Having demonstrated that SST analogues induced the expression of miR-7 and miR-148a and that both these miRNAs are expressed in NETs, we examined how these miRNAs would affect the growth of both the NCI-H727 and CNDT2 carcinoid cell lines by either over expressing or inhibiting them. Both miR-7 and miR-148a significantly inhibit the growth of carcinoid cells in vitro indicating that these miRNA could mediate some of the growth inhibitory effect induced by SST analogues (Figure 5A,B). We subsequently blocked the effect of miR-7 and miR-148a by transfecting LNA-inhibitors that increased the growth of both carcinoid cell lines (Figure 5C,D).

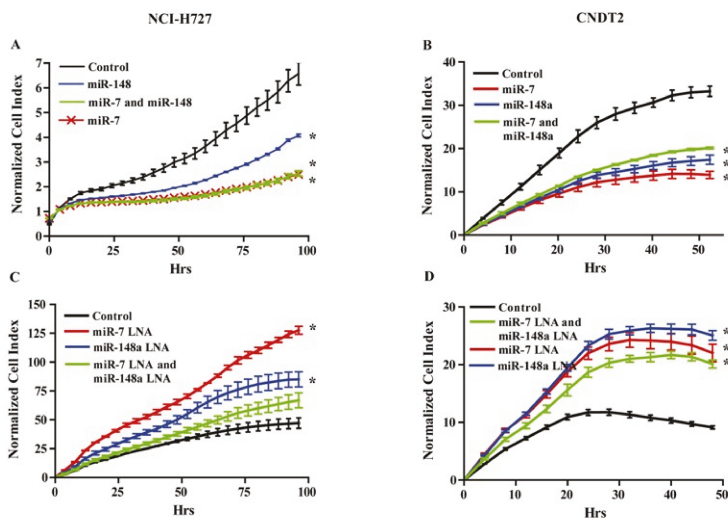


Figure 5. miR-7 and miR-148a regulates growth of the carcinoid cell lines NCI-H727 and CNDT2. (A) and (B) Both miR-7 and miR-148a reduces cellular growth of the carcinoid cell lines NCI-H727 and CNDT2 where the highest inhibitory effect is caused by transfecting with miR-7. (C) and (D) Inhibiting miR-7 and miR-148a by transfecting the carcinoid cell lines NCI-H727 and CNDT2 with miRNA inhibitors alleviate growth repression and cause the cells to grow more than the controls. * growth rates statistically significant compared to control.

3.6. Somatostatin Modulates the Expression of the Let-7 Family

We have previously shown that the expression of several Let-7 family members is reduced during NET carcinogenesis [37] and we therefore specifically examined how the SSTR and DRD2 analogues affected the expression of the let-7 miRNA family. We found that the expression of 4 of the Let-7 members increased after treatment with SSTR analogues, the expression of one was slightly reduced and the expression of 4 was unaffected (Table 5).

Table 5. Differentially regulated let-7 family members based on SSA treatment.

miRNA	Change	BIM-23014		BIM-23A760		BIM-53097	
		FC	<i>p</i> -Value	FC	<i>p</i> -Value	FC	<i>p</i> -Value
let-7a	→	1.2	0.000	1.2	0.000	−1.0	NS
let-7b	↓	−1.4	0.000	−1.4	0.001	−1.0	NS
let-7c	→	−1.0	NS	1.0	NS	−1.1	NS
let-7d	→	−1.0	NS	1.0	NS	1.1	NS
let-7e	→	−1.2	NS	−1.1	NS	1.1	NS
let-7f	↑	1.8	0.002	2.1	0.000	−1.2	NS
let-7g	↑	1.7	0.001	1.9	0.000	−1.1	NS
let-7i	↑	1.3	0.000	1.4	0.000	−1.1	NS
miR-98	↑	1.8	0.004	2.2	0.000	−1.1	NS

The alteration in the expression of the let-7 family of miRNAs in NIH-H727 cells after 24 h treatment with somatostatin analogue BIM-23014, dopamine analogue BIM-53097 and the chimeric somatostatin-dopamine analogue BIM-23a760. N = 4, NS—*p*-value not significant.

Thus, SST analogues had the ability to reintroduce the expression of the let-7 family and possibly revert some of the pathways otherwise involved in NET carcinogenesis and malignancy.

4. Discussion

We have shown that in our NET model system SST analogue treatment primarily induces changes in miRNA expression profiles in carcinoid cell lines. SST analogues are widely used to treat patients with NETs, as the SST analogues both alleviate the carcinoid syndrome and inhibit growth of the tumors [19,38–40]. Here we show that SST analogues inhibit the growth of two carcinoid cell lines and that the inhibitory potential depends on the analogue used. This is in good concordance with other studies that also show that SST analogues inhibit the growth of the NCI-H727 carcinoid cell line [30,41] and the CNDT2 carcinoid cell line [42] and that a chimeric compound targeting SSTRs and DRD2s has the most potent growth inhibitory effect [30,41,42]. The growth inhibitory effects of SST analogues are mediated directly by binding of SST to the SSTRs, which leads to cell cycle arrest or apoptosis and indirectly by the inhibition of growth factors and the suppression of oncogenic signal transducing pathways. However, the molecular mechanisms linking neuroendocrine proliferation and tumor progression are not yet fully understood. The PI3K/MAPK/mTOR pathway is known to play an important role in NETs [43,44]. Already existing therapies used to treat patients with NETs consist of agents targeting this specific signaling pathway [45–49]. SSTR2 has been shown to bind directly to the p85 subunit of PI3K, which belongs to the class I_A PI3Ks and is the class specifically involved in promoting cell survival, growth and proliferation and the most important subclass involved in human cancers [50,51]. SST analogue treatment of a pancreatic cell line can inhibit the binding between SSTR2 and p85 which is critical for the down-regulation of PI3K activity and resulting decreased cell survival [50].

We found miR-7 and miR-148a among the most up-regulated miRNAs after treatment with SST analogues and increased expression of the let-7 family members, all of which have growth inhibitory effects. miR-7 has been shown to be down-regulated in several cancers including colorectal cancer [52] and gastric cancer [53] and to be endocrine specific [54], with evidence going both ways to whether or not miR-7 acts as a tumor suppressor [55]. Here we demonstrated miR-7 to be highly present in NETs.

Both miR-7 and SST targets the PI3K/MAPK/mTOR pathway underlining the significance for the up-regulation of this specific miRNA by SST. In a study in lung cancers, including the carcinoid cell line NCI-H727 showed miR-7 to directly target PI3KR3, the regulatory subunit of PI3K, and to reduce the metastatic potential by reducing the effect of TLR9 signaling [56]. In hepatocellular carcinoma miR-7 has been shown to directly target and repress PI3KCD an integral component of the PI3K signaling pathway, thereby both inhibiting cellular growth, invasion and migration in vitro and more importantly tumorigenesis and metastasis in vivo [57].

Another important aspect of the induction of miR-7 expression in cancer lies in the fact that miR-7, through its inhibitory actions on central cancerous signaling pathways, can increase the sensitivity and improve otherwise chemo- or radiotherapy resistant tumor cells [55]. SST mediates up-regulation of miR-7 thereby increasing the cells sensitivity to the growth inhibitory actions of SST itself resulting in a positive feed-forward loop.

miR-148a has been shown to be down-regulated in early gastric- [58], pancreatic- [59] and colon cancer [60], as well as in breast cancer [61], and, when down-regulated, suggested as a biomarker for the diagnosis and prognosis of gastrointestinal cancers [58,62,63]. Shivapurkar et al. showed that a panel of six miRNAs including miR-148a could predict the risk of recurrence of colon cancer [60]; thus the up-regulation of miR-148a by SST analogue treatment might also be used as a biomarker in NETs. We found miR-148a localized to endocrine cells and tumor tissue, all of which points towards it being a specific and important miRNA in endocrine tumor formation and initiation. Our finding that miR-148a has a growth inhibitory effect on carcinoid cell lines makes it an important miRNA to up-regulate for SST analogues as it will give the SST analogues the ability to “pack that extra punch” in stopping cell growth, and which is possibly why we see that SST analogue treatment primarily gives a change in miRNA expression rather than mRNA expression.

SST itself attenuates Insulin Growth Factor 1 (IGF-1) signaling [64] and IGF-1 Receptor (IGF-1R) is important in GEP-NET tumor growth factor biology [51]. Since miR-148a targets IGF-1, SST analogue induced miR-148a expression could again lead to the interference with the PI3K signaling pathway by the inhibition of growth factors like IGF-1 and its receptor IGF-1R, which frequently are overexpressed in NETs [43,51]. Inhibition of IGF-1R signaling has been shown to decrease PI3K signaling and the induction of cell cycle arrest and apoptosis [51]. miR-148a directly targets and down-regulates IGF-1R in breast cancer and over-expression of miR-148a decreased phosphorylated Akt, a component of the PI3K signaling pathway [61]. miR-7 also targets IGF-1R in a gastric cancer model with significant influence on inhibiting the metastatic potential in this model [53]. However, while SSTR agonists have been shown to have anti-tumor growth activity [30,41,42], IGF-1R inhibition proved to be ineffective in treatment of NETs [65]. One explanation could be that IGF-1R inhibition only targets the receptor activity but not directly the downstream signaling events. In contrast, SSTR activation and induction of the miRNAs could potentially inhibit IGF-1R signaling at multiple levels and hence be more effective. We have previously shown the let-7 family to be among the most down-regulated miRNA in NETs. We also identified and that it targets HMGA2, BACH1 and MMP1 and reduce the expression of these oncogenes all present in NETs [37]. Here we find that treatment with SST analogues modulates the expression of the let-7 family and leads to the up-regulation of several family members and the down-regulation of only one.

Recently, several reports have established the let-7 miRNAs as key players in metabolic pathways, especially in the glucose metabolic pathway through the inhibition of IGF-1R, which is a key target in the PI3K/mTOR signaling pathway [66]. Others have shown the let-7 family to directly target and inhibit IGF-1 and IGF-1R, which harbors three let-7 binding sites in its 3'UTR. This leads to the elimination of PI3K activation, and hence abrogates the signaling pathway leading towards cell division, differentiation and survival. placing let-7 up-stream of IGF-1/IGF-1R with downstream effects on the PI3K signaling cascade [67,68]. The positive modulation of the let-7 family by SST analogues targets and inhibits the IGF-1/PI3K signaling pathway in addition to maybe inhibiting glucose nutrient in aiding the rapid growth of the cancer cells. By this mechanism, they are targeted

on both survival signaling as well as their supply of nutrients. In conclusion miR-7, miR-148a and the let-7 family most up-regulated by SST have been shown to play a role in inhibiting the PI3K signaling pathway at different levels reducing the cancer cell's ability to escape and circumvent inhibition of a single step. The fact that the miRNAs up-regulated by the SST analogues, the analogues themselves and the already existing therapies all seem to target the same signaling pathway underlines the importance of targeting this pathway when treating NETs.

Author Contributions: K.B.V.D. and L.F.-H. conceived and designed the experiments; K.B.V.D.; C.K. and T.B. performed the experiments; K.B.V.D.; L.F.-H.; C.K., J.V. and T.B. interpreted and analyzed the data; all authors contributed reagents and materials; K.B.V.D. and L.F.-H. wrote the paper; obtained funding (K.B.V.D., B.F., A.K., L.F.-H.); study supervision (L.F.-H.). All authors critically revised the manuscript for important intellectual content and approved the paper.

Funding: This study was supported by The Lundbeck Foundation (L.F.-H., K.B.V.D. R31-A2394), The Desiree and Niels Yde Foundation (L.F.-H., K.B.V.D.), The Danish National Research Foundation (L.F.-H.), The Novo Nordisk Foundation (L.F.-H.), The National Advanced Technology Foundation (A.K.; 005-2007-2), The Svend Andersen Foundation (A.K.; 2008), Rigshospitalets Research Council (A.K.; 604-18), and The Capital Region of Denmark (A.K.; R101-A1945).

Acknowledgments: Mette Moldaschl is thanked for excellent technical assistance. We also thank Lee M. Ellis, Department of Surgical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77230 for the CDNT2 cells.

Conflicts of Interest: M.D. Culler is an employee of Ipsen-Biomeasure. No other potential conflict of interest that could be perceived as prejudicing the impartiality of the research is reported.

References

1. Modlin, I.M.; Oberg, K.; Chung, D.C.; Jensen, R.T.; de Herder, W.W.; Thakker, R.V.; Caplin, M.; Delle Fave, G.; Kaltsas, G.A.; Krenning, E.P.; et al. Gastroenteropancreatic neuroendocrine tumours. *Lancet Oncol.* **2008**, *9*, 61–72. [[CrossRef](#)]
2. Kloppel, G. Classification and pathology of gastroenteropancreatic neuroendocrine neoplasms. *Endocr. Relat. Cancer* **2011**, *18* (Suppl. 1), S1–S16. [[CrossRef](#)] [[PubMed](#)]
3. Janson, E.T.; Sorbye, H.; Welin, S.; Federspiel, B.; Gronbaek, H.; Hellman, P.; Ladekarl, M.; Langer, S.W.; Mortensen, J.; Schalim-Jantti, C.; et al. Nordic guidelines 2014 for diagnosis and treatment of gastroenteropancreatic neuroendocrine neoplasms. *Acta Oncol.* **2014**, *53*, 1284–1297. [[CrossRef](#)] [[PubMed](#)]
4. Travis, W.D. The 2015 World Health Organization classification of lung tumors. *J. Thorac. Oncol.* **2015**, *10*, 1243–1260. [[CrossRef](#)] [[PubMed](#)]
5. Ramirez, R.A.; Beyer, D.T.; Diebold, A.E.; Voros, B.A.; Chester, M.M.; Wang, Y.Z.; Boudreaux, J.P.; Woltering, E.A.; Uhlhorn, A.P.; Ryan, P.; et al. Prognostic factors in typical and atypical pulmonary carcinoids. *Ochsner. J.* **2017**, *17*, 335–340. [[PubMed](#)]
6. Lepage, C.; Rachet, B.; Coleman, M.P. Survival from malignant digestive endocrine tumors in England and Wales: A population-based study. *Gastroenterology* **2007**, *132*, 899–904. [[CrossRef](#)] [[PubMed](#)]
7. Niederle, M.B.; Hackl, M.; Kaserer, K.; Niederle, B. Gastroenteropancreatic neuroendocrine tumours: The current incidence and staging based on the Who and European neuroendocrine tumour society classification: An analysis based on prospectively collected parameters. *Endocr. Relat. Cancer* **2010**, *17*, 909–918. [[CrossRef](#)] [[PubMed](#)]
8. Caplin, M.E.; Baudin, E.; Ferolla, P.; Filosso, P.; Garcia-Yuste, M.; Lim, E.; Oberg, K.; Pelosi, G.; Perren, A.; Rossi, R.E.; et al. Pulmonary neuroendocrine (carcinoid) tumors: European neuroendocrine tumor society expert consensus and recommendations for best practice for typical and atypical pulmonary carcinoids. *Ann. Oncol.* **2015**, *26*, 1604–1620. [[CrossRef](#)] [[PubMed](#)]
9. Hashmi, H.; Vanberkel, V.; Bade, B.C.; Kloecker, G. Clinical presentation, diagnosis, and management of typical and atypical bronchopulmonary carcinoid. *J. Community Support* **2017**, *15*, E303–E308. [[CrossRef](#)]
10. Eriksson, B.; Arnberg, H.; Lindgren, P.G.; Lorelius, L.E.; Magnusson, A.; Lundqvist, G.; Skogseid, B.; Wide, L.; Wilander, E.; Oberg, K. Neuroendocrine pancreatic tumours: Clinical presentation, biochemical and histopathological findings in 84 patients. *J. Intern. Med.* **1990**, *228*, 103–113. [[CrossRef](#)] [[PubMed](#)]

11. Hendifar, A.E.; Marchevsky, A.M.; Tuli, R. Neuroendocrine tumors of the lung: Current challenges and advances in the diagnosis and management of well-differentiated disease. *J. Thorac. Oncol.* **2017**, *12*, 425–436. [[CrossRef](#)] [[PubMed](#)]
12. Ramage, J.K.; Davies, A.H.; Ardill, J.; Bax, N.; Caplin, M.; Grossman, A.; Hawkins, R.; McNicol, A.M.; Reed, N.; Sutton, R.; et al. Guidelines for the management of gastroenteropancreatic neuroendocrine (including carcinoid) tumours. *Gut* **2005**, *54* (Suppl. 4), iv1–iv16. [[CrossRef](#)] [[PubMed](#)]
13. Plockinger, U.; Wiedenmann, B. Treatment of gastroenteropancreatic neuroendocrine tumors. *Virchows Arch.* **2007**, *451* (Suppl. 1), S71–S80. [[CrossRef](#)] [[PubMed](#)]
14. Cameselle-Teijeiro, J.M.; Mato Mato, J.A.; Fernandez Calvo, O.; Garcia Mata, J. Neuroendocrine pulmonary tumors of low, intermediate and high grade: Anatomopathological diagnosis-prognostic and predictive factors. *Mol. Diagn. Ther.* **2018**, *22*, 169–177. [[CrossRef](#)] [[PubMed](#)]
15. Reichlin, S. Secretion of somatostatin and its physiological function. *J. Lab. Clin. Med.* **1987**, *109*, 320–326. [[PubMed](#)]
16. Larsson, L.I.; Goltermann, N.; de Magistris, L.; Rehfeld, J.F.; Schwartz, T.W. Somatostatin cell processes as pathways for paracrine secretion. *Science* **1979**, *205*, 1393–1395. [[CrossRef](#)] [[PubMed](#)]
17. Yamada, Y.; Post, S.R.; Wang, K.; Tager, H.S.; Bell, G.I.; Seino, S. Cloning and functional characterization of a family of human and mouse somatostatin receptors expressed in brain, gastrointestinal tract, and kidney. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 251–255. [[CrossRef](#)] [[PubMed](#)]
18. Papotti, M.; Bongiovanni, M.; Volante, M.; Allia, E.; Landolfi, S.; Helboe, L.; Schindler, M.; Cole, S.L.; Bussolati, G. Expression of somatostatin receptor types 1–5 in 81 cases of gastrointestinal and pancreatic endocrine tumors. A correlative immunohistochemical and reverse-transcriptase polymerase chain reaction analysis. *Virchows Arch.* **2002**, *440*, 461–475. [[CrossRef](#)] [[PubMed](#)]
19. Modlin, I.M.; Pavel, M.; Kidd, M.; Gustafsson, B.I. Review article: Somatostatin analogues in the treatment of gastroenteropancreatic neuroendocrine (carcinoid) tumours. *Aliment. Pharmacol. Ther.* **2010**, *31*, 169–188. [[PubMed](#)]
20. Binderup, T.; Knigge, U.; Loft, A.; Mortensen, J.; Pfeifer, A.; Federspiel, B.; Hansen, C.P.; Hojgaard, L.; Kjaer, A. Functional imaging of neuroendocrine tumors: A head-to-head comparison of somatostatin receptor scintigraphy, 123i-MIBG scintigraphy, and 18F-FDG PET. *J. Nucl. Med.* **2010**, *51*, 704–712. [[CrossRef](#)] [[PubMed](#)]
21. Buscail, L.; Saint-Laurent, N.; Chastre, E.; Vaillant, J.C.; Gespach, C.; Capella, G.; Kalthoff, H.; Lluís, F.; Vaysse, N.; Susini, C. Loss of sst2 somatostatin receptor gene expression in human pancreatic and colorectal cancer. *Cancer Res.* **1996**, *56*, 1823–1827. [[PubMed](#)]
22. Sheppard, M.; Shapiro, B.; Pimstone, B.; Kronheim, S.; Berelowitz, M.; Gregory, M. Metabolic clearance and plasma half-disappearance time of exogenous somatostatin in man. *J. Clin. Endocrinol. Metab.* **1979**, *48*, 50–53. [[CrossRef](#)] [[PubMed](#)]
23. Appetecchia, M.; Baldelli, R. Somatostatin analogues in the treatment of gastroenteropancreatic neuroendocrine tumours, current aspects and new perspectives. *J. Exp. Clin. Cancer Res.* **2010**, *29*, 19. [[CrossRef](#)] [[PubMed](#)]
24. Cives, M.; Kunz, P.L.; Morse, B.; Coppola, D.; Schell, M.J.; Campos, T.; Nguyen, P.T.; Nandoskar, P.; Khandelwal, V.; Strosberg, J.R. Phase II clinical trial of pasireotide long-acting repeatable in patients with metastatic neuroendocrine tumors. *Endocr. Relat. Cancer* **2015**, *22*, 1–9. [[CrossRef](#)] [[PubMed](#)]
25. Alexandraki, K.I.; Karapanagioti, A.; Karoumpalis, I.; Boutzios, G.; Kaltsas, G.A. Advances and current concepts in the medical management of gastroenteropancreatic neuroendocrine neoplasms. *Biomed Res. Int.* **2017**, *2017*, 9856140. [[CrossRef](#)] [[PubMed](#)]
26. 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. [[CrossRef](#)] [[PubMed](#)]
27. Ha, T.Y. MicroRNAs in human diseases: From cancer to cardiovascular disease. *Immune Netw.* **2011**, *11*, 135–154. [[CrossRef](#)] [[PubMed](#)]
28. Van Buren, G., 2nd; Rashid, A.; Yang, A.D.; Abdalla, E.K.; Gray, M.J.; Liu, W.; Somcio, R.; Fan, F.; Camp, E.R.; Yao, J.C.; et al. The development and characterization of a human midgut carcinoid cell line. *Clin. Cancer Res.* **2007**, *13*, 4704–4712. [[CrossRef](#)] [[PubMed](#)]

29. Stilling, G.A.; Zhang, H.; Ruebel, K.H.; Leontovich, A.A.; Jin, L.; Tanizaki, Y.; Zhang, S.; Erickson, L.A.; Hobday, T.; Lloyd, R.V. Characterization of the functional and growth properties of cell lines established from ileal and rectal carcinoid tumors. *Endocr. Pathol.* **2007**, *18*, 223–232. [[CrossRef](#)] [[PubMed](#)]
30. Kidd, M.; Drozdov, I.; Joseph, R.; Pfragner, R.; Culler, M.; Modlin, I. Differential cytotoxicity of novel somatostatin and dopamine chimeric compounds on bronchopulmonary and small intestinal neuroendocrine tumor cell lines. *Cancer* **2008**, *113*, 690–700. [[CrossRef](#)] [[PubMed](#)]
31. Arvigo, M.; Gatto, F.; Ruscica, M.; Ameri, P.; Dozio, E.; Albertelli, M.; Culler, M.D.; Motta, M.; Minuto, F.; Magni, P.; et al. Somatostatin and dopamine receptor interaction in prostate and lung cancer cell lines. *J. Endocrinol.* **2010**, *207*, 309–317. [[CrossRef](#)] [[PubMed](#)]
32. Kustermann, S.; Boess, F.; Bunes, A.; Schmitz, M.; Watzele, M.; Weiser, T.; Singer, T.; Suter, L.; Roth, A. A label-free, impedance-based real time assay to identify drug-induced toxicities and differentiate cytostatic from cytotoxic effects. *Toxicol. In Vitro* **2013**, *27*, 1589–1595. [[CrossRef](#)] [[PubMed](#)]
33. Ke, N.; Wang, X.; Xu, X.; Abassi, Y.A. The xCELLigence system for real-time and label-free monitoring of cell viability. *Methods Mol. Biol.* **2011**, *740*, 33–43. [[PubMed](#)]
34. Binderup, T.; Knigge, U.; Mellon Mogensen, A.; Palnaes Hansen, C.; Kjaer, A. Quantitative gene expression of somatostatin receptors and noradrenaline transporter underlying scintigraphic results in patients with neuroendocrine tumors. *Neuroendocrinology* **2008**, *87*, 223–232. [[CrossRef](#)] [[PubMed](#)]
35. Peltier, H.J.; Latham, G.J. Normalization of microRNA expression levels in quantitative RT-PCR assays: Identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA* **2008**, *14*, 844–852. [[CrossRef](#)] [[PubMed](#)]
36. Jorgensen, S.; Baker, A.; Moller, S.; Nielsen, B.S. Robust one-day in situ hybridization protocol for detection of microRNAs in paraffin samples using LNA probes. *Methods* **2010**, *52*, 375–381. [[CrossRef](#)] [[PubMed](#)]
37. Dossing, K.B.; Binderup, T.; Kaczkowski, B.; Jacobsen, A.; Rossing, M.; Winther, O.; Federspiel, B.; Knigge, U.; Kjaer, A.; Friis-Hansen, L. Down-regulation of miR-129-5p and the let-7 family in neuroendocrine tumors and metastases leads to up-regulation of their targets Egr1, G3bp1, Hmga2 and Bach1. *Genes* **2014**, *6*, 1–21. [[CrossRef](#)] [[PubMed](#)]
38. Strosberg, J.R.; Fisher, G.A.; Benson, A.B.; Malin, J.L.; Panel, G.T.C.; Cherepanov, D.; Broder, M.S.; Anthony, L.B.; Arslan, B.; Fisher, G.A.; et al. Systemic treatment in unresectable metastatic well-differentiated carcinoid tumors: Consensus results from a modified delphi process. *Pancreas* **2013**, *42*, 397–404. [[CrossRef](#)] [[PubMed](#)]
39. Rinke, A.; Muller, H.H.; Schade-Brittinger, C.; Klose, K.J.; Barth, P.; Wied, M.; Mayer, C.; Aminossadati, B.; Pape, U.F.; Blaker, M.; et al. Placebo-controlled, double-blind, prospective, randomized study on the effect of octreotide lar in the control of tumor growth in patients with metastatic neuroendocrine midgut tumors: A report from the PROMID study group. *J. Clin. Oncol.* **2009**, *27*, 4656–4663. [[CrossRef](#)] [[PubMed](#)]
40. Caplin, M.E.; Pavel, M.; Ruzzniewski, P. Lanreotide in metastatic enteropancreatic neuroendocrine tumors. *N. Engl. J. Med.* **2014**, *371*, 1556–1557. [[CrossRef](#)] [[PubMed](#)]
41. Kidd, M.; Schally, A.V.; Pfragner, R.; Malfertheiner, M.V.; Modlin, I.M. Inhibition of proliferation of small intestinal and bronchopulmonary neuroendocrine cell lines by using peptide analogs targeting receptors. *Cancer* **2008**, *112*, 1404–1414. [[CrossRef](#)] [[PubMed](#)]
42. Li, S.C.; Martijn, C.; Cui, T.; Essaghir, A.; Luque, R.M.; Demoulin, J.B.; Castano, J.P.; Oberg, K.; Giandomenico, V. The somatostatin analogue octreotide inhibits growth of small intestine neuroendocrine tumour cells. *PLoS ONE* **2012**, *7*, e48411. [[CrossRef](#)] [[PubMed](#)]
43. Kharmate, G.; Rajput, P.S.; Lin, Y.C.; Kumar, U. Inhibition of tumor promoting signals by activation of SSTR2 and opioid receptors in human breast cancer cells. *Cancer Cell Int.* **2013**, *13*, 93. [[CrossRef](#)] [[PubMed](#)]
44. Fernandes, I.; Pacheco, T.R.; Costa, A.; Santos, A.C.; Fernandes, A.R.; Santos, M.; Oliveira, A.G.; Casimiro, S.; Quintela, A.; Fernandes, A.; et al. Prognostic significance of AKT/mTOR signaling in advanced neuroendocrine tumors treated with somatostatin analogs. *Onco Targets Ther.* **2012**, *5*, 409–416. [[CrossRef](#)] [[PubMed](#)]
45. Porta, C.; Paglino, C.; Mosca, A. Targeting PI3K/AKT/mTOR signaling in cancer. *Front. Oncol.* **2014**, *4*, 64. [[CrossRef](#)] [[PubMed](#)]
46. Johnbeck, C.B.; Munk Jensen, M.; Haagen Nielsen, C.; Fisker Hag, A.M.; Knigge, U.; Kjaer, A. ¹⁸F-FDG and ¹⁸F-FLT-pet imaging for monitoring everolimus effect on tumor-growth in neuroendocrine tumors: Studies in human tumor xenografts in mice. *PLoS ONE* **2014**, *9*, e91387.

47. Raymond, E.; Dahan, L.; Raoul, J.L.; Bang, Y.J.; Borbath, I.; Lombard-Bohas, C.; Valle, J.; Metrakos, P.; Smith, D.; Vinik, A.; et al. Sunitinib malate for the treatment of pancreatic neuroendocrine tumors. *N. Engl. J. Med.* **2011**, *364*, 501–513. [[CrossRef](#)] [[PubMed](#)]
48. Yao, J.C.; Shah, M.H.; Ito, T.; Bohas, C.L.; Wolin, E.M.; Van Cutsem, E.; Hobday, T.J.; Okusaka, T.; Capdevila, J.; de Vries, E.G.; et al. Everolimus for advanced pancreatic neuroendocrine tumors. *N. Engl. J. Med.* **2011**, *364*, 514–523. [[CrossRef](#)] [[PubMed](#)]
49. Pavel, M.E.; Hainsworth, J.D.; Baudin, E.; Peeters, M.; Horsch, D.; Winkler, R.E.; Klimovsky, J.; Lebwahl, D.; Jehl, V.; Wolin, E.M.; et al. Everolimus plus octreotide long-acting repeatable for the treatment of advanced neuroendocrine tumours associated with carcinoid syndrome (RADIANT-2): A randomised, placebo-controlled, phase 3 study. *Lancet* **2011**, *378*, 2005–2012. [[CrossRef](#)]
50. Bousquet, C.; Guillermet-Guibert, J.; Saint-Laurent, N.; Archer-Lahlou, E.; Lopez, F.; Fanjul, M.; Ferrand, A.; Fourmy, D.; Pichereaux, C.; Monsarrat, B.; et al. Direct binding of p85 to sst2 somatostatin receptor reveals a novel mechanism for inhibiting PI3K pathway. *EMBO J.* **2006**, *25*, 3943–3954. [[CrossRef](#)] [[PubMed](#)]
51. Briest, F.; Grabowski, P. PI3K-AKT-mTOR-signaling and beyond: The complex network in gastroenteropancreatic neuroendocrine neoplasms. *Theranostics* **2014**, *4*, 336–365. [[CrossRef](#)] [[PubMed](#)]
52. Zhang, N.; Li, X.; Wu, C.W.; Dong, Y.; Cai, M.; Mok, M.T.; Wang, H.; Chen, J.; Ng, S.S.; Chen, M.; et al. MicroRNA-7 is a novel inhibitor of YY1 contributing to colorectal tumorigenesis. *Oncogene* **2013**, *32*, 5078–5088. [[CrossRef](#)] [[PubMed](#)]
53. Zhao, X.; Dou, W.; He, L.; Liang, S.; Tie, J.; Liu, C.; Li, T.; Lu, Y.; Mo, P.; Shi, Y.; et al. MicroRNA-7 functions as an anti-metastatic microRNA in gastric cancer by targeting insulin-like growth factor-1 receptor. *Oncogene* **2013**, *32*, 1363–1372. [[CrossRef](#)] [[PubMed](#)]
54. Kredo-Russo, S.; Mandelbaum, A.D.; Ness, A.; Alon, I.; Lennox, K.A.; Behlke, M.A.; Hornstein, E. Pancreas-enriched miRNA refines endocrine cell differentiation. *Development* **2012**, *139*, 3021–3031. [[CrossRef](#)] [[PubMed](#)]
55. Gu, D.N.; Huang, Q.; Tian, L. The molecular mechanisms and therapeutic potential of microRNA-7 in cancer. *Expert Opin. Ther. Targets* **2015**, *19*, 415–426. [[CrossRef](#)] [[PubMed](#)]
56. Xu, L.; Wen, Z.; Zhou, Y.; Liu, Z.; Li, Q.; Fei, G.; Luo, J.; Ren, T. MicroRNA-7-regulated TLR9 signaling-enhanced growth and metastatic potential of human lung cancer cells by altering the phosphoinositide-3-kinase, regulatory subunit 3/Akt pathway. *Mol. Biol. Cell* **2013**, *24*, 42–55. [[CrossRef](#)] [[PubMed](#)]
57. Fang, Y.; Xue, J.L.; Shen, Q.; Chen, J.; Tian, L. MicroRNA-7 inhibits tumor growth and metastasis by targeting the phosphoinositide 3-kinase/Akt pathway in hepatocellular carcinoma. *Hepatology* **2012**, *55*, 1852–1862. [[CrossRef](#)] [[PubMed](#)]
58. Zheng, G.; Xiong, Y.; Xu, W.; Wang, Y.; Chen, F.; Wang, Z.; Yan, Z. A two-microRNA signature as a potential biomarker for early gastric cancer. *Oncol. Lett.* **2014**, *7*, 679–684. [[CrossRef](#)] [[PubMed](#)]
59. Hanoun, N.; Delpu, Y.; Suriawinata, A.A.; Bournet, B.; Bureau, C.; Selves, J.; Tsongalis, G.J.; Dufresne, M.; Buscail, L.; Cordelier, P.; et al. The silencing of microRNA 148a production by DNA hypermethylation is an early event in pancreatic carcinogenesis. *Clin. Chem.* **2010**, *56*, 1107–1118. [[CrossRef](#)] [[PubMed](#)]
60. Shivapurkar, N.; Weiner, L.M.; Marshall, J.L.; Madhavan, S.; Deslattes Mays, A.; Juhl, H.; Wellstein, A. Recurrence of early stage colon cancer predicted by expression pattern of circulating microRNAs. *PLoS ONE* **2014**, *9*, e84686. [[CrossRef](#)]
61. Xu, Q.; Jiang, Y.; Yin, Y.; Li, Q.; He, J.; Jing, Y.; Qi, Y.T.; Xu, Q.; Li, W.; Lu, B.; et al. A regulatory circuit of miR-148a/152 and DNMT1 in modulating cell transformation and tumor angiogenesis through IGF-IR and IRS1. *J. Mol. Cell Biol.* **2013**, *5*, 3–13. [[CrossRef](#)] [[PubMed](#)]
62. Sakamoto, N.; Naito, Y.; Oue, N.; Sentani, K.; Uraoka, N.; Zarni Oo, H.; Yanagihara, K.; Aoyagi, K.; Sasaki, H.; Yasui, W. MicroRNA-148a is downregulated in gastric cancer, targets MMP7, and indicates tumor invasiveness and poor prognosis. *Cancer Sci.* **2014**, *105*, 236–243. [[CrossRef](#)] [[PubMed](#)]
63. Sun, J.; Song, Y.; Wang, Z.; Wang, G.; Gao, P.; Chen, X.; Gao, Z.; Xu, H. Clinical significance of promoter region hypermethylation of microRNA-148a in gastrointestinal cancers. *Onco Targets Ther.* **2014**, *7*, 853–863. [[PubMed](#)]
64. Murray, R.D.; Kim, K.; Ren, S.G.; Chelly, M.; Umehara, Y.; Melmed, S. Central and peripheral actions of somatostatin on the growth hormone-IGF-I axis. *J. Clin. Investig.* **2004**, *114*, 349–356. [[CrossRef](#)] [[PubMed](#)]

65. Libutti, S.K. Therapy: Blockade of IGF-1R-not effective in neuroendocrine tumours. *Nat. Rev. Endocrinol.* **2013**, *9*, 389–390. [[CrossRef](#)] [[PubMed](#)]
66. Thornton, J.E.; Gregory, R.I. How does Lin28 let-7 control development and disease? *Trends Cell Biol.* **2012**, *22*, 474–482. [[CrossRef](#)] [[PubMed](#)]
67. Shen, G.; Wu, R.; Liu, B.; Dong, W.; Tu, Z.; Yang, J.; Xu, Z.; Pan, T. Upstream and downstream mechanisms for the promoting effects of IGF-1 on differentiation of spermatogonia to primary spermatocytes. *Life Sci.* **2014**, *101*, 49–55. [[CrossRef](#)] [[PubMed](#)]
68. Alajez, N.M.; Shi, W.; Wong, D.; Lenarduzzi, M.; Waldron, J.; Weinreb, I.; Liu, F.F. Lin28b promotes head and neck cancer progression via modulation of the insulin-like growth factor survival pathway. *Oncotarget* **2012**, *3*, 1641–1652. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

Complemented Palindromic Small RNAs First Discovered from SARS Coronavirus

Chang Liu ^{1,†}, Ze Chen ^{2,3,†}, Yue Hu ¹, Haishuo Ji ^{4,5}, Deshui Yu ⁴, Wenyuan Shen ¹, Siyu Li ⁴, Jishou Ruan ⁶, Wenjun Bu ^{4,*} and Shan Gao ^{4,5,*}

¹ Laboratory of Medical Molecular Virology, School of Medicine, Nankai University, Tianjin 300071, China; changliu@nankai.edu.cn (C.L.); huyue2016@mail.nankai.edu.cn (Y.H.); shenwy@mail.nankai.edu.cn (W.S.)

² State Key Laboratory of Veterinary Etiological Biology and Key Laboratory of Veterinary Parasitology of Gansu Province, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Science, Lanzhou 730046, China; chenze@caas.cn

³ Co-Innovation Center for Prevention and Control of Important Animal Infectious Diseases and Zoonoses, Yangzhou 225009, China

⁴ College of Life Sciences, Nankai University, Tianjin 300071, China; haishuo_ji@mail.nankai.edu.cn (H.J.); deshui_yu@mail.nankai.edu.cn (D.Y.); siyu_li@mail.nankai.edu.cn (S.L.)

⁵ Institute of Statistics, Nankai University, Tianjin 300071, China

⁶ School of Mathematical Sciences, Nankai University, Tianjin 300071, China; jsruan@nankai.edu.cn

* Correspondence: wenjunbu@nankai.edu.cn (W.B.); gao_shan@mail.nankai.edu.cn (S.G.)

† These authors contributed equally to this work.

Received: 4 July 2018; Accepted: 22 August 2018; Published: 5 September 2018

Abstract: In this study, we report for the first time the existence of complemented palindromic small RNAs (cpsRNAs) and propose that cpsRNAs and palindromic small RNAs (psRNAs) constitute a novel class of small RNAs. The first discovered 19-nt cpsRNA UUAACAAGCUUGUUAAGA, named SARS-CoV-cpsR-19, was detected from a 22-bp DNA complemented palindrome TCTTTACAAGCTTGTTAAAGA in the severe acute respiratory syndrome coronavirus (SARS-CoV) genome. The phylogenetic analysis supported that this DNA complemented palindrome originated from bat betacoronavirus. The results of RNA interference (RNAi) experiments showed that one 19-nt segment corresponding to SARS-CoV-cpsR-19 significantly induced cell apoptosis. Using this joint analysis of the molecular function and phylogeny, our results suggested that SARS-CoV-cpsR-19 could play a role in SARS-CoV infection or pathogenesis. The discovery of cpsRNAs has paved a way to find novel markers for pathogen detection and to reveal the mechanisms underlying infection or pathogenesis from a different point of view. Researchers can use cpsRNAs to study the infection or pathogenesis of pathogenic viruses when these viruses are not available. The discovery of psRNAs and cpsRNAs, as a novel class of small RNAs, also inspire researchers to investigate DNA palindromes and DNA complemented palindromes with lengths of psRNAs and cpsRNAs in viral genomes.

Keywords: palindromic small RNA; complemented palindromic small RNA; small RNA; DNA complemented palindrome; severe acute respiratory syndrome coronavirus

1. Introduction

Small RNA sequencing (small RNA-seq or sRNA-seq) is used to obtain thousands of short RNA sequences with lengths that are usually less than 50 bp. With sRNA-seq, many novel non-coding RNAs (ncRNAs) have been discovered. For example, two featured series of ribosomal RNA (rRNA)-derived RNA fragments (rRFs) constitute a novel class of small RNAs [1]. Small RNA-seq has also been used for virus detection in plants [2–4] and invertebrates [5]. In 2016, Wang et al. first used sRNA-seq data

from the National Center for Biology Information Sequence Read Archive database (NCBI SRA) to show that sRNA-seq can be used to detect and identify human viruses [6], however the detection results were not as robust as those of plant or invertebrate viruses. To improve virus detection in mammals, our strategy was to detect and compare featured RNA fragments in plants, invertebrates, and mammals using sRNA-seq data. In a previous study [7], we detected siRNA duplexes that were induced by plant viruses and analyzed these small interfering RNA (siRNA) duplexes as an important class of featured RNA fragments. In this study, we aimed to investigate siRNA duplexes that were induced by invertebrate and mammalian viruses and unexpectedly discovered another important class of featured RNA fragments: complemented palindromic small RNAs (cpsRNAs). Among all of the detected cpsRNAs, the first discovered cpsRNA, named SARS-CoV-cpsR-19, from the severe acute respiratory syndrome coronavirus (SARS-CoV) strain MA15 merited further study, because mice infected with SARS-CoV MA15 died from an overwhelming viral infection with viral-mediated destruction of pneumocytes and ciliated epithelial cells [8]. SARS-CoV-cpsR-19 was detected from a DNA complemented palindrome TCTTTAACAAGCTTGTTAAAGA in the SARS-CoV genome. In our previous study of mitochondrial genomes, we reported palindromic small RNAs (psRNAs) for the first time [9]. Both of psRNAs and cpsRNAs could comprise a novel class of small RNAs due to their special sequence structures.

In this study, we compared the features of the siRNA duplexes that were induced by mammalian viruses with those that were induced by plant and invertebrate viruses. We found that the detected siRNA duplexes that were induced by mammalian viruses had significantly lower percentages of total sequenced reads than those that were induced by plant and invertebrate viruses, and it seemed that they were produced only from a few sites in the viral genomes. One possible reason could be that a large proportion of the sRNA-seq data is from other small RNA fragments, owing to the presence of a number of double-stranded RNA (dsRNA)-triggered nonspecific responses, such as type I interferon (IFN) synthesis [10]. Another possible reason could be that the missing siRNA duplexes or siRNA fragments function in cells by interacting with host RNAs or proteins. Based on this idea, we hypothesized that SARS-CoV-cpsR-19 or other cpsRNAs from SARS-CoV MA15 could play a role in SARS-CoV infection or pathogenesis.

To test our hypothesis, we conducted the joint analysis of the molecular function and phylogeny. First, we investigated the origins of SARS-CoV-cpsR-19 by the phylogenetic analysis of coronavirus genome sequences that were associated with bats, palm civets, rats, mice, monkeys, dogs, bovines, hedgehogs, giraffes, waterbucks, and equines. Subsequently, we performed RNA interference (RNAi) experiments to test the possible cellular effects that were induced by SARS-CoV-cpsR-19. The phylogenetic analysis supported that the DNA complemented palindrome TCTTTAACAAGCTTGTTAAAGA could originate from bat betacoronaviruses. The results of the RNAi experiments showed that one 19-nt segment corresponding to SARS-CoV-cpsR-19 significantly induced cell apoptosis. This study aimed to provide a different point of view for pathogen detection and pathogenesis studies.

2. Materials and Methods

2.1. Datasets and Data Analysis

All sRNA-seq data were downloaded from the NCBI SRA database. In our previous study, 6 mammalian viruses (HPV-18, HBV, HCV, HIV-1, SMRV, and EBV) were detected from 36 runs of sRNA-seq data [6]. In this study, 11 invertebrate viruses were detected from 51 runs of sRNA-seq data (Supplementary file 1) and 2 mammalian viruses (H1N1 and SARS-CoV) were detected from 20 runs of sRNA-seq data (NCBI SRA: SRP012018) [11]. In total, 11 invertebrate viruses and 8 mammalian viruses were detected from 107 runs of sRNA-seq data using VirusDetect [4], and their genome sequences were downloaded from the NCBI GenBank database. Among 107 runs of sRNA-seq data, four runs (NCBI SRA: SRR452404, SRR452406, SRR452408, and SRR452410) had been sequenced from lung

tissue in mice that were infected with SARS-CoV MA15 [11] and they were used to detect SARS-CoV. The cleaning and quality control of the sRNA-seq data were performed using pipeline Fastq_clean [12], which was optimized to clean the raw reads from the Illumina platforms. Using the software Bowtie v0.12.7 [13] with one mismatch, we aligned all of the cleaned sRNA-seq reads to viral genome sequences and obtained alignment results in sequence alignment map (SAM) format for the detection of siRNA duplexes using the program duplexfinder [7]. Statistical computation and plotting were performed using the software R v2.15.3 (R Core Team, Vienna, Austria) with the Bioconductor packages [14]. The *ORF3b* gene from human betacoronavirus (GenBank: DQ497008.1), 20 homologous sequences from the bat betacoronaviruses, and nine homologous sequences from the civet betacoronaviruses (Supplementary file 1) were aligned using ClustalW2 [15] with curation. After the removal of the identical sequences, the *ORF3b* gene from human betacoronavirus, eight homologous sequences from bat betacoronaviruses, and two homologous sequences from civet betacoronaviruses were used for phylogenetic analysis. Since these homologous sequences had high identities (from 85.16% to 99.78%) to the *ORF3b* gene from DQ497008, the Neighbor Joining (NJ) method was used for phylogenetic analysis.

2.2. RNAi and Cellular Experiments

Based on the short hairpin RNA (shRNA) design protocol [1], the 16-nt, 18-nt, 19-nt, 20-nt, and 22-nt segments from the DNA complemented palindrome TCTTTAACAAGCTTGTAAAGA and their control “CGTACGCGGAATACTTCGA” were selected as the target sequences for pSIREN-RetroQ vector construction (Clontech, Mountain View, CA, USA). PC-9 cells were provided by Dr. Qingsong Wang from Tianjin Medical University and they were divided into six groups—namely, 16, 18, 19, 20, 22, and control—for transfection using plasmids containing the 16-nt, 18-nt, 19-nt, 20-nt, and 22-nt segments and the control sequences. Each group had three replicate samples for plasmid transfection and cell apoptosis measurement. Each sample was processed following the procedure described below. The PC-9 cells were washed with phosphate buffer saline (PBS) and were trypsinized 12 h prior to transfection. Gbico RPMI-1640 medium (Thermo Fisher Scientific, Waltham, MA, USA) was added to the cells, which were then centrifuged at 1000 rpm for 10 min at 4 °C to remove the supernatant. Gbico RPMI-1640 medium containing 10% fetal bovine serum was added to adjust the solvent to reach a volume of 2 µL and contain 2×10^5 cells. These cells were seeded into one well of a 6-well plate for plasmid transfection. Transfection of 2 µg of plasmid was performed using 5 µL Lipofectamine 2000 (Life technology, Carlsbad, CA, USA), following the manufacturer’s instructions. Cell apoptosis was measured with the FITC Annexin V Apoptosis Detection Kit I (BD Biosciences, Franklin Lakes, NJ, USA), following the procedure described below. The cells were washed with PBS, were trypsinized, and were collected using a 5-mL culture tube 48 or 72 h after transfection. The culture tube was then centrifuged at 1000 rpm for 10 min at 4 °C to remove the supernatant. The cells were washed twice with cold PBS and were resuspended in 1X Binding Buffer at a concentration of 1×10^6 cells/mL. Then, 100 µL of the solution (1×10^5 cells) was transferred to a new culture tube with 5 µL of FITC-Annexin V and 5 µL PI. The cells were gently vortexed and were incubated for 15 min at room temperature in the dark. 1X Binding Buffer (400 µL) was added to the tube. Finally, the sample was analyzed using a FACSCalibur flow cytometer (BD Biosciences, Franklin Lakes, NJ, USA) within 1 h. Apoptotic cells were quantified by summing the count of the early apoptotic cells (FITC-Annexin V+/PI−) and the late apoptotic cells (FITC-Annexin V+/PI+). To confirm the results using Annexin V/PI staining and detection, the expression levels of three cell-apoptosis marker genes (*BAX*, *BCL2*, and *CASP3*) were also measured by quantitative PCR (qPCR) assays using 16-nt, 18-nt, 19-nt, 20-nt, and 22-nt segments to produce siRNA duplexes by pSIREN-RetroQ plasmid transfection. The cells were collected 24 h after transfection. For each gene, three RNAi samples and three control samples were tested for relative quantification. After transfection, RNA extraction, complementary DNA (cDNA) synthesis, and cDNA amplification were performed following the same procedure described below. For each sample, total RNA was isolated using RNAiso Plus Reagent (TaKaRa Bio Inc., Kusatsu, Shiga, Japan) and the cDNA was synthesized by

M-MuLV (New England Biolabs, Hitchin, UK). The cDNA product was amplified by qPCR (Eppendorf, Hamburg, Germany) using *GAPDH* as internal control and the qPCR reaction mixture using Real Time PCR Easy (Foregene, Chengdu, China) was incubated at 95 °C for 10 min, followed by 35 PCR cycles (10 s at 95 °C, 20 s at 55 °C, and 30 s at 72 °C for each cycle). The primers of five genes are listed in Table 1.

Table 1. Primers for quantitative PCR (qPCR) assays.

Gene Symbol	Forward Primer	Reverse Primer
<i>GAPDH</i>	ACATCGCTCAGACACCATG	TGTAGTTGAGGTCAATGAAGGG
<i>BAX</i>	AGTAACATGGAGCTGCAGAG	AGTAGAAAAGGGCGACAACC
<i>BCL2</i>	GTGGATGACTGAGTACCTGAAC	GCCAGGAGAAATCAAACAGAGG
<i>CASP3</i>	ACTGGACTGTGGCATTGAG	GAGCCATCCTTTGAATTTCGC
<i>MDL1</i> *	CCCAATCCACATCAAAACCC	GGACGAGAAGGGATTGACTG

* The primers of *MDL1* were used to amplify both *MDL1* and *MDL1AS* as they are sense–antisense transcripts from the same loci [9] and were predicted to be markers which can indicate the activities of mitochondria and even the whole cells.

3. Results

3.1. Comparison of siRNA Duplexes Induced by Plant, Invertebrate, and Mammalian Viruses

In total, 11 invertebrate and 8 mammalian viruses (HPV-18, HBV, HCV, HIV-1, SMRV, and EBV that were detected in the previous study, and H1N1 and SARS-CoV that were detected in this study) were used to detect siRNA duplexes (see Materials and Methods). Next, we compared the features of the siRNA duplexes that were induced by invertebrate viruses (Figure 1A) with those that were induced by plant viruses (Figure 1B). The results showed that duplex length was the principal factor that determined the read count in both plants and invertebrates. The 21-nt siRNA duplexes were the most abundant duplexes in both plants and invertebrates, followed by the 22-nt siRNA duplexes in plants, whereas it was the 20-nt siRNA duplexes in invertebrates. The 21-nt siRNA duplexes with 2-nt overhangs were the most abundant 21-nt duplexes in plants, whereas the 21-nt siRNA duplexes with 1-nt overhangs were the most abundant 21-nt duplexes in invertebrates, however they had a very similar read count to that of the 21-nt siRNA duplexes with 2-nt overhangs. The 18-nt, 19-nt, 20-nt, and 22-nt siRNA duplexes in invertebrates had much higher percentages of total sequenced reads than those in plants. In addition, 18-nt and 19-nt siRNA duplexes had very similar read counts, and the 20-nt and 22-nt siRNA duplexes had very similar read counts in invertebrates. As the siRNA duplexes that were induced by mammalian viruses had significantly lower percentages of total sequenced reads, comparison of the siRNA-duplex features between mammals and invertebrates or plants could not provide meaningful results using our data.

However, as an unexpected result of the comparison, we discovered cpsRNAs from both invertebrate and mammalian viruses. As this study was to focus on the study of cpsRNAs from SARS-CoV, we did not include cpsRNAs from other animal viruses. One 19-nt cpsRNA UUAACAAGCUUGUAAAAGA from a DNA complemented palindrome TCTTTAACAAGCTTGTTAAAGA (DQ497008: 25962-25983), located in the *ORF3b* gene of the SARS-CoV MA15 genome (GenBank: DQ497008.1), was detected in four runs of siRNA-seq data (see Materials and Methods). This 19-nt cpsRNA was named SARS-CoV-cpsR-19 and the DNA complemented palindrome contained 22 nucleotides which perfectly matched its reverse complement sequence. From this DNA complemented palindrome, we also detected one 18-nt and one 21-nt cpsRNA named SARS-CoV-cpsR-18 and SARS-CoV-cpsR-21 (Figure 2A), respectively. However we did not detect cpsRNAs of other lengths (e.g., 16-nt, 20-nt, or 22-nt) using our data. We speculated that 18-nt, 19-nt, and one 21-nt cpsRNA could be derived from siRNA duplexes (Figure 2B), as the presence of these overhanging nucleotides is a hallmark of small RNAs that are produced by silencing-related

ribonucleases (e.g., Drosha or Dicer). However, we did not have any evidence to prove the existence of siRNA duplexes.

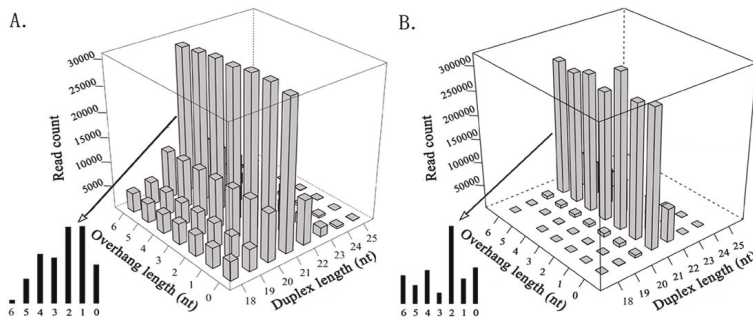


Figure 1. Comparison of small interfering RNA (siRNA) duplexes that were induced by plant and invertebrate viruses. All of the cleaned small RNA sequencing (sRNA-seq) reads were aligned to viral genome sequences using the software Bowtie v0.12.7 [13] with one mismatch. The detection of siRNA duplexes was performed using the program duplexfinder [7]. (A). The read count of siRNA duplexes varies with the duplex length and the overhang length, using data from 11 invertebrate viral genomes. (B). The read count of siRNA duplexes varies with the duplex length and the overhang length, using data from seven plant viral genomes [7].

3.2. Discovery of psRNAs and cpsRNAs

Palindromes have been discovered in the published genomes of most species and play important roles in biological processes. The well-known samples of DNA palindromes include restriction enzyme sites, methylation sites, and palindromic motifs in T cell receptors [16]. In this study, we classified DNA palindromes into DNA palindromes and DNA complemented palindromes, from which psRNAs and cpsRNAs could be produced, respectively. A DNA palindrome is classically defined as a nucleic acid sequence that is reverse complementary to itself, while small RNAs that are reverse complementary to themselves are defined as cpsRNAs in this study. Accordingly, a typical psRNA should have a sequence that is 100% identical to its reverse sequence, however most psRNAs are semipalindromic or heteropalindromic, such as hsa-tiR-MDL1-16 AAAGACACCCCCACA from a DNA palindrome AAAGACACCCCCACAGTTT (NC_012920: 561-580) [9]. As a heteropalindromic psRNA, hsa-tiR-MDL1-16 is derived by cleavage after it starts transcription of the long heavy strand (H-strand) primary transcript at the position 561 of the human mitochondrial genome (Figure 2A). All cpsRNAs that are discovered from SARS-CoV MA15 are semipalindromic or heteropalindromic, such as SARS-CoV-cpsR-19 and SARS-CoV-cpsR-21. Although SARS-CoV-cpsR-21 contains almost 100% of the total nucleotides which contribute to the matches (Figure 2A), most cpsRNAs have mismatches or insertions/deletions (InDels). One example is a new Epstein-Barr virus (EBV) microRNA precursor (pre-miRNA) with a length of 89 nt, as was reported in our previous study [6]. This pre-miRNA sequence contains only 87.64% (78/89) of the total nucleotides which contribute to the matches.

In this study, we also found that DNA palindromes with sizes ranging from 14 to 31 nt and DNA complemented palindromes with sizes ranging from 14 to 53 nt existed ubiquitously in the animal virus genomes, however only a few of them were detected as transcribed or processed into psRNAs or cpsRNAs. For example, we only detected psRNAs from two (CTACTGACCAGTCATC and AAGGTCTCCCTCTGGAA) of 14 DNA palindromes and cpsRNA from two (GCAAATTGCACAATTTGC and TCTTTAACAAGCTTGTTAAAGA) of 29 DNA complemented palindromes (Table 2) in the SARS-CoV genome (GenBank: DQ497008.1) using four runs of sRNA-seq data (see Materials and Methods). One possible reason could be that only a few of cpsRNAs had been ligated to adapters during the sRNA-seq library preparation process as they could form hairpins

(Figure 2A) at room temperature (~20 °C). From Table 2, it can be seen that the T_m (melting temperature) of cpsRNA hairpins in the SARS-CoV genome is distributed ranging from 14 °C to 26 °C. This finding provided a clue to explain why the detected siRNA duplexes that were induced by mammalian viruses had significantly lower percentages of total sequenced reads than those that were induced by plant and invertebrate viruses, and could help to improve the virus detection of mammalian viruses.

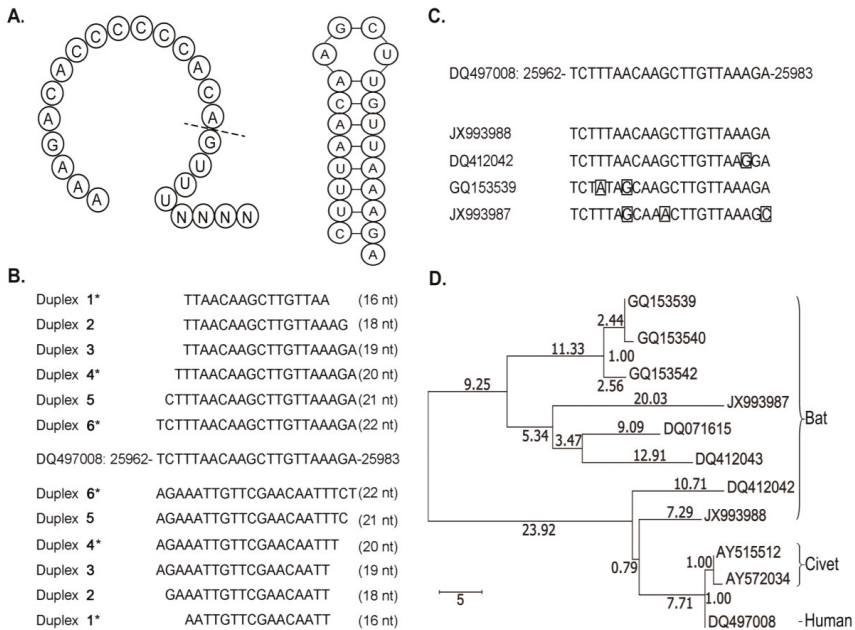


Figure 2. Clues to the origins of SARS-CoV-cpsR-19. All of the genome sequences are represented by their GenBank accession numbers (e.g., DQ497008). (A). hsa-tiR-MDL1-16 is derived by cleavage after it starts transcription of the long H-strand primary transcript at the position 561 of human mitochondrial genome (left). SARS-CoV-cpsR-21 can form a hairpin (right). (B). 16-nt, 18-nt, 19-nt, 20-nt, and 22-nt siRNA duplexes were used for RNA interference (RNAi) experiments. * 16-nt, 20-nt, and 22-nt palindromic small RNAs (cpsRNAs) were not detected in this study. (C). SARS-CoV-cpsR-19 was detected from a DNA complemented palindrome TCTTTAAACAAGCTTGTTAAAGA in the SARS-CoV genome. All of the 22-nt homologous sequences in the civet betacoronavirus genomes were identical to the DNA complemented palindrome, whereas four genotypes of 22-nt homologous sequences were detected in the bat betacoronavirus genomes, however only one of them was identical to it. (D). The phylogenetic tree was built by the Neighbor Joining (NJ) method using the *ORF3b* gene from human betacoronavirus, eight homologous sequences from bat betacoronaviruses, and two homologous sequences from civet betacoronaviruses. The branch's length corresponds to an average number of nucleotide changes per 100 nucleotides.

3.3. Clues to Origins of SARS-CoV-cpsR-19

The previously unknown SARS virus generated widespread panic in 2002 and 2003 when it caused 774 deaths and more than 8000 cases of illness. Scientists immediately suspected that civet cats, which are only distant relatives of house cats, may have been the springboard for the transmission of SARS-CoV to humans [17]. Later, scientists concluded that civets were not the original source of SARS. Further investigation showed that the genetic diversity of coronaviruses in bats increased the possibility of variants crossing the species barrier to cause disease outbreaks in the human population [18]. To investigate the origins of SARS-CoV-cpsR-19, we obtained coronavirus genome sequences that

were associated with bats, palm civets, rats, mice, monkeys, dogs, bovines, hedgehogs, giraffes, waterbucks, and equines from the NCBI GenBank database. The results of the sequence analysis showed that the DNA complemented palindrome TCTTTAACAAGCTTGTTAAAGA was only located in the *ORF3b* genes of betacoronaviruses. Next, we blasted the *ORF3b* gene of human betacoronavirus (GenBank: DQ497008.1) with all of the obtained betacoronavirus genomes, except for those that were obtained from experiments with mice and monkeys. The results showed that the *ORF3b* gene from human betacoronavirus had homologous genes from the betacoronaviruses of bats and palm civets (Supplementary file 1) rather than from those of other species. The DNA complemented palindrome also had 22-nt homologous sequences in the bat and civet betacoronavirus genomes. All of the 22-nt homologous sequences in the civet betacoronavirus genomes were identical to the DNA complemented palindrome, whereas four genotypes of 22-nt homologous sequences were detected in the bat betacoronavirus genomes, however only one of them was identical to it (Figure 2C). Four genotypes had no, one, two, and three mismatches with the DNA complemented palindrome and their corresponding *ORF3b* homologous sequences had identities of 96.77%, 96.13%, 87.96%, and 85.16%, respectively. This suggested that one betacoronavirus variant containing the DNA complemented palindrome could have originated from bats and was then passed onto palm civets and finally to humans. This was consistent with the results of the phylogenetic analysis using the *ORF3b* homologous sequences from the bat and civet betacoronavirus genomes (Figure 2D). In the phylogenetic tree, all of the human and civet betacoronaviruses containing the DNA complemented palindrome were grouped into one clade. The nearest relative of the human and civet clade was the bat betacoronavirus (GenBank: JX993988.1) containing the DNA complemented palindrome, and the next nearest relative was the bat betacoronavirus (GenBank: DQ412042.1) containing a homologous 22-nt sequence with one mismatch with the DNA complemented palindrome.

Table 2. DNA complemented palindromes in the severe acute respiratory syndrome coronavirus (SARS-CoV) genome.

DNA Complementing Palindrome	Start	End	Length	GC%	Tm	MFE
GGTAACTATAAAGTTACC	1783	1800	18	33	20	-6.9
AATGTGAGAATCACATT	2779	2795	17	29	18	-4.4
AAGAAACTAAGTTTCTT	3923	3939	17	24	18	-3.5
ATGGTAAGCTTACCAT	3971	3987	17	35	18	-4.5
AAATGCAAAATCTGCATT	4234	4251	18	28	18	-4.9
ATATGTCTATGACATAT	4949	4965	17	24	18	-4.3
CCTCATGTAAATCATGAGG	5020	5038	19	42	22	-9.0
ATAACAATTGTTAT	5207	5220	14	14	12	-0.9
ACTTCAACAGCTTGAAGT	5241	5258	18	39	18	-4.7
ACTTCAAATTCATTGAAAGT	6256	6275	20	25	20	-5.5
GTACTTTTACTAAAAGTAC	6734	6752	19	26	20	-4.6
ATCTACCAGTGGTAGAT	9189	9205	17	41	20	-6.5
TTACCTTCCAAGGTAA	10,892	10,907	16	38	16	-4.0
CCACITTATAAGTGG	14,160	14,174	15	40	18	-4.7
CCCATTTAATAAATGGG	14,882	14,898	17	35	20	-5.7
CAGTGACAATGTCACTG	16,463	16,479	17	47	22	-7.7
CACCTTTGAAAAAGGTG	16,760	16,776	17	41	20	-5.6
TGTAAGAGAATTTCTTACA	17,651	17,669	19	26	18	-5.3
TGAATATGACTATGCATATTCA	17,783	17,805	23	26	26	-9.7
CTACTTTAAGAAAAGTAG	20,081	20,097	17	29	18	-3.6
AGCAATCTTGGAAATGCT	21,106	21,122	17	41	20	-6.1
TTCCTCTTAAATTAAGAGGAA	21,337	21,357	21	29	22	-9.1
GCAITACTACAGAAGTAATGC	23,599	23,619	21	38	22	-7.9
AGCCCTTTATAAGGGCT	25,480	25,496	17	47	22	-8.7
TCTTTAACAAGCTTGTTAAAGA *	25,962	25,983	22	27	20	-7.2
CAACGGTACTATTACCGTTG	26,406	26,425	20	45	24	-9.0
ACCTTCATGAAGGT	28,048	28,061	14	43	14	-2.6
GCAAATTCACAAATTTGC *	29,028	29,045	18	39	18	-4.0
TAAAATTAATTTTA	29,668	29,681	14	0	NA	NA

In total, 29 DNA complemented palindromes were identified in the SARS-CoV genome (GenBank: DQ497008.1).

* In this study, we only detected psRNAs from two of 29 DNA complemented palindromes. Minimum Free Energy (MFE) was calculated using RNAfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>). Tm (melting temperature) of cpsRNA hairpins was calculated by the formula: $Tm = 4*(C+G) + 2*(A+T)$ and only using the nucleic acids in the stems of DNA complemented palindromes.

3.4. Preliminary Studies on Biological Functions of SARS-CoV-cpsR-19

Our previous study showed that the psRNA hsa-tiR-MDL1-16 contained the Transcription Initiation Site (TIS) of the human mitochondrial H-strand and could be involved in mtDNA transcription regulation [9]. This inspired us to speculate that cpsRNAs could also have specific biological functions and we investigated SARS-CoV-cpsR-19 using RNAi and Annexin V/PI staining and detection. 16-nt, 18-nt, 19-nt, 20-nt, and 22-nt segments from the DNA complemented palindrome TCTTTAACAAGCTTGTTAAAGA were used to produce siRNA duplexes by pSIREN-RetroQ plasmid transfection (see Materials and Methods). As a result, the 19-nt and 20-nt segments significantly induced cell apoptosis 2.76- and 1.48-fold 48 h after their transfection, respectively. Particularly, the 19-nt segment significantly induced cell apoptosis 7.94-fold (36.04%/4.54%) 72 h after its transfection into PC-9 cells (Figure 3). Using the 19-nt segment, we also tested cell apoptosis in five other human cell lines and one mouse cell line. The results showed that the 19-nt segment significantly induced cell apoptosis in the A549, MCF-7, and H1299 cell lines, however it did not in the MB231, H520, and L929 (mouse) cell lines. The siRNA duplexes that were induced by the 19-nt segment could silence cell-specific transcripts to induce cell apoptosis through RNAi. These results suggested that SARS-CoV-cpsR-19 had significant biological functions and could play a role in SARS-CoV infection or pathogenesis.

To confirm the results using Annexin V/PI staining and detection, the expression levels of three cell-apoptosis marker genes (*BAX*, *BCL2*, and *CASP3*) were measured by qPCR assays using the same RNAi protocol in the HeLa cell line (see Materials and Methods). The results showed 1.08-, 1.65-, 1.60-, 1.24-, and 1.30-fold increases in *BAX/BCL2* ratios which were caused by 16-nt, 18-nt, 19-nt, 20-nt, and 22-nt segments, respectively (Figure 3E). As a low *BAX/BCL2* ratio (<1) and a high *BAX/BCL2* ratio (>1) is against and in favor of cell apoptosis, it suggested that 18-nt and 19-nt segments significantly induced cell apoptosis. The RNAi using 18-nt, 19-nt, and 20-nt segments also caused significantly increased expression of *CASP3* by 1.72, 2.58, and 1.89 fold changes in HeLa cells. Therefore, the experimental results by qPCR assays were consistent with those using Annexin V/PI staining and detection. In addition, the expression levels of two novel long non-coding RNAs (lncRNAs) (*MDL1* and *MDLIAS*) from mitochondrial genome were measured to investigate their expression changes that were caused by RNAi using 16-nt, 18-nt, 19-nt, 20-nt, and 22-nt segments. *MDL1* and *MDLIAS* were recently discovered [9] and were predicted to be markers which can indicate the activities of mitochondria and even the whole cells. The RNAi using 16-nt, 18-nt, 19-nt, 20-nt, and 22-nt segments also caused increased relative expression levels of *MDL1* by 1.28, 5.43, 1.58, 9.25, and 1.41 fold changes in HeLa cells (Figure 3E). It suggested that qPCR of *MDL1* produce higher sensitivities than that of *BAX/BCL2* and *CASP3* in the detection of cell apoptosis.

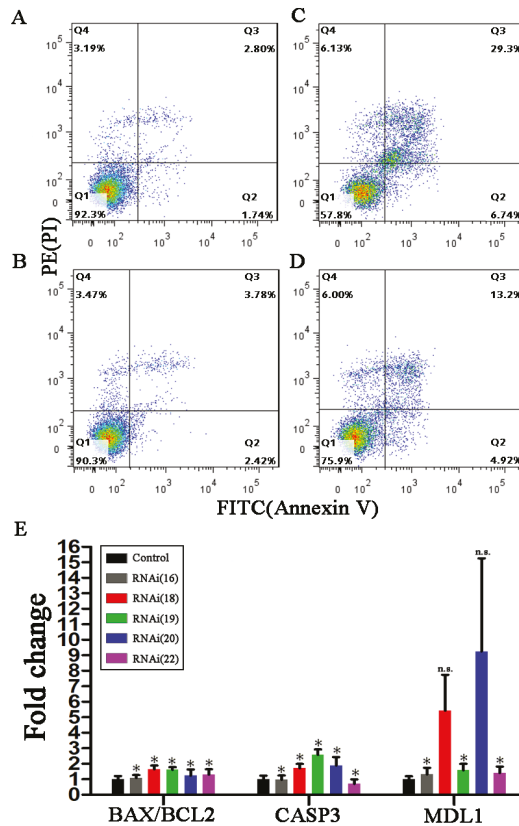


Figure 3. RNAi and cellular experiments for validation. PC-9 cells were divided into six groups named 16, 18, 19, 20, 22, and control for transfection using plasmids containing 16-nt, 18-nt, 19-nt, 20-nt, 22-nt segments, and their controls (Figure 2B). Each group had three replicate samples for plasmid transfection and cell apoptosis measurement. The 19-nt and 20-nt segments significantly induced cell apoptosis, whereas the 16-nt, 18-nt, and 22-nt did not show significantly positive results. The samples in this figure were selected randomly from the control (A), 18 (B), 19 (C), and 20 (D) group. (E). The experimental results by qPCR assays were consistent with those using Annexin V/PI staining and detection. The primers of *MDL1* were used to amplify both *MDL1* and *MDLIAS*, as they are sense–antisense transcripts from the same loci [9] and are predicted to be markers which can indicate the activities of mitochondria and even the whole cells.

4. Conclusions and Discussion

SARS-CoV-cpsR-19 from the DNA complemented palindrome TCTTTAACAAGCTTGTTAAAGA (DQ497008: 25962–25983) is located in the *ORF3b* gene of the SARS-CoV MA15 genome. In one previous study, this 22-nt DNA palindrome was studied using two mathematics models (M0 and M1) to conclude that this DNA palindrome was quite unlikely to occur by chance [19]. These authors also found that the underrepresentation of 4-nt DNA palindromes existed in all of the coronaviruses and the significant underrepresentation of 6-nt DNA palindromes existed only in SARS but not in the other six coronaviruses. Then, they hypothesized that this avoidance of 6-nt DNA palindromes in the SARS genome would offer a protective effect on the virus, making it comparatively more difficult to be destroyed by some host-defence mechanisms (e.g., restriction enzymes). In this study, we unexpectedly discovered cpsRNAs from the DNA palindrome TCTTTAACAAGCTTGTTAAAGA and reported

that SARS-CoV-cpsR-19 could play a role in SARS-CoV infection or pathogenesis. We hypothesized that cpsRNAs, as a novel class of small RNAs, were produced by some host-defence mechanisms (e.g., RNAi) and that they could interact with the immune system of the host cells to help virus survival or infection.

Another previous study reported that another sRNA AGGAACUGGCCAGGAAGCUUC in the SARS-CoV genome, named small viral RNA-N (svRNA-N), contributed to SARS-CoV pathogenesis [20]. SARS-CoV-cpsR-19 was detected in very low abundance using four runs of sRNA-seq data (see Materials and Methods) in this study but was not detected using the sRNA-seq data (NCBI SRA: SRP094035) in that study [20]. One possible reason could be that SARS-CoV-cpsR-19 can form one hairpin and the T_m (melting temperature) of this hairpin is about 20 °C. Therefore, most of SARS-CoV-cpsR-19 cannot be ligated to adapters during the sRNA-seq library preparation process at room temperature (~20 °C). As the optimal secondary structure of svRNA-N (Minimum Free Energy (MFE) = -1.7) is less stable than those of SARS-CoV-cpsR-19 (MFE = -2.6) and all the other possible cpsRNAs (Table 2), svRNA-N is more easily captured during the sRNA-seq library preparation process. SARS-CoV-cpsR-19 induced cell apoptosis, whereas svRNA-N caused pulmonary inflammation. In Vivo experiments revealed that the biogenesis pathway that was responsible for svRNA-N's synthesis was not dependent on Dicer or Drosha as the canonical biogenesis of miRNAs and svRNA-N were able to silence mRNA expression by targeting its 3' UTR. However, the mechanisms underlying infection or pathogenesis of both SARS-CoV-cpsR-19 and svRNA-N are still unknown.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4425/9/9/442/s1>, Table S1: 51 runs of sRNA-seq data for invertebrate virus detection, Table S2: Homologous sequences of ORF3b.

Author Contributions: S.G. conceived this project. S.G. and W.B. supervised this project. S.G., Z.C., and D.Y. analyzed the data. H.J. and S.L. curated the sequences and prepared all of the figures, tables, and additional files. C.L., Y.H., and W.S. performed the experiments. S.G. drafted the main manuscript. S.G. and J.R. revised the manuscript. All of the authors have read and approved the manuscript.

Funding: This work was supported by grants from the National Key Research and Development Program of China (2016YFC0502304-03) to Defu Chen, Central Public-Interest Scientific Institution Basal Research Fund of Lanzhou Veterinary Research Institute of CAAS to Ze Chen, Natural Science Foundation of China (81472052) to Qingsong Wang, and Internationalization of Outstanding Postdoctoral Training Program from Tianjin government to Shan Gao.

Acknowledgments: We thank Dr. Qingsong Wang from Tianjin Medical University for her cell lines, Yangguang Han and Jinke Huang from Tianjin KeYiJiaXin Technology Co., Ltd. for his professional guidance on the RNAi and cellular experiments. This manuscript has been released as a pre-print at <https://www.biorxiv.org/content/early/2017/12/13/185876>.

Conflicts of Interest: Non-financial competing interests are claimed in this study.

References

- Chen, Z.; Sun, Y.; Yang, X.; Wu, Z.; Guo, K.; Niu, X.; Wang, Q.; Ruan, J.; Bu, W.; Gao, S. Two featured series of rRNA-derived RNA fragments (rRFs) constitute a novel class of small RNAs. *PLoS ONE* **2017**, *12*, e0176458. [[CrossRef](#)] [[PubMed](#)]
- Kreuze, J.F.; Perez, A.; Untiveros, M.; Quispe, D.; Fuentes, S.; Barker, I.; Simon, R. Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: A generic method for diagnosis, discovery and sequencing of viruses. *Virology* **2009**, *388*, 1–7. [[CrossRef](#)] [[PubMed](#)]
- Li, R.; Gao, S.; Hernandez, A.G.; Wechter, W.P.; Fei, Z.; Ling, K. Deep sequencing of small RNAs in tomato for virus and viroid identification and strain differentiation. *PLoS ONE* **2012**, *7*, e37127. [[CrossRef](#)] [[PubMed](#)]
- Zheng, Y.; Gao, S.; Chellappan, P.; Li, R.; Marco, G.; Dina, G.; Segundo, F.; Ling, K.; Jan, K.; Fei, Z. VirusDetect: An automated pipeline for efficient virus discovery using deep sequencing of small RNAs. *Virology* **2017**, *500*, 130–138. [[CrossRef](#)] [[PubMed](#)]
- Nayak, A.; Tassetto, M.; Kunitomi, M.; Andino, R. *RNA Interference-Mediated Intrinsic Antiviral Immunity in Invertebrates*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 371, pp. 183–200.
- Wang, F.; Sun, Y.; Ruan, J.; Chen, R.; Chen, X.; Chen, C.; Kreuze, J.F.; Fei, Z.; Zhu, X.; Gao, S. Using small RNA deep sequencing to detect human viruses. *BioMed Res. Int.* **2016**, *2016*, 2596782. [[CrossRef](#)] [[PubMed](#)]

7. Niu, X.; Sun, Y.; Chen, Z.; Li, R.; Padmanabhan, C.; Ruan, J.; Kreuze, J.F.; Ling, K.; Fei, Z.; Gao, S. Using small RNA-seq data to detect siRNA duplexes induced by plant viruses. *Genes* **2017**, *8*, 163. [[CrossRef](#)] [[PubMed](#)]
8. Roberts, A.; Deming, D.; Paddock, C.D.; Cheng, A.; Yount, B.; Vogel, L.; Herman, B.D.; Sheahan, T.; Heise, M.; Genrich, G.L. A mouse-adapted SARS-coronavirus causes disease and mortality in BALB/c mice. *PLoS Pathog.* **2007**, *3*, e5. [[CrossRef](#)] [[PubMed](#)]
9. Gao, S.; Tian, X.; Chang, H.; Sun, Y.; Wu, Z.; Cheng, Z.; Dong, P.; Zhao, Q.; Ruan, J.; Bu, W. Two novel lncRNAs discovered in human mitochondrial DNA using PacBio full-length transcriptome data. *Mitochondrion* **2017**, *36*. [[CrossRef](#)] [[PubMed](#)]
10. Elbashir, S.M.; Harborth, J.; Weber, K.; Tuschl, T. Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* **2002**, *26*, 199–213. [[CrossRef](#)]
11. Peng, X.; Gralinski, L.; Ferris, M.T.; Frieman, M.B.; Thomas, M.J.; Proll, S.; Korth, M.J.; Tisoncik, J.R.; Heise, M.; Luo, S. Integrative deep sequencing of the mouse lung transcriptome reveals differential expression of diverse classes of small RNAs in response to respiratory virus infection. *Mbio* **2011**, *2*, e00198-11. [[CrossRef](#)] [[PubMed](#)]
12. Zhang, M.; Zhan, F.; Sun, H.; Gong, X.; Fei, Z.; Gao, S. Fastq_clean: An optimized pipeline to clean the Illumina sequencing data with quality control. In Proceedings of the 2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Belfast, UK, 2–5 November 2014.
13. Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **2009**, *10*, R25. [[CrossRef](#)] [[PubMed](#)]
14. Gao, S.; Ou, J.; Xiao, K. *R Language and Bioconductor in Bioinformatics Applications, Chinese ed.*; Tianjin Science and Technology Translation Publishing Co., Ltd.: Tianjin, China, 2014.
15. Clustal Omega. Available online: <http://www.ebi.ac.uk/Tools/msa/clustalo/> (accessed on 18th May 2014).
16. Srivastava, S.K.; Robins, H.S. Palindromic nucleotide analysis in human T cell receptor rearrangements. *PLoS ONE* **2012**, *7*, e52250. [[CrossRef](#)] [[PubMed](#)]
17. Enserink, M. Clues to the animal origins of SARS. *Science* **2003**, *300*, 1351. [[CrossRef](#)] [[PubMed](#)]
18. Li, W.; Shi, Z.; Yu, M.; Ren, W.; Smith, C.; Epstein, J.H.; Wang, H.; Crameri, G.; Hu, Z.; Zhang, H. Bats are natural reservoirs of SARS-like coronaviruses. *Science* **2005**, *310*, 676–679. [[CrossRef](#)] [[PubMed](#)]
19. Chew, D.S.H.; Choi, K.P.; Heidner, H.; Leung, M.Y. Palindromes in SARS and other coronaviruses. *Inform. J. Comput.* **2004**, *16*, 331–340. [[CrossRef](#)] [[PubMed](#)]
20. Morales, L.; Oliveros, J.C.; Fernandez-Delgado, R.; Tenoever, B.R.; Enjuanes, L.; Sola, I. SARS-CoV-encoded small RNAs contribute to infection-associated lung pathology. *Cell Host Microbe* **2017**, *21*, 344–355. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

Serum and Lipoprotein Particle miRNA Profile in Uremia Patients

Markus Axmann ¹, Sabine M. Meier ¹, Andreas Karner ², Witta Strobl ¹, Herbert Stangl ^{1,*} and Birgit Plochberger ^{2,*}

¹ Center for Pathobiochemistry and Genetics, Institute of Medical Chemistry and Pathobiochemistry, Medical University Vienna, 1090 Vienna, Austria; markus.axmann@meduniwien.ac.at (M.A.); sabine.meier@meduniwien.ac.at (S.M.M.); witta.strobl@meduniwien.ac.at (W.S.)

² University of Applied Sciences Upper Austria, School of Medical Engineering and Applied Social Sciences, 4020 Linz, Austria; andreas.karner@fh-linz.at

* Correspondence: herbert.stangl@meduniwien.ac.at (H.S.); birgit.plochberger@fh-linz.at (B.P.); Tel.: +43-1-40160-38023 (H.S.); +43-5-0804-52131 (B.P.)

Received: 6 September 2018; Accepted: 30 October 2018; Published: 5 November 2018

Abstract: microRNAs (miRNAs) are post-transcriptional regulators of messenger RNA (mRNA), and transported through the whole organism by—but not limited to—lipoprotein particles. Here, we address the miRNA profile in serum and lipoprotein particles of healthy individuals in comparison with patients with uremia. Moreover, we quantitatively determined the cellular lipoprotein-particle-uptake dependence on the density of lipoprotein particle receptors and present a method for enhancement of the transfer efficiency. We observed a significant increase of the cellular miRNA level using reconstituted high-density lipoprotein (HDL) particles artificially loaded with miRNA, whereas incubation with native HDL particles yielded no measurable effect. Thus, we conclude that no relevant effect of lipoprotein-particle-mediated miRNA-transfer exists under in vivo conditions though the miRNA profile of lipoprotein particles can be used as a diagnostic marker.

Keywords: microRNA; lipoprotein; transfer; uremia

1. Introduction

In general, intercellular communication occurs via cell-to-cell adhesion or by secretory messengers via e.g., hormones. A majority of circulating microRNAs (miRNAs) are derived from shedding of plasma membrane vesicles [1–4]. However, the protein Argonaute2 was also reported to carry a population of circulating miRNA without the need for vesicles [5]. Recently, high-density lipoprotein (HDL) particles were shown to transfer proteins and miRNAs from donor to recipient cells [6,7]. miRNAs are a class of small, non-coding RNAs which have emerged as important post-transcriptional regulators [8]. They intracellularly target specific messenger RNAs (mRNAs) to degrade or repress the translation process [8–11]. In extracellular space, miRNAs were thought to be relative unstable molecules; however, surprisingly, they were found to circulate in a highly stable form in various body fluids like blood, milk, saliva and urine [12]. Turchinovich et al. [13] proposed that extracellular miRNA carried by HDL is associated with the Argonaute family of proteins. However, proteome analysis did not reveal these proteins as a component of HDL [14–19]. The profile of circulating miRNAs display characteristic changes when pathogenic alterations occur. This dysregulation is linked to diseases such as cancer, cardiovascular [20–24] and kidney diseases [25–28]. Altered miRNA profiles in the HDL fraction of patients suffering from familial hypercholesterolemia were seen [6]. Thus, they may serve as markers for various diseases [29] and their progression [12] or the development of new therapeutic agents [8,30]. Thereby, mode and specificity

of HDL uptake routes may influence the functional role of secreted miRNAs. This is one of the least understood issues in the field of miRNAs.

Here, we present data that miRNA is present—besides serum—in all lipoprotein particles and find that the miRNA profile is different in patients with uremia. Finally, we aimed to follow the uptake of miRNA to cells using quantitative real-time PCR (qPCR). For this purpose, we derived a protocol to increase the uptake for miRNA from HDL particles significantly. Taken together, our data indicate that—despite the presence of miRNA on lipoprotein particles—its abundance is rather low, which makes any influence on the cell metabolism via lipoprotein-mediated transfer in vivo generally highly unlikely.

2. Materials and Methods

The manuscript and the used nomenclature are written according to the minimum information for publication of quantitative real-time PCR (qPCR) experiments (MIQE) guidelines [31].

2.1. Patients

For the miRNA TaqMan™ (Applied Biosystems, Carlsbad, CA, USA) array analyses, 17 adult chronic renal failure (CRF) patients suffering from CRF Kidney Disease (Outcomes Quality Initiative stage 3–5 without hemodialysis) and 14 adult CRF patients on maintenance hemodialysis and matched controls were recruited (for lipid analyses see Table 1 of Meier et al. [32]). This study was approved by the Ethics Committee, Medical University of Vienna (EK-Nr. 511/2007). Written informed consent was obtained from all participants. Patients with an age over 55 years, with diabetes mellitus, nephrotic syndrome, severe hyperlipidemia, inflammatory diseases, malignancy and infections within the last three months, and patients on corticosteroids, lipid lowering or immunosuppressive drugs were excluded. Venous blood was collected after an overnight fast. Sera were frozen in aliquots at -20°C . Lipoprotein particles were isolated immediately and stored frozen in aliquots at -20°C until analysis. For detailed lipid and apolipoprotein analyses see Meier et al. [32].

2.2. microRNA in Serum and Lipoprotein Particles

As there is no known endogenous control for miRNAs in serum or lipoprotein particles, we added cel-miR-39 from the *C. elegans* species as internal control and for reference [33]. Total miRNA was reverse transcribed using the TaqMan™ microRNA Reverse Transcription Kit and the appropriate reverse transcription (RT) primers (Applied Biosystems), according to the manufacturer's protocol. Resulting samples were subjected to RT-qPCR using TaqMan™ Gene Expression Universal Master Mix (Applied Biosystems) and TaqMan™ miRNA Assays (Applied Biosystems) following the manufacturer's instructions. The amplification was conducted in a StepOne Real-Time PCR-System (48-well, Applied Biosystems); data was collected using the StepOne Software v2.1 (Applied Biosystems).

2.3. TaqMan™ Arrays

Equal volumes of patient samples were pooled and miRNA was isolated from 100 μL serum or 500 μg lipoprotein particles using the miRNeasy Kit (QIAGEN GmbH, Hilden, Germany), according to the supplier's instructions, in two independent experiments. RNA quantity and purity was measured using a NanoDrop ND-1000 Spectrophotometer (peqlab Biotechnologie GmbH, Erlangen, Germany) and RNA was stored at -80°C . Total miRNA was reverse transcribed using the TaqMan™ microRNA RT Kit with MegaPlex RT Primers (Applied Biosystems). To ensure the sample content for the TaqMan™ arrays, miRNAs from lipoprotein particles and serum were further processed by applying preamplification using the appropriate primers and the PreAmp Master Mix (Applied Biosystems). Afterwards, samples were loaded into TaqMan™ Array Cards A+B (in total 754 miRNAs were detected) and analyzed using a 7900HT Fast Real-Time PCR System (Applied Biosystems). Data was collected and analyzed using appropriate software from Applied Biosystems.

(SDS 2.4 and Data assist v3.0) yielding $\Delta\Delta c_q$. The value RQ (relative quantification) is defined as $RQ = 2^{-\Delta\Delta c_q}$.

2.4. Synthetic miRNA

Human mature miRNAs hsa-miR145 (5'-GUC CAG UUU UCC CAG GAA UCC CU-3'), hsa-miR155 (5'-UUA AUG CUA AUC GUC AUA GGG GU-3') and hsa-miR223 (5'-UGU CAG UUU GUC AAA UAC CCC A-3') were synthesized by Microsynth (Microsynth, Vienna, Austria). The manufacturer did purification with HPLC & dialysis. miRNAs were solubilized in 10 mM tris(hydroxymethyl) aminomethane (TRIS) buffer, pH 7.5 (Thermo Fisher Scientific, Vienna, Austria) and stored at $-20\text{ }^{\circ}\text{C}$ in aliquots of 100 μL (final storage concentration 10 μM).

2.5. Lipoprotein Particle Isolation

For the reconstitution/labeling experiments, human plasma was collected from two normolipidemic healthy volunteers twice (time between donations was roughly one month), in accordance with the medical and ethical guidelines of the Medical University of Vienna. This part of the study was approved by the Ethics Committee, Medical University of Vienna (EK-Nr. 1414/2016). Written informed consent was obtained from all participants. Individual lipoprotein particle (HDL and low-density lipoprotein (LDL)) fractions were isolated by serial ultra-centrifugation at a density of 1.21 g/mL or 1.06 g/mL, respectively [34]. Final protein concentration was determined photometrically (Bradford assay) and samples were stored under an inert atmosphere at $+4\text{ }^{\circ}\text{C}$. For the preparation of lipoprotein particle deficient serum (LPDS), human sera from both donors were spun using ultra-centrifugation at a density of 1.21 g/mL, dialyzed and stored at $-20\text{ }^{\circ}\text{C}$.

2.6. Reconstitution of HDL Particles

HDL particles were reconstituted by a modified protocol, previously published in [35]. In short, lipids from HDL particles were extracted two times with ethanol : diethyl ether (3:2) at $-20\text{ }^{\circ}\text{C}$ for 2 h. Precipitate was dried under nitrogen gas flow and resuspended in buffer A (150 mM NaCl, 0.1% ethylenediaminetetraacetic acid (EDTA), 10 mM TRIS/HCl, pH 8.0, all Sigma Aldrich, Vienna, Austria). Protein concentration was determined photometrically (Bradford assay). A lipid mixture, consisting of L- α -phosphatidylcholine, cholesterol oleate and cholesterol (all Sigma Aldrich) at a molar ratio of 100:22:4.8 dissolved in chloroform : methanol (2:1), was dried under nitrogen gas and resuspended in buffer A. Aliquots of synthetic miRNAs (100 μL , 10 μM) were mixed with freshly prepared spermine solution (final concentration 15 mM, Sigma Aldrich) for 30 min at $30\text{ }^{\circ}\text{C}$. Lipid suspension and miRNA/spermine solution were mixed and sodium deoxycholate (Sigma Aldrich) was added for lipid solubilization at a final concentration of 15 mM. In negative control experiments, HDL particles were reconstituted without addition of miRNA and/or spermine. The mixture was stirred at $+4\text{ }^{\circ}\text{C}$ for 2 h. Delipidated HDL was added at a final molar ratio of L- α -phosphatidylcholine, cholesterol oleate, cholesterol and protein (HDL) of 100:22:4.8:1, and the mixture was stirred at $+4\text{ }^{\circ}\text{C}$ overnight. Extensive dialysis using Slide-A-Lyzer™ dialysis cassettes (cut-off 20 kDa, Thermo Fisher Scientific) against phosphate-buffered saline (PBS, Roth, Graz, Austria) + Amberlite XAD-2 polymeric adsorbent (15 g/L, Sigma Aldrich) was performed to separate generated reconstituted HDL (rHDL) particles from free miRNA, spermine and detergent. Final protein concentration was determined photometrically (Bradford assay) and samples were stored under inert atmosphere at $+4\text{ }^{\circ}\text{C}$.

2.7. Labeling of LDL Particles

LDL particles were labeled with miRNA by a modified protocol, previously published in [36]. In short, LDL particle solution was incubated with 0.3 mM EDTA/0.1 mM ethylene glycol-bis(β -aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA, Roth) for 10 min at $+4\text{ }^{\circ}\text{C}$. Aliquots of miRNA were mixed with freshly prepared spermine solution (final concentration 15 mM, Sigma Aldrich) for 30 min at $30\text{ }^{\circ}\text{C}$. After addition of an equal volume of dimethyl sulfoxide

DMSO (Thermo Fisher Scientific), the mixture was diluted five-fold with buffer containing 150 mM NaCl/0.3 mM EDTA/0.1 mM EGTA. LDL particle solution and miRNA mixture were combined and incubated for 2 h at 40 °C. In negative control experiments, LDL particles were labeled without addition of miRNA and/or spermine. Extensive dialysis using Slide-A-Lyzer™ dialysis cassettes (cut-off 20 kDa, Thermo Fisher Scientific) against PBS (Roth) + Amberlite XAD-2 polymeric adsorbent (15 g/L, Sigma Aldrich) was performed to separate labeled LDL particles from free miRNA, spermine and organic solvent. Final protein concentration was determined photometrically (Bradford assay) and samples were stored under inert atmosphere at +4 °C.

2.8. Quality Control of Reconstituted/Labeled Lipoprotein Particles

High-speed atomic force microscopy (HS-AFM) (see Figure A1) was performed in tapping mode with free amplitudes of 1.5 nm–2.5 nm. The amplitude setpoint was larger than 90% of the free oscillation amplitude. USC-F1.2-k0.15 cantilevers (Nanoworld AG, Neuchâtel, Switzerland) were used. HDL and LDL particle aliquots were diluted 1:10³ in PBS (in order to observe individual particles) and incubated for 5 min on freshly cleaved mica. After rinsing with PBS, the samples were imaged in PBS at room temperature. Image processing and particle analysis were done in Gwyddion 2.49 (CMI, Brno, Czech Republic) and SPIP (Image Metrology, Hørsholm, Denmark). The probability density functions (pdfs) representing the height distribution of particles were calculated using the ‘ks-density’ algorithm in MATLAB (Mathworks, Natick, MA, USA).

2.9. Cells

Chinese hamster ovary (CHO) cell lines CHOK1, IdIA7 (reduced expression of LDL-receptor), and IdIA7-SR-B1 (overexpression of scavenger receptor class B type 1 (SR-B1)) were cultured as previously described in [37,38]. In short, cells were maintained in Dulbecco’s modified Eagle Medium/Nutrient Mixture F-12 (Sigma Aldrich) supplemented with 10% fetal calf serum (FCS, Sigma Aldrich), Penicillin/Streptomycin (100 U/mL and 0.1 mg/mL final concentration, Sigma Aldrich) and 2 mM L-Glutamine (Roth). Geneticin sulphate (Thermo Fisher Scientific) was added to IdIA7-SR-B1 cells at a final concentration of 0.5 mg/mL as selection antibiotic.

Cells were grown until just reaching confluency in individual LabTek™ chambers (eight chambers per slide, each with a 4 mm × 4 mm surface area and ~700 µL total volume, Thermo Fisher Scientific). Cells were washed three times with Hanks’ balanced salt solution (HBSS, Roth) and incubated in FCS-free culture medium. Next, HDL particle solution (final concentration 50 µg/mL or 5 µg/mL) or LDL particle solution (final concentration 5 µg/mL) from the different preparations (native/reconstituted/labeled) was added. Incubation was performed at 37 °C and 5% CO₂. After 16 h, cells were washed three times with HBSS and covered with 100 µL of medium without FCS. For negative control experiments, chambers without cells and with cells without any addition of particles were used. For blocking experiments (testing of unspecific binding), native lipoprotein particle solution or bovine serum albumin (BSA, Sigma Aldrich) was used. Cell number was determined for each experiment in two independent chambers, which were similarly treated as above (medium containing 50 µg/mL HDL or 5 µg/mL LDL particle solution) and the average cell number of both chambers was used for normalization.

2.10. miRNA-Extraction

miRNA-extraction was performed according to the manufacturer’s protocol of the miRNeasy Mini Kit (Qiagen, Vienna, Austria). In short, a sample volume (ranging from 1 µL to 100 µL (lowest protein concentration)) of native/reconstituted HDL particles or native/labeled LDL particles normalized to the lowest protein concentration was lysed and homogenized. The cell sample volume used for miRNA-extraction contained the pooled cells from two independent chambers. For the negative control sample, 100 µL of RNase-free water (Roth) was used. A phenol/chloroform extraction was performed subsequently to separate proteins/DNA from RNA.

2.11. Reverse Transcription

Reverse transcription was performed according to the manufacturer's protocol of the TaqMan™ microRNA Reverse Transcription Kit (Thermo Fisher Scientific). Briefly, 7 µL of the kit's mastermix solution were added to 5 µL of the extracted miRNA-sample and 3 µL of miRNA-specific primer (Assay ID 002623, 000526, and 002278, Thermo Fisher Scientific). In order to use the same cell number for each cell line, the sample volume of the cell line with a higher overall cell number was reduced accordingly; the residual volume to reach the total sample volume of 5 µL was RNase-free water (thus, sample volume from the individual cell lines varied between 5 µL and 2 µL). As a sample for standard curve preparation, an aliquot of miRNA was diluted sequentially in RNase-free water. Reverse transcription was performed using a Thermocycler (LabCycler Sensoquest, Göttingen, Germany) with the following program: (1) 30 min at 16 °C, (2) 30 min at 42 °C, (3) 5 min at 85 °C, (4) ∞ at 4 °C. Usually, the qPCR step was done immediately after the reverse transcription—otherwise the complementary DNA synthesized from the miRNA samples was stored at −20 °C.

2.12. qPCR

qPCR of the reverse transcribed samples was performed according to the manufacturer's protocol of the TaqMan™ assay (Thermo Fisher Scientific). In short, per sample, 7.5 µL of mastermix solution (iTaq™, Biorad, Vienna, Austria) was diluted with 4.75 µL RNase-free water and 0.756 µL hydrolysis probe (TaqMan™ Assay ID 002623, 000526, and 002278, Thermo Fisher Scientific). Each sample contained further 2 µL of cDNA sample or RNase-free water and was measured twice. qPCR was performed using a Corbett RG-6000 PCR machine (Corbett Research, Cambridge, UK) with the following program: (1) 2 min at 50 °C, (2) 10 min at 95 °C, (3) 15 s at 95 °C, (4) 60 s at 60 °C. Steps (3) and (4) were repeated up to 50 times. For the analysis, RotorGene 6000 Software Version 1.7 (Qiagen, Vienna, Austria) was used with activated 'DynamicTube Normalization' (for compensation of different background levels using the second derivative of each sample trace) and 'Noise Slope Correction' (normalization to the noise level). The threshold for calculation of the quantification cycle c_q for each miRNA was determined from the standard curves of each miRNA individually, via the 'Auto-Find Threshold' function of the software package—hereby, the software maximizes the R-value of the fit of the standard curve—and kept equal for each specific miRNA. The c_q value of the negative control was always at least six cycles higher than the highest sample c_q value (in the range between 30–40 cycles). In each run, at least one of the standard dilutions of miRNA was added as a sample to calibrate each individual run for the same reaction efficiency.

2.13. Calculation of miRNA Strands Number

For standard curves (Figure A2), the number of miRNA strands of each sequential dilution step was calculated from the initial concentration and the subsequent dilution steps during reverse transcription and PCR. Thereby, we assumed that the reverse transcription performed linearly over the low concentration range used here and in a 1:1 stoichiometry. Accordingly, the number of strands corresponding to the c_q value of each sample run was calculated from the standard curve regression line and normalized to the number of cells or the lipoprotein particle number, respectively. Lipoprotein particle number was estimated from the initial protein concentration and its average molecular weight (molecular weight MW_{HDL} ~250 kDa, MW_{LDL} ~3 MDa). We assumed no lipid contribution to the molecular weight—thus, we slightly overestimated the number of miRNA strands per lipoprotein particle. We further assumed a 100% recovery of miRNA during the miRNA extraction step.

3. Results

3.1. miRNA in Lipoprotein Particle Fractions

First, we tested the presence of miRNA in serum and all lipoprotein particle fractions of healthy volunteers. To do so, we separated serum and lipoprotein particle fractions and purified the contained miRNAs. For analysis, we selected two miRNAs known to be involved in renal and cardiovascular disease, namely miR-21 [39] and miR-223 [40]. Both miRNAs were detectable in all fractions (Table 1) with the highest relative amount in HDL particles; however, the abundance was low (absolute c_q values varied between 26 and 36 cycles). This data indicates a lipoprotein-particle-selective association of miRNAs. To demonstrate that HDL particles did not contain remaining serum constituents, we diluted the HDL particle fraction with excess PBS and centrifuged it a second time. Still, this fraction contained the highest miRNA amount among the lipoprotein particle classes ($\langle \Delta \Delta c_q \rangle = -2.5$ (SD = 0.2) and -1.5 (SD = 0.1) for miR-223 and miR-21, respectively), indicating that miRNAs are relatively enriched at least in the HDL particle fraction.

Table 1. miRNAs circulating in human serum and associated to lipoprotein particles.

Sample	miRNA	$\langle \Delta \Delta c_q \rangle$	SD $\langle \Delta \Delta c_q \rangle$
LPDS	223	0.0	0.1
	21	0.0	0.2
VLDL	223	5.7	0.2
	21	4.6	0.2
LDL	223	2.8	0.2
	21	1.8	0.2
HDL	223	-3.5	0.1
	21	-2.5	0.1

Lipoprotein particle deficient serum (LPDS) and lipoprotein particle fractions (very-low density lipoprotein (VLDL), low density lipoprotein (LDL) and high density lipoprotein (HDL) particles) were analyzed for two distinct microRNAs (miRNAs). Values are in relation to the levels of cel-miR-39 and subsequently LPDS. ($\langle x \rangle$ = average value of x , SD = standard deviation, $n = 3$ independent experiments).

Next, we investigated miRNA profiles in serum and HDL particles of uremic patients recruited for a study on the influence of chronic renal failure (CRF) on cholesterol efflux from macrophages [32]. Figure 1 depicts the expression profile of miRNAs found in the serum/HDL particles of patients suffering from CRF (top) and from the more advanced disease course (hemodialysis, bottom) relative to their matched healthy controls. In both settings, the miRNA profile difference between serum and HDL particles was diverse: some miRNAs were represented in HDL particles and serum in a similar portion (see Figure 1 overlap of the blue and orange line); however, some miRNAs were increased at different rates in serum and HDL particles. The most regulated miRNA in CRF patients—miR-192—was about 3.8 times increased in HDL particles while it was decreased in serum to one-third. Nevertheless, there was a substantial overlap in the miRNA signature between serum and HDL particles in both CRF and hemodialysis patients.

Furthermore, we assessed the differences in the serum and HDL signatures between the two patient cohorts; most of the changes in the miRNA levels were in a similar direction (Figure 2). The miR-122 level increased sharply in the HDL fraction with disease progression ($\Delta RQ \sim +6.6$), while it hardly decreased in the serum ($\Delta RQ \sim -0.7$). On the other hand, some miRNA levels like miR-24, miR574-3p, miR-222, miR-27b and miR-29a were increased in the serum with disease progression, whereas in HDL particles, the levels remained nearly constant. While these miRNA levels were significantly different between HDL particles and serum, a substantial overlap in the miRNA profile exists.

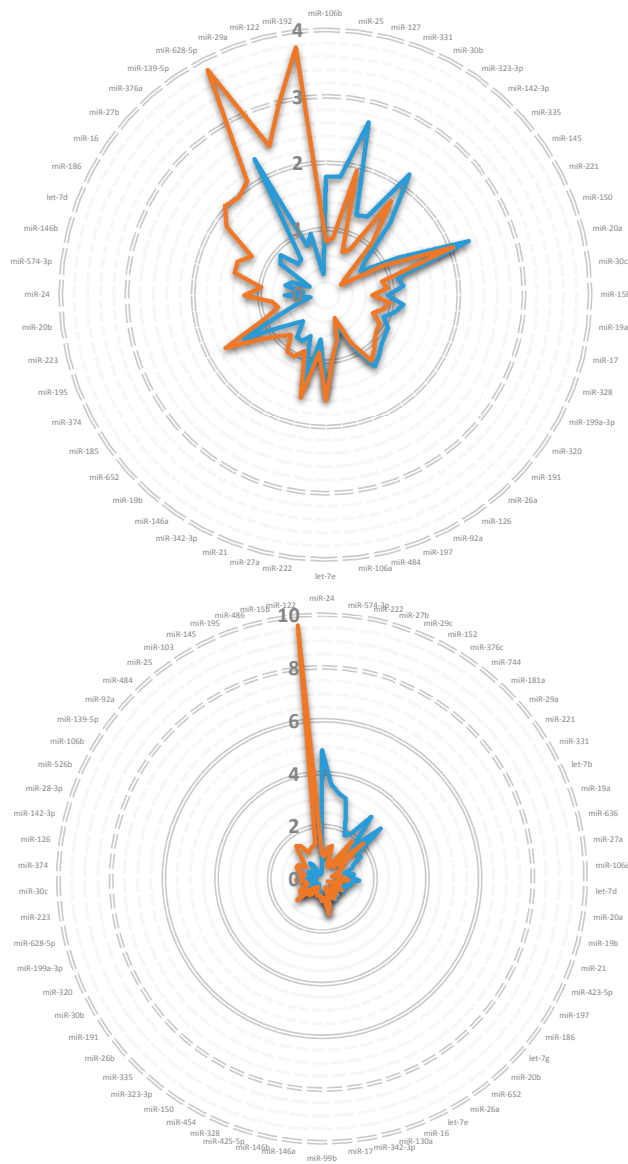


Figure 1. miRNA profile in serum and high-density lipoprotein (HDL) particles of chronic renal failure (CRF) and hemodialysis patients. Pooled serum and HDL particles of adult CRF (top) and hemodialysis patients (bottom) and their matched controls was analyzed by the TaqMan™ miRNA array A (duplicates). From the 381 different miRNAs, 57 (serum) and 90 (HDL particles) for CRF patients and 75 (serum) and 89 (HDL particles) for hemodialysis patients were detectable (plotted were only miRNAs detectable in serum and HDL particles) and the relative quantitation (RQ) was calculated subsequently (the corresponding control miRNAs were set to RQ = 1). RQ-values were arranged in the following manner: starting at the twelve o'clock position with the maximum negative difference between the RQ values of HDL particles (orange) and serum (blue), the values increase clockwise (reaching nearly equality at the six o'clock position) and reach the maximum positive difference before closing the circle again (see Table A2 (in Appendix A) for the complete data set).

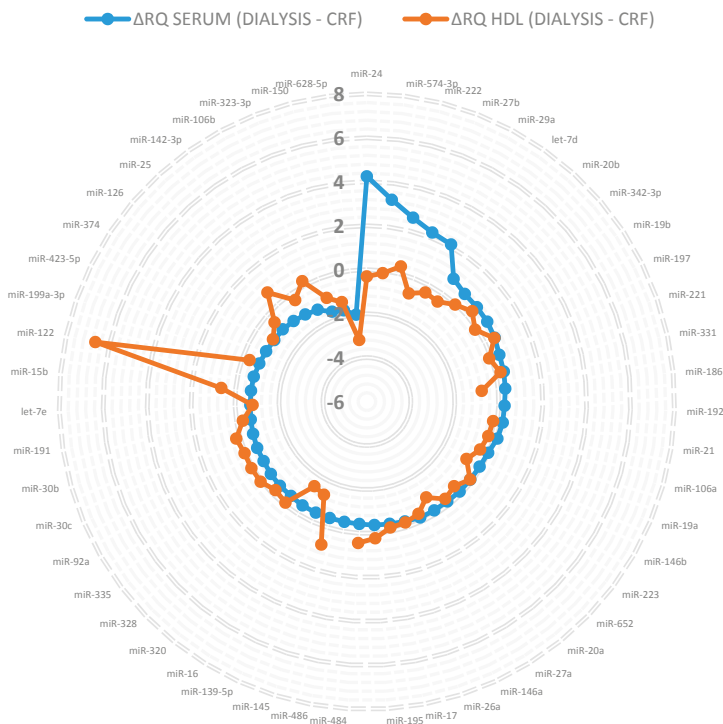


Figure 2. Differences in the miRNA profile in serum and HDL particles during disease progression from CRF to hemodialysis. Data is arranged in the following manner: starting at the twelve o’clock position with the maximum difference between the RQ values of serum (blue) during disease progression from CRF to hemodialysis ($=\Delta RQ$), the values decrease clockwise (reaching nearly no change at the four o’clock position) and reach the maximum negative difference before closing the circle again (plotted are only miRNAs detectable in both disease progression steps). For comparison, ΔRQ values were superimposed for the same miRNAs (if available) in the HDL particle fraction (orange).

3.2. miRNA Associated to Lipoprotein Particles

After having shown that miRNA also circulates with lipoprotein particles and that their expression profile was altered in the disease progression, we assessed if we can reconstitute this system in order to observe and follow cellular miRNA uptake. Therefore, we tested the uptake of native HDL/LDL and reconstituted HDL/labeled LDL particles with two selected miRNAs: miR-223, which shows a high abundance and miR-155, which is rare in lipoprotein particles [41]. To assess the validity of the reconstituted and labeled lipoprotein particles in comparison to their native counterpart, we imaged single lipoprotein particles by HS-AFM and analyzed their size (Figure A1). We observed globular particles comparable to the native forms. However, the reconstituted HDL particle variants were slightly smaller than the native ones; reasonably due to the partial reconstitution process. For LDL particles, the size distributions of the native and labeled particles largely overlap; obviously, spermine addition accounts for the observed size increase. According to our data, the ratio of tested miRNA strands to the number of native lipoprotein particles was in the range of one strand of a specific miRNA to around 1 billion lipoprotein particles (Tables 2 and 3).

Table 2. * miRNA content of HDL particle variants.

Sample	Donor	miRNA	< # miRNA/particle >	SD < # miRNA/particle >
Native HDL	1	223	6.10^{-9}	7.10^{-10}
		155	3.10^{-9}	9.10^{-10}
	2	223	7.10^{-9}	4.10^{-9}
		155	4.10^{-8}	2.10^{-8}
rHDL	1	223	4.10^{-9}	8.10^{-10}
		155	5.10^{-9}	3.10^{-9}
	2	223	7.10^{-9}	4.10^{-9}
		155	1.10^{-8}	1.10^{-8}
rHDL + miR-155	1	155	1.10^{-4}	1.10^{-4}
	2	155	8.10^{-5}	5.10^{-5}
rHDL + miR-155 & spermine	1	155	1.10^{-4}	5.10^{-5}
	2	155	7.10^{-5}	2.10^{-5}

* Each condition was tested twice with two independently obtained samples. rHDL: reconstituted HDL.

Table 3. * miRNA content of LDL particle variants.

Sample	Donor	miRNA	< # miRNA/particle >	SD < # miRNA/particle >
Native LDL	1	223	3.10^{-8}	2.10^{-8}
		155	9.10^{-8}	5.10^{-8}
	2	223	5.10^{-9}	3.10^{-9}
		155	9.10^{-8}	2.10^{-8}
Labeled LDL	1	223	5.10^{-8}	3.10^{-8}
		155	9.10^{-8}	5.10^{-8}
	2	223	4.10^{-9}	3.10^{-9}
		155	9.10^{-8}	2.10^{-8}
Labeled LDL + miR-155	1	155	1.10^{-4}	6.10^{-5}
	2	155	2.10^{-3}	3.10^{-4}
Labeled LDL + miR-155 & spermine	1	155	5.10^{-4}	3.10^{-4}
	2	155	5.10^{-3}	6.10^{-4}

* Each condition was tested twice with two independently obtained samples.

Nevertheless, we were able to increase the miRNA/particle ratio via reconstitution or labeling of the respective lipoprotein particle. The reconstitution (which involves thorough delipidation of the particle) or the labeling process of the lipoprotein particle itself did not influence the miRNA/particle ratio (compare miRNA/particle ratio of native and reconstituted/labeled particles). Thus, we found neither sequence-specificity (no significant difference between miRNA/particle ratio of miR-223 and miR-155), nor any significant effect of electrical-charge-compensation of the poly-anionic nature of the miRNA with addition of the natural poly-cationic nucleotide-sequence-stabilizer-agent spermine. However, our approach increased the miRNA/particle ratio by around five orders of magnitude (from 10^{-9} to 10^{-4} for HDL particles and 10^{-8} to 10^{-3} for LDL particles as shown in Tables 2 and 3).

Next, we tested whether the uptake of lipoprotein particles in cells depends on the cell membrane density of their respective lipoprotein receptor. In order to reach a steady state between uptake and potential degradation of miRNA, we incubated cells for 16 h with lipoprotein particles. We observed no increase of the cellular miRNA-content after incubation with native lipoprotein particles at a concentration in the physiological range ($5 \mu\text{g/mL}$ and $50 \mu\text{g/mL}$ for HDL or LDL particles) independent of the receptor density (Tables 4 and 5). We did neither observe a significant increase in the number of miRNA strands per cell for HDL particles (overexpression of

the HDL-receptor SR-B1 in ldlA7-SR-B1 cells in comparison with native HDL-receptor density in ldlA7 cells) nor for LDL particles (a low LDL-receptor density in ldlA7 cells in comparison with native LDL-receptor density in CHOK1).

However, incubation using lipoprotein particles with artificially increased miRNA content yielded—at least for HDL particles—a significant and receptor-density-dependent increase of the number of miRNA strands per cell (approximately by a factor 4 or 10 at native or overexpressed receptor densities). In contrast, incubation with LDL particles did not influence the miRNA level.

Table 4. * Influence of lipoprotein receptor density on cellular miRNA-content before/after HDL particle uptake.

Cell Line	HDL Conditions and Concentration	< # miRNA/cell >	SD < # miRNA/cell >
ldlA7	no HDL addition	240	35
	native, 50 µg/mL	250	25
	miR-155 and spermine, 50 µg/mL	980	55
ldlA7-SR-B1	no HDL addition	260	20
	native, 50 µg/mL	330	55
	miR-155 and spermine, 50 µg/mL	2500	190

* Each condition was tested twice with two independently obtained samples. The HDL particles with the highest < # miRNA/particle > = 1×10^{-4} were used for these experiments. The used sample volume in the qPCR step equated for both cell lines to 3100 cells.

Table 5. * Influence of lipoprotein particle receptor density on cellular miRNA-content before/after LDL particle uptake.

Cell Line	LDL Conditions and Concentration	< # miR/cell >	SD < # miR/cell >
ldlA7	no LDL addition	190	10
	native, 5 µg/mL	240	20
	miR-155 and spermine, 5 µg/mL	210	40
CHOK1	no LDL addition	120	10
	native, 5 µg/mL	130	20
	miR-155 and spermine, 5 µg/mL	160	70

* Each condition was tested twice with two independently obtained samples. The LDL particles with the highest < # miRNA/particle > = 5×10^{-3} were used for these experiments. The used sample volume in the qPCR step equated for both cell lines to 5800 cells.

In order to exclude the possibility of a non-cell-mediated unspecific interaction of HDL particles with the cell culture vessel itself, we performed control experiments with chambers without cells (see Table A1). Here, the determined numbers for the miRNA amount corresponding to an individual cell (comparable to numbers from Tables 4 and 5) were around 1% and were, therefore, neglected. An incubation-concentration-dependence of the unspecific binding could be seen, and it was further observed that pre-blockage with BSA or native HDL particles had no influence.

4. Discussion

We have shown that, for some miRNAs, there is a significant difference in the profile between the HDL particle fraction and serum derived from CRF as well as hemodialysis patients. For other miRNAs, their values are nearly identical. Moreover, during disease progression, a similar bimodal behavior was observable. Especially miR-24, which has been reported to be upregulated in patients with kidney transplants [42], showed an increase of more than four RQ values in the serum, while its value even decreased in the HDL particle fraction by more than three values. Contrarily, RQ values of miR-122 increased by more than six values during disease progression in the HDL particle fraction while the serum RQ value hardly decreased. Interestingly, miR-21 and miR-223—two miRNAs known to be involved in renal and cardiovascular disease—did not change, neither in serum nor

in the HDL particle fraction. Thus, we postulate that the miRNA profile may be suitable to identify certain diseases and follow their progression [43]; however, one has to be careful regarding its origin, as some miRNA levels differ between serum and HDL particle fraction.

Vickers et al. showed that HDL particles deliver miR-223 to SR-B1–transfected baby hamster kidney (BHK) cells, leading to repression of the Renilla-SR-B1-3′UTR luciferase reporters [6]. Therefore, HDL particles can be expected to transport miRNAs to target cells, leading to altered gene expression. Tabet et al. demonstrated that HDL-transported miR-223 down-regulated ICAM-1 expression in endothelial cells [7]. In the present study, we verified the principal functionality of HDL particles as miRNA-transfer-vehicle. Nevertheless, the number of miRNAs transported via native HDL particles may be insignificant due to the extremely low miRNA/particle ratio. This number may slightly vary depending on the miRNA and individual; however, any relevant influence on the cellular mRNA profile due to an uptake of miRNA-containing lipoprotein particles seems highly unlikely—at least in normolipidemic healthy individuals. We like to note that our numbers (strands/particle) are comparable to data published by Dimmler et al. (10^4 strands/ μg lipoproteins are comparable to 10^{-9} strands/particle) [41].

Moreover, we have estimated that for the observed increase of miRNA strands per cell at standard receptor density, each cell has to interact, on average, with at least 7.5 million HDL particles (assuming a miRNA/particle ratio of $1:10^4$, see Table 1) during the incubation period of 16 h. Under the assumption of no intracellular degradation of miRNA, this value represents only a lower limit. This yields a minimum of roughly 60 uptake events of HDL particles per second—a calculation previously unavailable, which enables a rough estimation of the absolute cargo transfer. According to our measured ratio of 10^{-9} miRNA strands per native HDL particle, on average, it would take more than half a year until a single miRNA strand has been taken up by a single cell. In contrast to HDL-mediated miRNA-uptake, we did not observe a significant increase of the miRNA level using LDL particles—an observation in agreement with the different lysosomal uptake pathway of LDL, which leads to degradation of the cargo itself.

Furthermore, our observation that reconstitution and the associated delipidation step did not influence the miRNA/particle ratio leads us to speculate that miRNA itself is not dominantly lipid-associated, and thus, not associated to the particle surface. Experiments using fluorescently labeled miRNA added to planar supported lipid bilayers (data not shown) yielded no observable signal—independent of the tested conditions (lipid composition, miRNA-concentration, temperature, pH-value, label). Moreover, extra-particle electrostatic association seems unlikely, due to the negative surface potential of lipoprotein particles and the poly-anionic nature of RNA strands. In addition, miRNAs are usually incorporated into a protein of the Argonaute family (AGO1-4) and have been proposed to circulate with HDL [13]. More than 150 proteins were found within the HDL particle fraction using proteome analysis; 95 of these proteins were identified to be specifically bound to the HDL particle. However, neither of the AGO proteins were detected [14–19]. Additionally, using Western Blot analysis and fluorescence imaging, we were unable to detect any AGO-2 protein (unpublished data). Therefore, it seems to be unlikely that this protein family is carrying the miRNA of HDL particles. The protein moiety enabling some of the HDL particles to bind miRNA still remains elusive.

In summary, we conclude that miRNA is transported and transferred to cells via lipoprotein particles, however, most likely plays no relevant role in vivo regarding the alteration of the cellular miRNA, and thus, the mRNA profile. However, lipoprotein particles may serve as diagnostic tools for certain disease-specific miRNAs—as exemplarily demonstrated here by us—or moreover, as an artificial drug delivery system.

Author Contributions: Conceptualization, M.A., S.M.M., W.S., H.S. and B.P.; Methodology, M.A., S.M.M., W.S., H.S. and B.P.; Validation, M.A., H.S. and B.P.; Formal Analysis, M.A., A.K.; Investigation, M.A., S.M.M. and A.K.; Resources, H.S. and B.P.; Writing—Original Draft Preparation, M.A., H.S. and B.P.; Visualization, M.A., A.K. and B.P.; Supervision, H.S. and B.P.; Project Administration, H.S. and B.P.; Funding Acquisition, S.M.M., H.S. and B.P.

Funding: This work was supported by the Austrian Science Fund Project P29110-B21, the “Hochschuljubiläumsstiftung der Stadt Wien zur Förderung der Wissenschaft” Project H-3065/2011, the European Fund for Regional Development (EFRE, IWB2020), the Federal State of Upper Austria and the “Land OÖ Basisfinanzierung”.

Acknowledgments: We thank C. Röhrle and M. Ogris for helpful discussions. We thank J. Strasser for providing the miRNA-image used in the graphical abstract.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Unspecific binding of HDL particles to the chamber.

Chamber Conditions	< # miRNA >	SD < # miRNA >
Blocking with 5% BSA for 1 h + medium for 16 h	2600	5
Blocking with 500 µg/mL native HDL for 1 h + medium for 16 h	2200	90
Blocking with 5% BSA for 1 h + 50 µg/mL rHDL + miR-155 & spermine for 16 h	14500	2800
Blocking with 500 µg/mL native HDL for 1 h + 50 µg/mL rHDL + miR-155 & spermine for 16 h	11000	2800
50 µg/mL rHDL + miR-155 & spermine for 16 h	11000	2100
5 µg/mL rHDL + miR-155 & spermine for 16 h	2800	140

The average number of miRNA < # miRNA > corresponds to the same volume as is contained in the chamber with IdLA7-SR-B1 cells (Table 4). For the authors, the only reasonable way to compare these values to the numbers from Table 4 is to divide them by the number of cells used (there ~3100 cells).

Table A2. RQ Values from Figure 1 of CRF (top) and hemodialysis patients (bottom).

Assay	RQ _{SERUM}	RQ _{HDL}	Assay	RQ _{SERUM}	RQ _{HDL}	Assay	RQ _{SERUM}	RQ _{HDL}
miR-106b	1.78	0.82	miR-19b	0.52	0.80	miR-425-5p	N.D.	0.83
miR-25	1.80	0.87	miR-652	0.64	0.93	miR-148b	N.D.	0.85
miR-127	2.68	1.94	miR-185	0.86	1.16	miR-485-3p	N.D.	0.87
miR-331	1.29	0.70	miR-374	1.39	1.71	miR-99b	N.D.	0.88
miR-30b	1.34	0.78	miR-195	0.62	0.97	miR-28-3p	N.D.	0.96
miR-323-3p	2.21	1.73	miR-223	0.38	0.74	miR-376c	N.D.	0.97
miR-142-3p	1.46	1.04	miR-20b	0.23	0.82	let-7b	N.D.	1.12
miR-335	0.79	0.39	miR-24	0.62	1.23	let-7g	N.D.	1.34
miR-145	0.63	0.27	miR-574-3p	0.35	0.98	miR-29c	N.D.	1.39
miR-221	1.22	0.95	miR-146b	0.63	1.41	miR-26b	N.D.	1.45
miR-150	2.30	2.05	let-7d	0.55	1.42	miR-155	N.D.	1.66
miR-20a	1.11	0.87	miR-186	0.28	1.26	miR-494	N.D.	1.85
miR-30c	1.17	0.94	miR-16	0.83	1.82	miR-152	N.D.	1.97
miR-15b	0.93	0.70	miR-27b	0.91	2.03	miR-454	N.D.	2.05
miR-19a	1.17	0.97	miR-376a	0.63	1.97	miR-133a	N.D.	2.12
miR-17	1.07	0.88	miR-139-5p	0.65	2.09	miR-539	N.D.	2.37
miR-328	0.93	0.75	miR-628-5p	2.32	3.83	miR-148a	N.D.	2.85
miR-199a-3p	1.06	0.91	miR-29a	0.80	2.40	miR-518f	N.D.	2.99
miR-320	1.07	0.93	miR-122	0.96	2.97	miR-627	N.D.	3.02
miR-191	1.12	0.98	miR-192	0.32	3.76	miR-636	N.D.	3.73
miR-26a	1.20	1.09	miR-142-5p	N.D.	0.35	miR-301b	N.D.	3.83
miR-126	1.30	1.20	miR-526b	N.D.	0.38	miR-107	N.D.	5.30
miR-92a	0.90	0.81	miR-103	N.D.	0.45	miR-410	N.D.	5.77
miR-197	0.46	0.37	miR-744	N.D.	0.45	miR-424	N.D.	6.91
miR-484	0.75	0.71	miR-340	N.D.	0.48	miR-487b	N.D.	6.92
miR-106a	0.92	0.90	let-7a	N.D.	0.50	let-7c	N.D.	13.26
let-7e	1.53	1.58	miR-130b	N.D.	0.56	miR-486	0.61	N.D.
miR-222	0.67	0.87	miR-130a	N.D.	0.58	miR-532	1.11	N.D.
miR-27a	1.38	1.59	miR-301	N.D.	0.65	miR-375	1.20	N.D.
miR-21	0.66	0.90	miR-128a	N.D.	0.72	miR-423-5p	1.24	N.D.
miR-342-3p	0.78	1.04	miR-28	N.D.	0.75	miR-212	1.63	N.D.
miR-146a	0.77	1.03	miR-345	N.D.	0.82			

Table A2. Cont.

Assay	RQ _{SERUM}	RQ _{HDL}	Assay	RQ _{SERUM}	RQ _{HDL}	Assay	RQ _{SERUM}	RQ _{HDL}
miR-24	4.86	0.92	miR-99b	0.82	0.84	miR-483-5p	0.43	N.D.
miR-574-3p	3.59	0.86	miR-146a	0.58	0.65	miR-192	0.59	N.D.
miR-222	3.29	1.20	miR-146b	0.55	0.65	miR-99a	2.80	N.D.
miR-27b	3.16	1.30	miR-425-5p	0.56	0.66	miR-660	3.94	N.D.
miR-29c	2.36	0.52	miR-328	0.44	0.54	miR-193a-5p	6.52	N.D.
miR-152	1.84	0.50	miR-454	0.23	0.34	miR-125b	16.31	N.D.
miR-376c	2.05	0.73	miR-150	0.57	0.71	miR-193b	17.65	N.D.
miR-744	2.98	1.81	miR-323-3p	0.59	0.78	miR-200c	90.15	N.D.
miR-181a	1.67	0.61	miR-335	0.26	0.45	miR-203	1626.14	N.D.
miR-29a	2.90	2.03	miR-26b	1.03	1.24	miR-532	N.D.	0.39
miR-221	1.61	0.86	miR-191	0.46	0.69	miR-370	N.D.	0.47
miR-331	1.67	0.92	miR-30b	0.72	0.96	miR-339-3p	N.D.	0.52
let-7b	1.30	0.68	miR-320	0.58	0.83	miR-301	N.D.	0.54
miR-19a	1.17	0.56	miR-199a-3p	0.24	0.57	miR-125a-5p	N.D.	0.54
miR-636	0.87	0.37	miR-628-5p	0.29	0.65	miR-148b	N.D.	0.55
miR-27a	1.20	0.72	miR-223	0.27	0.63	miR-590-5p	N.D.	0.58
miR-106a	1.09	0.64	miR-30c	0.59	0.98	miR-142-5p	N.D.	0.62
let-7d	1.38	0.99	miR-374	0.45	0.83	miR-597	N.D.	0.64
miR-20a	0.94	0.55	miR-126	0.34	0.73	miR-133a	N.D.	0.65
miR-19b	1.09	0.71	miR-142-3p	0.30	0.69	miR-539	N.D.	0.70
miR-21	0.91	0.71	miR-28-3p	0.52	0.93	let-7a	N.D.	0.73
miR-423-5p	0.36	0.23	miR-526b	0.57	1.05	miR-127	N.D.	0.79
miR-197	0.95	0.85	miR-106b	0.52	1.03	miR-18a	N.D.	0.87
miR-186	0.60	0.51	miR-139-5p	0.22	0.76	miR-411	N.D.	0.92
let-7g	0.68	0.60	miR-92a	0.32	0.88	miR-324-3p	N.D.	1.10
miR-20b	0.85	0.78	miR-484	0.34	1.15	miR-185	N.D.	1.28
miR-652	0.53	0.46	miR-25	0.75	1.58	miR-148a	N.D.	1.31
miR-26a	0.93	0.87	miR-103	0.63	1.46	miR-130b	N.D.	1.35
let-7e	0.83	0.78	miR-145	0.20	1.11	miR-155	N.D.	1.65
miR-16	0.39	0.35	miR-195	0.25	1.20	miR-132	N.D.	2.02
miR-130a	1.07	1.04	miR-486	0.18	1.28	miR-487b	N.D.	2.11
miR-342-3p	1.37	1.35	miR-15b	0.23	1.35	miR-28	N.D.	2.22
miR-17	0.73	0.71	miR-122	0.23	9.61			

Arrangement of data is equal to the clockwise order shown in Figure 1; for miRNA data with only one available value (not shown in Figure 1), the miRNAs are arranged in an ascending order.

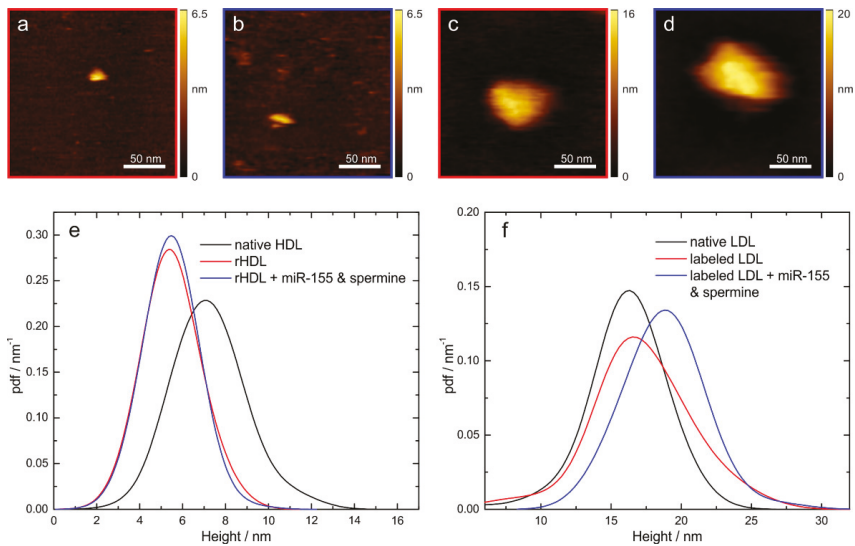


Figure A1. Spatial characterization of reconstituted/labeled lipoprotein particles. High-speed atomic force microscopy (HS-AFM) images of reconstituted HDL particles (a), rHDL particles with miR-155 & spermine (b), labeled LDL particles (c) and labeled LDL particles with miR-155 & spermine (d). The size analysis yielded distributions of particle heights for native HDL particles (N = 125 particles, $\langle \text{height} \rangle = 7.3 \text{ nm} \pm 1.5 \text{ nm}$), rHDL particles (N=110, $\langle \text{height} \rangle = 5.5 \text{ nm} \pm 1.0 \text{ nm}$) and rHDL particles + miR-155 & spermine (N = 122, $\langle \text{height} \rangle = 5.5 \text{ nm} \pm 0.9 \text{ nm}$) (e), as well as for native LDL particles (N = 79, $\langle \text{height} \rangle = 16.1 \text{ nm} \pm 2.4 \text{ nm}$), labeled LDL particles (N = 89, $\langle \text{height} \rangle = 17.2 \text{ nm} \pm 3.4 \text{ nm}$) and LDL particles + miR-155 & spermine (N = 57, $\langle \text{height} \rangle = 18.8 \text{ nm} \pm 2.4 \text{ nm}$) (f).

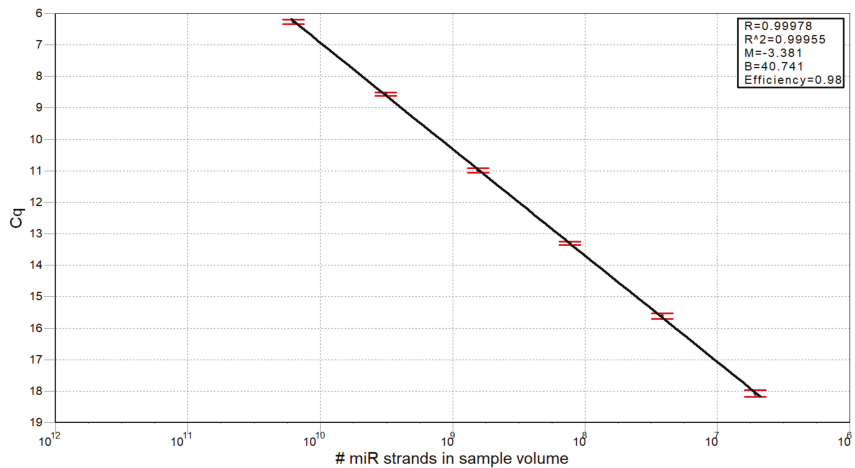


Figure A2. Cont.

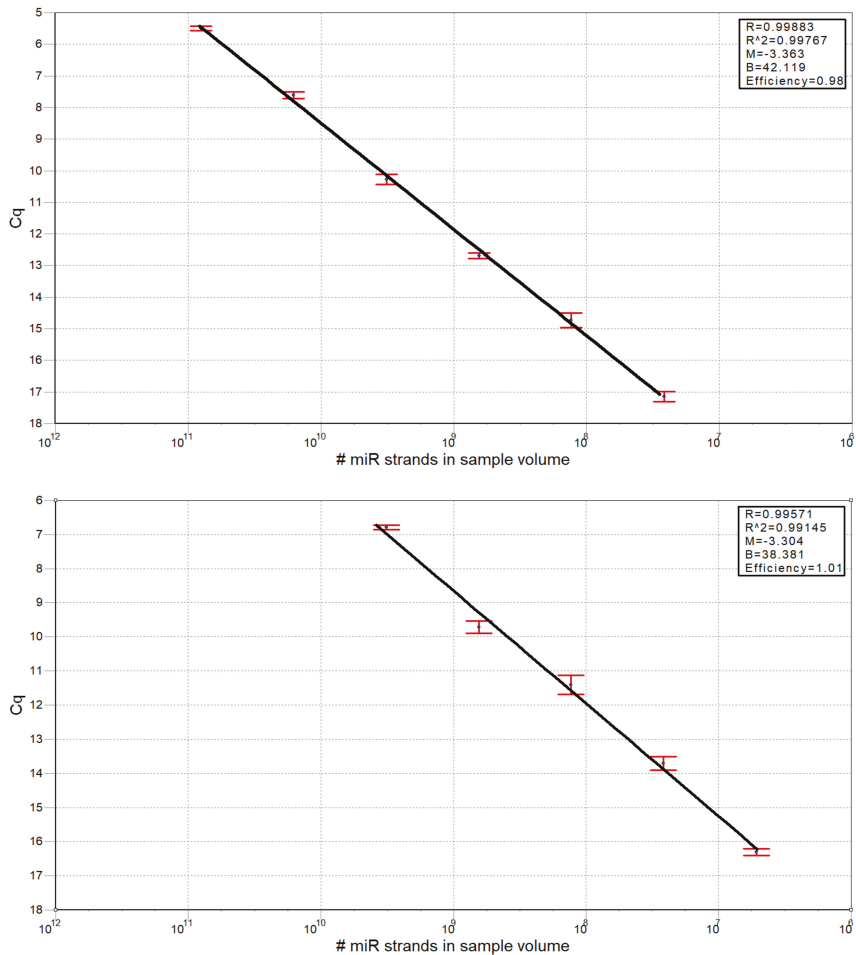


Figure A2. Standard curves for miR-145 (top), miR-155 (center) and miR-223 (bottom).

References

1. Simons, M.; Raposo, G. Exosomes—Vesicular carriers for intercellular communication. *Curr. Opin. Cell Biol.* **2009**, *21*, 575–581. [[CrossRef](#)] [[PubMed](#)]
2. Valadi, H.; Ekstrom, K.; Bossios, A.; Sjostrand, M.; Lee, J.J.; Lotvall, J.O. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **2007**, *9*, 654–659. [[CrossRef](#)] [[PubMed](#)]
3. Hunter, M.P.; Ismail, N.; Zhang, X.; Aguda, B.D.; Lee, E.J.; Yu, L.; Xiao, T.; Schafer, J.; Lee, M.L.; Schmittgen, T.D.; et al. Detection of microRNA expression in human peripheral blood microvesicles. *PLoS ONE* **2008**, *3*, e3694. [[CrossRef](#)] [[PubMed](#)]
4. Ratajczak, J.; Wysoczynski, M.; Hayek, F.; Janowska-Wieczorek, A.; Ratajczak, M.Z. Membrane-derived microvesicles: Important and underappreciated mediators of cell-to-cell communication. *Leukemia* **2006**, *20*, 1487–1495. [[CrossRef](#)] [[PubMed](#)]
5. Arroyo, J.D.; Chevillet, J.R.; Kroh, E.M.; Ruf, I.K.; Pritchard, C.C.; Gibson, D.F.; Mitchell, P.S.; Bennett, C.F.; Pogosova-Agadjanyan, E.L.; Stirewalt, D.L.; et al. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 5003–5008. [[CrossRef](#)] [[PubMed](#)]

6. Vickers, K.C.; Palmisano, B.T.; Shoucri, B.M.; Shamburek, R.D.; Remaley, A.T. MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nat. Cell Biol.* **2011**, *13*, 423–433. [[CrossRef](#)] [[PubMed](#)]
7. Tabet, F.; Vickers, K.C.; Cuesta Torres, L.F.; Wiese, C.B.; Shoucri, B.M.; Lambert, G.; Catherinet, C.; Prado-Lourenco, L.; Levin, M.G.; Thacker, S.; et al. Hdl-transferred microRNA-223 regulates icam-1 expression in endothelial cells. *Nat. Commun.* **2014**, *5*, 3292. [[CrossRef](#)] [[PubMed](#)]
8. Wahid, F.; Shehzad, A.; Khan, T.; Kim, Y.Y. MicroRNAs: Synthesis, mechanism, function, and recent clinical trials. *Biochim. Biophys. Acta* **2010**, *1803*, 1231–1243. [[CrossRef](#)] [[PubMed](#)]
9. Carthew, R.W.; Sontheimer, E.J. Origins and mechanisms of miRNAs and siRNAs. *Cell* **2009**, *136*, 642–655. [[CrossRef](#)] [[PubMed](#)]
10. Fabian, M.R.; Sonenberg, N.; Filipowicz, W. Regulation of mRNA translation and stability by micrnas. *Ann. Rev. Biochem.* **2010**, *79*, 351–379. [[CrossRef](#)] [[PubMed](#)]
11. Filipowicz, W.; Bhattacharyya, S.N.; Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? *Nat. Rev. Genet.* **2008**, *9*, 102–114. [[CrossRef](#)] [[PubMed](#)]
12. Chen, X.; Liang, H.; Zhang, J.; Zen, K.; Zhang, C.Y. Secreted microRNAs: A new form of intercellular communication. *Trends Cell Biol.* **2012**, *22*, 125–132. [[CrossRef](#)] [[PubMed](#)]
13. Turchinovich, A.; Samatov, T.R.; Tonevitsky, A.G.; Burwinkel, B. Circulating miRNAs: Cell-cell communication function? *Front. Genet.* **2013**, *4*, 119. [[CrossRef](#)] [[PubMed](#)]
14. Vaisar, T.; Pennathur, S.; Green, P.S.; Gharib, S.A.; Hoofnagle, A.N.; Cheung, M.C.; Byun, J.; Vuletic, S.; Kassim, S.; Singh, P.; et al. Shotgun proteomics implicates protease inhibition and complement activation in the antiinflammatory properties of HDL. *J. Clin. Investig.* **2007**, *117*, 746–756. [[CrossRef](#)] [[PubMed](#)]
15. Weichhart, T.; Kopecky, C.; Kubicek, M.; Haidinger, M.; Doller, D.; Katholnig, K.; Suarna, C.; Eller, P.; Tolle, M.; Gerner, C.; et al. Serum amyloid a in uremic HDL promotes inflammation. *J. Am. Soc. Nephrol.* **2012**, *23*, 934–947. [[CrossRef](#)] [[PubMed](#)]
16. Holzer, M.; Wolf, P.; Curcic, S.; Birner-Gruenberger, R.; Weger, W.; Inzinger, M.; El-Gamal, D.; Wadsack, C.; Heinemann, A.; Marsche, G. Psoriasis alters HDL composition and cholesterol efflux capacity. *J. Lipid Res.* **2012**, *53*, 1618–1624. [[CrossRef](#)] [[PubMed](#)]
17. Holzer, M.; Birner-Gruenberger, R.; Stojakovic, T.; El-Gamal, D.; Binder, V.; Wadsack, C.; Heinemann, A.; Marsche, G. Uremia alters HDL composition and function. *J. Am. Soc. Nephrol.* **2011**, *22*, 1631–1641. [[CrossRef](#)] [[PubMed](#)]
18. Singh, S.A.; Andraski, A.B.; Pieper, B.; Goh, W.; Mendivil, C.O.; Sacks, F.M.; Aikawa, M. Multiple apolipoprotein kinetics measured in human HDL by high-resolution/accurate mass parallel reaction monitoring. *J. Lipid Res.* **2016**, *57*, 714–728. [[CrossRef](#)] [[PubMed](#)]
19. Singh, S.A.; Aikawa, M. Unbiased and targeted mass spectrometry for the HDL proteome. *Curr. Opin. Lipidol.* **2017**, *28*, 68–77. [[CrossRef](#)] [[PubMed](#)]
20. Ha, T.Y. MicroRNAs in human diseases: From cancer to cardiovascular disease. *Immune Netw.* **2011**, *11*, 135–154. [[CrossRef](#)] [[PubMed](#)]
21. Jamaluddin, M.S.; Weakley, S.M.; Zhang, L.; Kougias, P.; Lin, P.H.; Yao, Q.; Chen, C. miRNAs: Roles and clinical applications in vascular disease. *Expert Rev. Mol. Diagn.* **2011**, *11*, 79–89. [[CrossRef](#)] [[PubMed](#)]
22. Kosaka, N.; Iguchi, H.; Yoshioka, Y.; Takeshita, F.; Matsuki, Y.; Ochiya, T. Secretory mechanisms and intercellular transfer of microRNAs in living cells. *J. Biol. Chem.* **2010**, *285*, 17442–17452. [[CrossRef](#)] [[PubMed](#)]
23. Qin, S.; Zhang, C. MicroRNAs in vascular disease. *J. Cardiovasc. Pharm.* **2011**, *57*, 8–12. [[CrossRef](#)] [[PubMed](#)]
24. Vickers, K.C.; Remaley, A.T. MicroRNAs in atherosclerosis and lipoprotein metabolism. *Curr. Opin. Endocrinol.* **2010**, *17*, 150–155. [[CrossRef](#)] [[PubMed](#)]
25. Akkina, S.; Becker, B.N. MicroRNAs in kidney function and disease. *Transl. Res.* **2011**, *157*, 236–240. [[CrossRef](#)] [[PubMed](#)]
26. Bhatt, K.; Mi, Q.S.; Dong, Z. MicroRNAs in kidneys: Biogenesis, regulation, and pathophysiological roles. *Am. J. Physiol. Renal Physiol.* **2011**, *300*, F602–F610. [[CrossRef](#)] [[PubMed](#)]
27. Li, J.Y.; Yong, T.Y.; Michael, M.Z.; Gleadle, J.M. Review: The role of microRNAs in kidney disease. *Nephrology* **2010**, *15*, 599–608. [[CrossRef](#)] [[PubMed](#)]

28. Neal, C.S.; Michael, M.Z.; Pimlott, L.K.; Yong, T.Y.; Li, J.Y.; Gleadle, J.M. Circulating microRNA expression is reduced in chronic kidney disease. *Nephrol. Dial. Transplant.* **2011**, *26*, 3794–3802. [[CrossRef](#)] [[PubMed](#)]
29. Mitchell, P.S.; Parkin, R.K.; Kroh, E.M.; Fritz, B.R.; Wyman, S.K.; Pogosova-Agadjanyan, E.L.; Peterson, A.; Noteboom, J.; O'Briant, K.C.; Allen, A.; et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10513–10518. [[CrossRef](#)] [[PubMed](#)]
30. Olson, E.N. MicroRNAs as therapeutic targets and biomarkers of cardiovascular disease. *Sci. Transl. Med.* **2014**, *6*, 239ps3. [[CrossRef](#)] [[PubMed](#)]
31. Bustin, S.A.; Benes, V.; Garson, J.A.; Hellemans, J.; Huggett, J.; Kubista, M.; Mueller, R.; Nolan, T.; Pfaffl, M.W.; Shipley, G.L.; et al. The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* **2009**, *55*, 611–622. [[CrossRef](#)] [[PubMed](#)]
32. Meier, S.M.; Wultsch, A.; Hollaus, M.; Ammann, M.; Pemberger, E.; Liebscher, F.; Lambers, B.; Fruhwurth, S.; Stojakovic, T.; Scharnagl, H.; et al. Effect of chronic kidney disease on macrophage cholesterol efflux. *Life Sci.* **2015**, *136*, 1–6. [[CrossRef](#)] [[PubMed](#)]
33. Fichtlscherer, S.; De Rosa, S.; Fox, H.; Schwietz, T.; Fischer, A.; Liebetrau, C.; Weber, M.; Hamm, C.W.; Roxe, T.; Muller-Ardogan, M.; et al. Circulating microRNAs in patients with coronary artery disease. *Circ. Res.* **2010**, *107*, 677–684. [[CrossRef](#)] [[PubMed](#)]
34. Schumaker, V.N.; Puppione, D.L. Sequential flotation ultracentrifugation. *Method Enzymol.* **1986**, *128*, 155–170.
35. Jonas, A. Reconstitution of high-density lipoproteins. *Method Enzymol.* **1986**, *128*, 553–582.
36. Krieger, M. Reconstitution of the hydrophobic core of low-density lipoprotein. *Method Enzymol.* **1986**, *128*, 608–613.
37. Stangl, H.; Cao, G.; Wyne, K.L.; Hobbs, H.H. Scavenger receptor, class b, type I-dependent stimulation of cholesterol esterification by high density lipoproteins, low density lipoproteins, and nonlipoprotein cholesterol. *J. Biol. Chem.* **1998**, *273*, 31002–31008. [[CrossRef](#)] [[PubMed](#)]
38. Stangl, H.; Hyatt, M.; Hobbs, H.H. Transport of lipids from high- and low-density lipoproteins via scavenger receptor-bi. *J. Biol. Chem.* **1999**, *274*, 32692–32698. [[CrossRef](#)] [[PubMed](#)]
39. Zhou, T.B.; Jiang, Z.P. Role of miR-21 and its signaling pathways in renal diseases. *J. Recept. Signal Transduct.* **2014**, *34*, 335–337. [[CrossRef](#)] [[PubMed](#)]
40. Yuan, X.; Wang, X.; Chen, C.; Zhou, J.; Han, M. Bone mesenchymal stem cells ameliorate ischemia/reperfusion-induced damage in renal epithelial cells via microRNA-223. *Stem Cell Res. Ther.* **2017**, *8*, 146. [[CrossRef](#)] [[PubMed](#)]
41. Wagner, J.; Riwanto, M.; Besler, C.; Knau, A.; Fichtlscherer, S.; Roxe, T.; Zeiher, A.M.; Landmesser, U.; Dimmeler, S. Characterization of levels and cellular transfer of circulating lipoprotein-bound microRNAs. *Arterioscler. Thromb. Vasc.* **2013**, *33*, 1392–1400. [[CrossRef](#)] [[PubMed](#)]
42. Lorenzen, J.M.; Kaucsar, T.; Schauerte, C.; Schmitt, R.; Rong, S.; Hubner, A.; Scherf, K.; Fiedler, J.; Martino, F.; Kumarswamy, R.; et al. MicroRNA-24 antagonism prevents renal ischemia reperfusion injury. *J. Am. Soc. Nephrol.* **2014**, *25*, 2717–2729. [[CrossRef](#)] [[PubMed](#)]
43. Desgagne, V.; Guerin, R.; Guay, S.P.; Corbin, F.; Couture, P.; Lamarche, B.; Bouchard, L. Changes in high-density lipoprotein-carried miRNA contribution to the plasmatic pool after consumption of dietary trans fat in healthy men. *Epigenomics* **2017**, *9*, 669–688. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

A Two-Cohort RNA-seq Study Reveals Changes in Endometrial and Blood miRNome in Fertile and Infertile Women

Kadri Rekker ^{1,2,*†}, Signe Altmäe ^{2,3,†}, Marina Suhorutshenko ^{1,2}, Maire Peters ^{1,2}, Juan F. Martinez-Blanch ⁴, Francisco M. Codoñer ⁴, Felipe Vilella ^{5,6}, Carlos Simón ^{5,6,7}, Andres Salumets ^{1,2,8,9} and Agne Velthut-Meikas ^{2,10}

¹ Institute of Clinical Medicine, Department of Obstetrics and Gynecology, University of Tartu, 50406 Tartu, Estonia; marina.suhorutsenko@ut.ee (M.S.); maire.peters@ut.ee (M.P.); andres.salumets@ccht.ee (A.S.)

² Competence Centre on Health Technologies, 50410 Tartu, Estonia; signealtmae@ugr.es (S.A.); agnevelthut@gmail.com (A.V.-M.)

³ Department of Biochemistry and Molecular Biology, Faculty of Sciences, University of Granada, 18071 Granada, Spain

⁴ Lifesequencing SL, 46980 Valencia, Spain; juan.martinez@lifesequencing.com (J.F.M.-B.); francisco.codoner@adm.com (F.M.C.)

⁵ Igenomix Foundation/INCLIVA, 46010 Valencia, Spain; felipe.vilella@igenomix.com (F.V.); carlos.simon@igenomix.com (C.S.)

⁶ Igenomix SL, 46980 Valencia, Spain

⁷ Department of Obstetrics and Gynecology, Valencia University, 46010 Valencia, Spain

⁸ Institute of Biomedicine and Translational Medicine, Department of Biomedicine, University of Tartu, 50412 Tartu, Estonia

⁹ Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital, FI-00014 Helsinki, Finland

¹⁰ Department of Chemistry and Biotechnology, Tallinn University of Technology, 12618 Tallinn, Estonia

* Correspondence: kadri.rekker@gmail.com; Tel.: +372-733-0402

† These authors contributed equally to this work.

Received: 25 September 2018; Accepted: 7 November 2018; Published: 23 November 2018

Abstract: The endometrium undergoes extensive changes to prepare for embryo implantation and microRNAs (miRNAs) have been described as playing a significant role in the regulation of endometrial receptivity. However, there is no consensus about the miRNAs involved in mid-secretory endometrial functions. We analysed the complete endometrial miRNome from early secretory (pre-receptive) and mid-secretory (receptive) phases from fertile women and from patients with recurrent implantation failure (RIF) to reveal differentially expressed (DE) miRNAs in the mid-secretory endometrium. Furthermore, we investigated whether the overall changes during early to mid-secretory phase transition and with RIF condition could be reflected in blood miRNA profiles. In total, 116 endometrial and 114 matched blood samples collected from two different population cohorts were subjected to small RNA sequencing. Among fertile women, 91 DE miRNAs were identified in the mid-secretory vs. early secretory endometrium, while no differences were found in the corresponding blood samples. The comparison of mid-secretory phase samples between fertile and infertile women revealed 21 DE miRNAs from the endometrium and one from blood samples. Among discovered novel miRNAs, chr2_4401 was validated and showed up-regulation in the mid-secretory endometrium. Besides novel findings, we confirmed the involvement of miR-30 and miR-200 family members in mid-secretory endometrial functions.

Keywords: endometrial receptivity; infertility; microRNA; recurrent implantation failure; small RNA-seq

1. Introduction

The establishment of a receptive endometrium is essential for successful embryo implantation. The endometrium is receptive within a limited time frame during the mid-secretory cycle phase and aberrations in the processes involved in transition to this stage can lead to infertility [1,2]. Indeed, impaired endometrial receptivity is suspected to play a major role in female infertility among women suffering from recurrent implantation failure (RIF). Gaining insights into the complex mechanisms controlling changes within the endometrium is crucial to understanding not only embryo implantation but also endometrial dysfunction that can lead to infertility. Although hundreds of simultaneously up- and down-regulated genes have been implicated in the processes of acquiring endometrial receptivity [3,4] and the development of RIF [5–10], the precise molecular mechanisms regulating gene expression necessary for endometrial function in fertility and infertility-associated diseases are not well understood.

MicroRNAs (miRNAs), non-coding RNAs (ncRNAs) of ~22 nucleotides in length, act as post-transcriptional regulators of gene expression by inhibiting the stability or repressing the translation of their target messenger RNA (mRNA) molecules [11]. To date, more than 2500 annotated miRNAs in the human genome are known, and as each miRNA may regulate hundreds of genes, it is estimated that miRNAs collectively adjust the expression of over one third of genes in the human genome [11,12]. Hence, miRNAs orchestrate a large variety of processes, including the cyclic changes in the female reproductive tract [13] and cellular processes involved in implantation, such as cellular differentiation, proliferation and apoptosis [14].

The involvement of miRNAs in the mid-secretory endometrial functions has been demonstrated [15–17] and aberrant miRNA profiles in RIF have been identified in a few previous microarray-based [18–21] and two RNA-sequencing (RNA-seq) based studies [22,23]. However, as the analysis settings and platforms have been different and only a small number of samples have been analysed, there is no consensus about the list of miRNAs involved in endometrial receptivity and in endometrial dysfunction(s), advocating further investigations.

Therefore, we set out to analyse the complete spectrum of endometrial miRNAs—miRNome from paired early secretory (ES) and mid-secretory (MS) phase samples collected from two independent population cohorts using the comprehensive RNA-seq technology. Additionally, MS endometrial samples from RIF patients were included in order to identify the dysregulated miRNAs in infertility. Furthermore, we investigated whether the overall cyclic changes in the female body during ES to MS transition and with RIF condition could also be reflected in blood miRNA profiles.

2. Materials and Methods

2.1. Study Participants

The study was approved by the Research Ethics Committee of the University of Tartu, Estonia (No 221/M-31), and the Ethical Clinical Research Committee of IVI Clinic, Valencia, Spain (No 1201-C-094-CS). Informed consent was signed by all women who entered the study and all research was performed in accordance with relevant guidelines and regulations.

The study participants were recruited and samples were collected by two independent research teams from two countries: Estonia (EST) and Spain (ESP). All protocols were standardized across the study centres with exceptions highlighted in the “Sample collection” paragraph.

The healthy fertile group consisted in total of 39 women: EST $n = 22$, age 30.2 ± 3.3 years (mean \pm standard deviation), body mass index (BMI) 23.1 ± 4.2 kg/m²; and ESP $n = 17$, age 29.4 ± 3.6 years, BMI 23.2 ± 2.8 kg/m²; with self-reported regular menstrual cycles. All women had at least one live-born child (1.5 ± 1.0 for EST cohort; 1.2 ± 0.4 for ESP cohort) within the last 10 years (5.3 ± 2.6 years, data available for the EST cohort) and no previous infertility records.

RIF patient group consisted in total of 38 women: EST $n = 21$, age 35.1 ± 3.9 years, BMI 22.3 ± 2.3 kg/m²; ESP $n = 17$, age 36.7 ± 3.3 years, BMI 25.2 ± 5.3 kg/m². All RIF patients had undergone at

least three (4.0 ± 1.5 for EST cohort; 3.0 ± 1.7 for ESP cohort) unsuccessful in vitro fertilisation (IVF) treatment cycles with embryo transfers. No donor eggs were used in the treatments. RIF patients were not undergoing hormonal stimulation for IVF at the time of sample collection. Among these patients, tubal and male factor were the main reasons for infertility along with endometriosis and unexplained infertility. Two participants were suffering from secondary infertility with the experience of childbirth in average 6.5 ± 0.5 years ago. None of the participants had received hormonal treatments for at least three months prior to the time of sample collection.

2.2. Sample Collection

All women performed urine-based ovulation tests to determine the luteinizing hormone (LH) surge using commercial kits (BabyTime® hLH urine cassette, Pharmanova, Beit Shemesh, Israel). The day with positive outcome of the ovulation test is referred to as LH + 0. Healthy women donated endometrium and blood samples during two time-points of the same menstrual cycle: LH + 1 to LH + 3 corresponding to the ES phase and LH + 7 to LH + 9 corresponding to the MS phase. Women in the RIF group provided samples at MS phase (LH + 7 to LH + 9).

Endometrial tissue was obtained using Pipelle catheter (Laboratoire CCD, Paris, France). The biopsy was divided into two parts: one part was rinsed in sterile phosphate-buffered saline (PBS) and frozen at $-80\text{ }^{\circ}\text{C}$ in RNAlater solution (Thermo Fisher Scientific, Waltham, MA, USA) in EST or without any additives in ESP, the other portion was placed in 10% buffered formalin for histological evaluation. Histological dating was conducted to confirm the endometrial phase. In addition, the receptivity status of all endometrial biopsies was assessed and confirmed by analysing a set of 57 receptivity biomarkers as part of our previous studies [4,24].

Blood samples were collected and processed differently by the two research centres. In EST, whole blood samples were collected into PAXgene Blood RNA Tubes (Qiagen, Hilden, Germany), incubated at room temperature for 2 h and then frozen until further use. In ESP, blood samples were collected with tubes containing $\text{K}_2\text{-EDTA}$ anticoagulant, the buffy coat fractions containing leukocytes and platelets were separated by a centrifugation procedure with Histopaque-1077 (Sigma-Aldrich, St. Louis, MO, USA), and were frozen until further use.

2.3. RNA Extraction from Endometrial Tissue

Up to 30 mg of the endometrial tissue was processed using miRNeasy Mini and RNeasy MinElute kits (Qiagen), following the manufacturer's protocol for isolating small RNA (<200 nucleotides) separately from large RNA molecules. DNase I treatment was performed on column using RNase-Free DNase Set (Qiagen). Purified RNA quantity was determined with Bioanalyzer 2100 Small RNA kit (Agilent Technologies, Santa Clara, CA, USA).

2.4. RNA Extraction from Blood Samples

In EST, RNA was extracted from the whole blood samples using PAXgene Blood miRNA Kit (Qiagen) according to manufacturer's instructions. In ESP, total RNA was isolated from buffy coat (leukocytes and platelets) using the miRNeasy Mini Kit and then cleaned with RNA Cleanup kit (Qiagen). On-column DNase I treatment was performed with RNase-Free DNase Set (Qiagen) for all samples.

2.5. Small RNA Sequencing

A total of 116 endometrial samples (EST $n = 65$; ESP $n = 51$) and 114 blood samples (EST $n = 65$; ESP $n = 49$) were subjected to small RNA sequencing in the current study (Figure 1). Small RNA libraries were constructed following the TruSeq Small RNA Library Preparation Guide (Illumina, San Diego, CA, USA). As input, $1\text{ }\mu\text{g}$ of total RNA or small RNA fraction was used; 12-plexing of samples was performed by inserting Illumina index sequences into each library. Final purification step was performed by manually selecting libraries corresponding to the length of inserted miRNAs (area

between 145–160 bp) using gel electrophoresis (6% Novex TBE gels). Libraries were quantified and validated with Agilent 2100 Bioanalyzer (Agilent Technologies), normalized, pooled, used for cluster generation on the cBot and sequenced on HiSeq 2000/2500 with a configuration of 50 cycles Single Reads following manufacturer’s instructions (Illumina) at Lifesequencing S.L., Valencia, Spain (ESP samples) or at Estonian Genome Center Core Facility, Tartu, Estonia (EST samples).

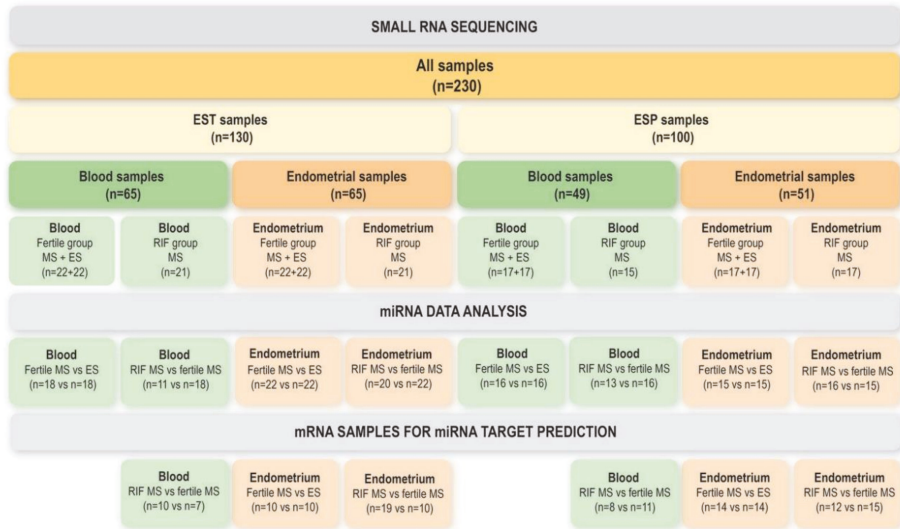


Figure 1. Number of samples included to small RNA sequencing, to final microRNA (miRNA) analysis and to messenger RNA (mRNA) sequencing analysis. EST—Estonian samples; ESP—Spanish samples; RIF—recurrent implantation failure; ES—early secretory phase; MS—mid-secretory phase.

2.6. SMALL RNA Sequencing Data Analysis

Small RNA sequencing data were deposited into the Gene Expression Omnibus (GEO accession number GSE108966). Quality control and adapter trimming of raw sequences was performed using Trimmomatic version 0.36 [25]. All sequences were scanned with 4-base long sliding window and reads were dropped if the average quality of base pair was below 30 according to the Phred +33 quality score system or if total read length was less than 17.

Identification of miRNAs from sequencing reads was performed by the miRDeep2 algorithm [26]. Firstly, mapper.pl algorithm was used to map the trimmed high-quality sequences to the UCSC human genome version hg19. Common reads were collapsed and passed as input to miRDeep2.pl algorithm together with mature and stem-loop sequences of human miRNAs from miRBase version 21 [27]. Standard miRDeep2 settings were used for the analysis.

Raw read counts for each miRNA from miRDeep2 analysis were used as input for differential expression analysis in edgeR package version 3.16.5 [28,29]. Analysis was run on R software environment version 3.3.1. Samples with library size below 500,000 reads and those performing as clear outliers according to the multidimensional scaling (MDS) plot were omitted. The remaining number of patient samples used for each comparison is displayed in Figure 1. Count per million (CPM) of at least 1 in 75% of samples in the smaller group was set as a threshold for each miRNA to be included in the analysis for each separate comparison. The remaining raw library sizes were scaled according to the method of weighted trimmed mean of M-values [30].

For comparing samples from ES and MS phases, a generalized linear model (GLM) likelihood ratio test with paired design was performed. For comparison of samples from fertile group and

from patients with RIF, GLM test with age adjustment was implemented as age between those two groups was statistically different (*t*-test, $p < 0.05$). EST and ESP datasets were analysed separately and miRNAs that were detected as differentially expressed (DE) in both sample sets were included for further analysis.

2.7. Novel miRNA Identification

Novel miRNAs were predicted from 92 blood samples and 110 endometrial samples (Table 1). The reads predicted as potential candidate miRNAs by miRDeep2 were subjected to BLAST in order to discriminate the sequences corresponding to other human ncRNAs such as rRNA, snRNA, tRNA, lncRNA etc. Mature miRNA sequences of *Pan troglodytes* were included as input for evolutionary conservation analysis, which provides more solid confirmation for finding novel unannotated miRNAs from high-throughput sequencing experiments [31]. Sequences were considered as potentially novel miRNAs if they met the following criteria: (a) miRDeep2 score ≥ 5 , (b) significant randfold *p*-value, (c) detected in at least 3 samples within a sample/biofluid type (endometrium or blood), (d) detected from two different datasets (EST + ESP) or detected from different sample types (endometrium and blood). miRDB was used for novel miRNA target prediction [32] and Ingenuity Pathway Analysis (IPA, 2017, Qiagen Bioinformatics Redwood City, CA, USA) was performed on predicted targets.

2.8. mRNA Data Analysis for miRNA Target Prediction

For miRNA target gene prediction purposes, mRNA data from our previously published RNA-seq study were retrieved from Gene Expression Omnibus (GSE98386) [4,24]. The mRNA dataset comprised of a subset of the same samples that were used for miRNA analysis in the current study (Figure 1). Raw mRNA sequencing read quality was inspected with FastQC v.0.11.3 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) before and after the preprocessing step. Adapter removal and trimming was performed with Trimmomatic tool v. 0.32 [25]. Quality threshold was set to 30 according to Phred-33 scoring system and reads <36 nucleotides were omitted. FASTQ Quality Filter from FASTX-Toolkit v. 0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit/index.html) was used as a second level of quality control, where the average quality threshold of 30 for every 50 nucleotides was set. Good quality reads were mapped to the human reference genome hg19 using TopHat v. 2.0.11 [33]. Read counts for every gene in Ensembl v.75 annotation file were generated with HTSeq v.0.6.1 [34]. Differential gene expression analysis was performed by using edgeR package according to the same parameters as described in miRNA data analysis. List of DE mRNAs that were common for EST and ESP samples were passed to miRNA-mRNA interaction analysis.

2.9. Integrated Analysis of miRNA and mRNA Data

Differentially expressed miRNAs and mRNAs in endometrial samples from ES vs. MS phases from fertile women and from MS phase of RIF patients were further analysed using the software IPA. The tool microRNA Target Filter was applied which allows prioritization of experimentally validated and predicted mRNA targets from TargetScan, TarBase, miRecords, and the Ingenuity Knowledge Base. Expression pairing of our miRNA and mRNA RNA-seq data was performed using the confidence filter of ‘experimentally observed’ and ‘high confidence’, and the expression pairing filter of opposite expression directions was applied (pairs of up-regulated miRNA with down-regulated mRNA, and down-regulated miRNA with up-regulated mRNA). The involvement of detected miRNA target genes in different canonical pathways was analysed. The overall level of gene expression in the canonical pathways was analysed by GOplot package version 1.0.2 in R statistical environment. The *z*-score refers to the level of under- or overrepresented mRNAs in each pathway and is calculated as $z = (\text{number of up-regulated genes} - \text{number of down-regulated genes}) / \sqrt{\text{count}}$ [35].

miRNA-mRNA interactions were not analysed on data from blood samples because no DE miRNAs between blood samples corresponding to MS vs. ES phase were identified in either of the

two cohorts of fertile women (Figure 2), and no consensus mRNAs were identified between EST and ESP cohorts when comparing blood mRNAs from MS cycle phase of fertile vs. RIF women.

2.10. Novel miRNA Validation Using Quantitative Real-Time Polymerase Chain Reaction

Custom TaqMan Small RNA assay (Thermo Fisher Scientific) was used for novel miRNA (chr2_4401, sequence: gaacacugaaguuuauggcug) validation with eight paired ES and MS endometrial samples from fertile women and eight MS endometrial samples from RIF women. The average level of miR-151a-5p and miR-196b-5p was used as reference for normalization. These reference miRNAs were chosen due to their stable expression levels according to the current small RNA sequencing data. Additionally, five paired endometrial stromal (CD13⁺) and epithelial cell (CD9⁺) samples (two from ES and three from MS phase) isolated by fluorescence-activated cell sorting (FACS) were used to determine the cell type specificity of the novel miRNA. Cell sorting was performed according to our previous publication [36]. cDNA synthesis was conducted with TaqMan MicroRNA Reverse Transcription Kit (Thermo Fisher Scientific) and quantitative real-time polymerase chain reaction (qRT-PCR) was performed with TaqMan Universal PCR Master Mix, No AmpErase UNG (Thermo Fisher Scientific). Real-time experiments were performed in duplicate. Relative miRNA expression levels were compared between the studied groups by paired (ES vs. MS) or unpaired (fertile vs. RIF) two-tailed *t*-test (Excel, Microsoft Corporation, Redmond, WA, USA) and *p*-value ≤ 0.05 was considered as significant. Fold change (FC) was calculated using the $2^{-\Delta\Delta Ct}$ method [37].

2.11. Data Availability

The datasets analysed during the current study are available in the Gene Expression Omnibus repository (GEO), accession numbers GSE108966 and GSE98386.

3. Results

3.1. miRNAs in the Endometrium

In this study, 65 endometrial samples from the Estonian cohort (EST) and 51 samples from independent Spanish validation cohort (ESP) were included for miRNA data analysis (Figure 1). In total, 615 miRNAs were detected from endometrial samples among the EST cohort and 624 miRNAs among ESP cohort (CPM of at least 1 in 75% of samples within a sample group of ES or MS of fertile group, or RIF). Most abundant miRNAs in endometrial tissues from both cohorts were miR-10b-5p, miR-10a-5p and miR-27b-3p (Table S1A,B).

miRNAs were considered as DE if they showed significantly altered levels (FDR < 0.05) also in validation cohort (both in EST and ESP). From fertile women, 91 DE endometrial miRNAs were confirmed from MS vs. ES (Figure 2A) of which 49 were down- and 42 up-regulated in MS endometrium (Table S2A).

The comparison of MS endometrial samples from RIF patients vs. fertile women in EST and ESP validation datasets revealed 21 DE miRNAs (Figure 2B), out of which eight miRNAs were more abundantly expressed in RIF patients and 13 miRNAs in fertile women (Table S2B).

miR-424-5p was the only DE miRNA present in both comparisons—between the ES and MS endometria of fertile women (Table S2A) as well as between fertile women and RIF patients (Table S2B). Interestingly, while down-regulated in the MS endometrium of fertile women (average FC = -1.87 between the two cohorts), it was up-regulated in RIF patients' samples from the same phase (average FC = 1.74). When predicting miR-424-5p target genes from our mRNA dataset (Figure 1, see Methods section), we detected 85 targets (Table S3A) with involvement in different canonical pathways important in MS endometrial functions such as glucocorticoid, insulin receptor, axonal guidance and interleukins signalling (Table S4A). The miRNA-mRNA target analysis among endometrial samples from RIF women, based on our experimental datasets, identified miR-424-5p to target solely Serine/Threonine Kinase 2 (SGK2) gene (Table S3B).

3.2. miRNAs in Blood

Altogether, 65 blood samples from EST and 49 samples from ESP validation cohort were included for miRNA data analysis (Figure 1). In total, 305 miRNAs were detected from blood samples among EST cohort and 710 miRNAs among the independent ESP cohort (Table S1C,D). miRNAs were isolated from whole blood in EST samples and from the buffy coat fraction in ESP samples. Possibly due to the differences in sample treatment protocols between the centres, the most abundant miRNAs in blood varied between the two cohorts: miR-486-5p, miR-92a-3p and miR-451a demonstrated the highest read counts among EST blood samples, while miR-26a-5p, miR-191-5p and miR-181a-5p exhibited the highest expression levels among ESP samples.

No DE miRNAs were detected between blood samples corresponding to ES and MS cycle phase in fertile women either in the EST dataset or in the validation ESP dataset (Figure 2C). The comparison of blood samples corresponding to MS cycle phase from fertile and RIF women revealed that miR-30a-5p was significantly up-regulated among RIF patients in EST cohort (FC = 3.0, FDR = 0.01); and the difference was also confirmed in ESP cohort (FC = 1.9, FDR = 0.03) (Figure 2D).

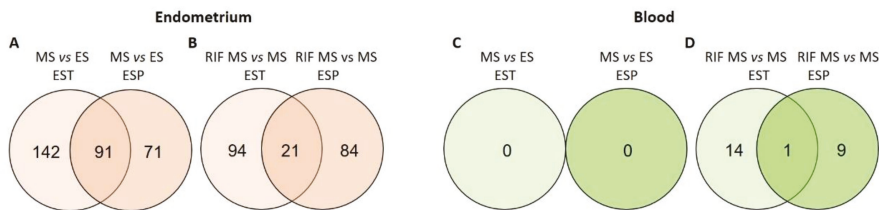


Figure 2. Number of differentially expressed miRNAs in EST and validation cohort ESP in (A) MS vs. ES phase endometrial samples of fertile women; (B) MS phase endometrial samples from infertile recurrent implantation failure (RIF) patient vs. fertile women; (C) MS vs. ES phase blood samples of fertile women; and (D) MS phase blood samples from RIF patients vs. fertile women. EST - Estonian samples; ESP—Spanish samples; RIF—recurrent implantation failure; ES—early secretory phase; MS—mid-secretory phase.

3.3. Novel miRNAs

In total, 18 novel miRNAs were determined from endometrial and blood samples (Tables 1 and S5). Out of these miRNAs, chr2_1900 (sequence: aucugaaauuugaaauggucc) and chr16_22077 (aggcuagcugggccacag) were detected only from MS endometrial samples (from 7 and 4 out of 73 MS samples, respectively) and chr2_4401 (gaacacugaaguuaauggcug) was found from the majority (75.3%; 55/73) of MS endometrial samples and only from 2.7% (1/37) of ES endometrial samples. chr14_10307 (ucugagcccgucuccuagg) was uniquely determined from blood samples of 16.7% RIF patients (4 out of 24 RIF blood samples).

We further focused on the most promising novel miRNA chr2_4401. Validation analysis by qRT-PCR confirmed the differential expression of chr2_4401 showing that the level of this novel miRNA was 37-fold higher ($p = 0.0001$) in MS compared to ES endometrium in fertile women (Figure 3A). No statistically significant differences in the expression levels of chr2_4401 between the MS endometrial samples from fertile and RIF women were detected (Figure 3A). Cell type-specific expression analysis showed very low levels of chr2_4401 in endometrial stromal cells, but 55-fold upregulation in epithelial cells was detected ($p = 0.003$, Figure 3B).

Table 1. Novel miRNAs detected from endometrial and blood samples.

Novel miRNA Provisional ID	Sample Type	Consensus Mature Sequence	Endometrium				Blood			
			Fertile Women		Recurrent Implantation Failure (RIF) Patients		Fertile Women		RIF Patients	
			Detected from Early Secretory (ES) Samples (n = 37)	Detected from Mid-Secretory (MS) Samples (n = 37)	Detected from MS Samples (n = 36)	Detected from ES Samples (n = 34)	Detected from MS Samples (n = 34)	Detected from MS Samples (n = 24)		
chr2_1900	Endometrium	aucugaaauuuugaauaggucc	-	3	4	-	-	-	-	-
chr2_4401	Endometrium	gaacacugaauguuaauaggcug	1	26	29	-	-	-	-	-
chr2_2219	Blood	cugagaagacagucgaacuuagac	-	-	-	11	7	5	5	5
chr3_7058	Endometrium, blood	aaagaauaacggucuuuugcc	-	2	-	3	4	5	5	5
chr3_1133	Endometrium, blood	caacaggccuugcucugcacaga	1	-	2	2	-	2	-	2
chr3_2060	Endometrium	cacggcaggagaccaccagc	7	2	1	-	-	-	-	-
chr4_6331	Endometrium	aacugggcaugtuuggaacuaagc	2	1	-	-	-	-	-	-
chr4_10018	Endometrium	aucagagaacuuucuuagac	3	1	9	-	-	-	-	-
chr8_10314	Endometrium, blood	uuauccuccagauagacuaggga	18	8	14	3	2	4	4	4
chr8_8702	Endometrium, blood	acacggggguuagagcucaacc	2	-	-	-	-	1	2	2
chr10_9545	Endometrium, blood	aaagruuuuugcgguuuuugc	1	-	1	1	2	2	2	2
chr12_27523	Endometrium, blood	ucuggccuuucuucauac	-	2	-	3	2	6	6	6
chr13_23821	Endometrium, blood	augugccuagugcugcuguc	5	5	4	6	6	7	7	7
chr14_10307	Blood	ucugagccctugtuucccuagg	-	-	-	-	-	-	-	4
chr14_3458	Endometrium	aaagucuaucugguuucucc	3	1	3	-	-	-	-	-
chr15_26374	Endometrium	aaaguuuuugggauuuugc	2	1	3	-	-	-	-	-
chr16_22077	Endometrium	cugactucccugccuccgaccag	-	2	2	-	-	-	-	-
chr17_11615	Endometrium, blood	uaacucuuaagaaucucccaag	1	-	-	18	16	10	10	10

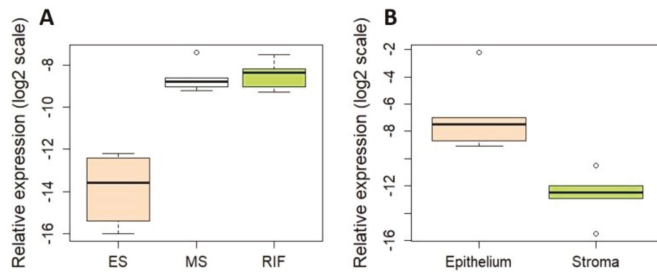


Figure 3. Validation of the novel miRNA chr2_4401 by quantitative real-time polymerase chain reaction (qRT-PCR). **(A)** miRNA expression level differences between paired ES ($n = 8$) and MS ($n = 8$) endometria of fertile women, and MS endometria of RIF ($n = 8$) patients. The relative miRNA expression levels (ΔCt) in the endometrium were 37-fold higher in MS compared to ES samples ($p = 0.0001$). No differences between MS samples of fertile and RIF patients were observed. **(B)** Relative expression level differences in epithelial ($n = 5$) and stromal ($n = 5$) fractions. The relative miRNA levels (ΔCt) were 55-fold higher in epithelial cells (CD9+) compared to stromal cells (CD13+) ($p = 0.003$). The average levels of miR-151a-5p and miR-196b-5p were used for data normalization. For illustrative purposes, relative expression levels (ΔCt) were multiplied by -1 . ES—early secretory phase; MS—mid-secretory phase.

The novel miRNA chr2_4401 precursor sequence (Figure 4A) is located within the short arm of chromosome 2 (p11.2) and is transcribed from intergenic region. No human miRNAs with high similarity sequences for the predicted novel miRNA was found from miRBase database, however, chr2_4401 shares the seed region (2–7 nt) with miR-200 family miRNAs (Figure 4B) and, therefore, target genes of the predicted miRNA are expected to be common with those of the miR-200 family. miRDB revealed more than 800 potential target genes for chr2_4401 (Table S6A). For pathway analysis, we focussed only on these potential targets that were detected as down-regulated in the MS endometrium in our mRNA-seq analysis from the same women (98 genes, Table S6B). The predicted target genes of this novel miRNA identified in our dataset were involved in pathways including estrogen-mediated S-phase entry ($p = 0.006$), cell cycle regulation by B-cell translocation gene (BTG) family proteins ($p = 0.01$), role of checkpoint kinase (CHK) proteins in cell cycle checkpoint control ($p = 0.03$), and epithelial adherens junction signalling ($p = 0.03$) (Table S4B).

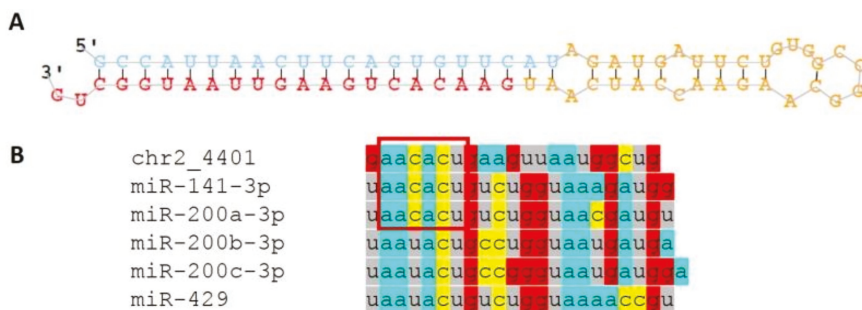


Figure 4. (A) Hairpin structure of the novel miRNA chr2_4401 precursor predicted by miRDeep2. (B) Sequence alignment between chr2_4401 and five miR-200 family miRNAs. The seed region (nucleotides 2–7, red square) of chr2_4401 is identical with miR-141-3p and miR-200a-3p.

3.4. miRNA–mRNA Interaction Prediction

The joint analysis of the miRNA and mRNA RNA-seq data from MS vs. ES endometrial samples by IPA software detected 81 (out of 91) DE miRNAs and 865 (out of 4240) mRNAs (Table S3A). Further analysis demonstrated that these 865 target mRNAs are involved in canonical pathways important

in several MS endometrial functions (Table S4C). According to the calculated z-scores, differential miRNA expression leads to the up-regulation of gene expression in the glucocorticoid, G-protein coupled receptor, IGF-1 and JAK/STAT signalling pathways (obtaining the highest z-score), while the most significant down-regulation of gene expression was observed within Wnt/beta-catenin signalling pathway (as demonstrated by the lowest z-score) (Figure 5).

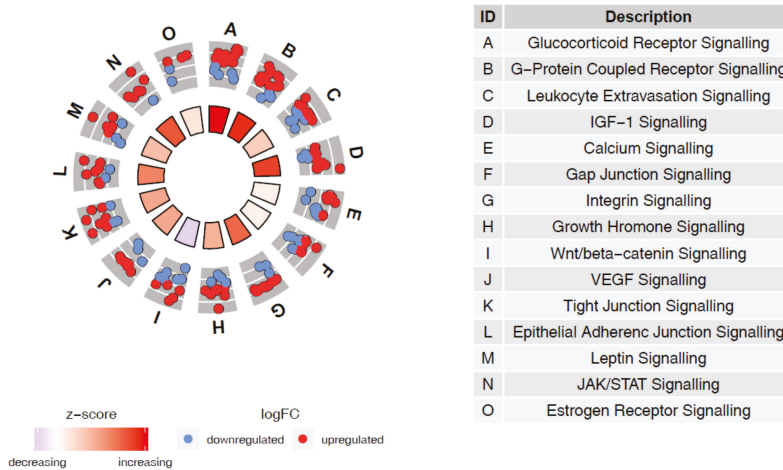


Figure 5. Circular plot of IPA canonical pathways enriched with miRNA targets differentially expressed between MS and ES endometrial samples of fertile women. The inner circle represents the overall up- or down-regulation of gene expression in each pathway according to the z-score. The outer circle depicts the fold change of expression level for each gene.

The analysis of our miRNA and mRNA sequencing data from MS endometrial samples from fertile vs. RIF women detected 21 DE miRNAs and 14 mRNAs. Out of those the IPA program identified 4 miRNAs in interaction with 5 mRNAs (Table S3B). The miRNA targets were detected to be involved in signal transducer and activator of transcription 3 (STAT3) and cyclin-dependent kinase 5 (CDK5) signalling pathways (Table S4D).

4. Discussion

Current knowledge about the involvement of miRNAs in endometrial transformation from ES (pre-receptive) to MS (receptive) stage both in fertility as well as in infertility conditions is limited. Moreover, it is not known whether the overall systemic changes in women during these processes are also reflected in blood miRNA profiles, which could serve as non-invasive biomarker candidates for evaluating fertility status. Therefore, using small RNA-seq technology, we profiled and validated the complete endometrial and blood miRNome from ES and MS phase samples from healthy fertile women and infertile RIF patients in two independent sample cohorts. This is the largest miRNA endometrial receptivity study to date with novel aspects of combining matched samples of the endometrium and blood from two different populations.

We identified 91 DE miRNAs between the ES and MS endometrium of healthy fertile individuals, showing the same directions in miRNA expression in both cohorts. Several miRNAs that have repeatedly been associated with endometrial receptivity were also confirmed by our study, including miR-30b-5p, miR-30d-3p, miR-30d-5p and miR-30a-5p that were up-regulated in the MS phase endometrium. Endometrial expression level changes of the miR-30 family members have been described in a number of previous endometrial receptivity studies [4,15–17,22,23,38]. Furthermore, it has been previously shown that endometrial miR-30d is taken up by the pre-implantation embryo,

resulting in modified transcriptome and embryo adhesion [16]. Therefore, our data together with previous studies collectively indicate that the miR-30 family is crucial in the regulation of endometrial receptivity and implantation processes in the endometrium.

Another interesting group of miRNAs in endometrial receptivity is the miR-200 family. We detected several miR-200 family members, such as miR-200a-3p, miR-200c-3p, miR-200c-5p, miR-141-3p, miR-141-5p and miR-429 to be up-regulated in the MS phase endometrium, being in line with previously published studies [16,23]. miR-200 family members are shown to target several genes in the endometrium that influence cell proliferation, migration and inflammation [39], which are all important cellular processes in MS endometrial functions. Interestingly, the miR-200 family is predominantly expressed in endometrial epithelial cells [36], which is the first cellular layer to interact with the implanting embryo. Nevertheless, as whole tissue biopsies were analysed in our study, we cannot confirm that the detected findings are epithelial cell specific.

Our study results also identified miRNAs that, to the best of our knowledge, have not been associated with human endometrial receptivity before (Table S2A). For instance, miR-873-3p was one of the most up-regulated and miR-3131 one of the most down-regulated miRNAs in the MS phase endometrium from fertile women. The roles of these miRNAs in endometrial tissue remain to be elucidated, but it has been demonstrated that miR-873-3p plays a role in the selection of bovine dominant follicle [40].

Interaction analysis of the DE miRNAs and their target genes identified the involvement of several canonical pathways important in endometrial receptivity, including JAK/STAT signalling, leptin signalling and growth hormone signalling. JAK/STAT signalling pathway transmits information from extracellular signals to the nucleus influencing transcription, and its involvement in embryo implantation is widely acknowledged [41–44]. Leptin signalling, regulated by cytokines and playing a role in inflammatory response, is another classically known signalling pathway to be involved in endometrial receptivity [43,45]. Furthermore, leptin is known to mediate the effects of growth hormone [46]. The beneficial effects of growth hormone administration on endometrial receptivity have recently been published, where growth hormone increased endometrial blood perfusion and the expression of different cytokines [47]. Growth hormone administration has also improved implantation, pregnancy and live birth rates in RIF patients [48]. The Wnt/beta-catenin signalling pathway has been demonstrated to control estrogen-dependent endometrial cell proliferation, decidualisation, trophoblast attachment and invasion. Fine-tuning of this pathway is particularly important as failures in Wnt signalling are associated with infertility, endometriosis, endometrial cancer and gestational diseases such as complete mole placentae and choriocarcinomas (reviewed in [49]). miRNAs differentially expressed upon the establishment of receptivity participate in the overall down-regulation of Wnt/beta-catenin signalling in the normal endometrium according to our results.

The endometrial biopsies from healthy participants were all taken at two time-points (ES and MS) within the same menstrual cycle. A recent study by Evans et al. demonstrated that the prior collection of an LH + 2 sample does not affect the general gene expression level of a LH + 7 sample [50]. Their study data suggested that the expression of tested genes exhibited a stable pattern, which was not affected even when sampled twice in one cycle and is valid to predict endometrial receptivity in the subsequent cycles. No such studies have been performed regarding miRNA expression, but similar variability in expression levels is expected within and between cycles as for mRNAs.

Besides the considerable changes in endometrial tissue, cyclic alterations take place throughout the female body during the menstrual cycle under the governance of steroid hormones. In order to find systemic miRNA changes as indirect markers that could reflect the receptive stage of the endometrium, we analysed blood samples that were collected simultaneously to endometrial biopsies. However, no DE blood miRNAs were found between the samples corresponding to ES and MS time-points in either of the cohorts, regardless of the sample isolation protocol. Prior to our current study, miRNAs from whole blood have not been investigated in abovementioned time-points, but we have previously shown that there are no differences in circulating plasma miRNA profile throughout the menstrual

cycle [51]. Therefore, it is probable that there are no blood-derived miRNAs that could reflect the changes which occur in female body during the establishment of the receptive endometrium.

Another interesting aspect of our study was the identification of dysregulated miRNAs among RIF women. Although this study group comprised of patients with various causes of infertility, they all had in common at least three unsuccessful IVF treatments with embryo transfers. We identified 21 DE miRNAs in the MS endometrium of RIF patients compared to fertile women. Notably, as has also been previously reported [21–23], miR-424-5p was up-regulated in RIF patients, while being down-regulated in the mid-secretory endometrium of healthy fertile women. Therefore, miR-424-5p may be a useful marker to determine endometrial insufficiency for embryo implantation. miR-424-5p has been previously shown to stimulate cell-to-cell adhesion during embryo implantation by targeting osteopontin (encoded by *SPP1* gene) [22,52,53]. Aberrant osteopontin levels in the endometrium have been linked to infertility [54].

Our target prediction analysis for miR-424-5p detected genes involved in numerous interleukins signalling pathways in the MS endometrium of fertile women, while the differential expression of a single miR-424-5p target gene, *SGK2*, was detected in the MS endometrium of RIF patients. Interleukins are a group of cytokines that regulate cell growth, differentiation, and they are particularly important in stimulating immune responses. The importance of immune responses in the pre- and peri-implantation periods in endometrial functions are widely acknowledged [4]. In order to provide a hospitable environment for the embryo, the balance between the maternal immune tolerance toward a semi-allogeneic implanting embryo and the protective anti-infectious mechanisms in the receptive phase uterus must be established [55]. Indeed, a recent study detected a range of interleukins to be down-regulated in infertile women with implantation failure in IVF [56]. *SGK2*, the potential target of miR-424-5p in RIF patients, is a protein kinase having powerful stimulating effect on K⁺ channels with a possible role in the regulation of epithelial transport and cell proliferation [57]. The importance of ion channels, including K⁺ in the regulation of endometrial receptivity has been summarised in a previous review [58]. However, to the best of our knowledge, *SGK2* has not been identified before in endometrial functions, and it could serve as an interesting target for future studies in women experiencing implantation failure.

Interaction analysis of our DE miRNAs and mRNAs identified the involvement of STAT3 and CDK5 signalling pathways in the development of RIF. CDK5 participates in a variety of cellular processes via the effects on angiogenesis, cell proliferation, cell adhesion, migration, and immune system (summarised in a recent review [59]).

Despite the differences in sample isolation protocols applied in the two cohorts, miR-30a-5p was elevated in the blood of RIF patients in both datasets. As miRNA levels in blood reflect its cellular composition, it is possible that the observed differences in miR-30a-5p levels between RIF patients and controls are derived from the alterations in the ratio of different cell types in blood samples. Although there are studies indicating that infertile women have altered levels of immune cells in blood, the results are highly conflicting [60], and therefore we cannot confirm nor rule out the possibility that miR-30a-5p alteration is due to variability in immune cell levels in RIF patients. Nevertheless, this miRNA could indirectly imply to the dysregulated physiological conditions resulting in implantation failure and serves as a potential minimally invasive biomarker in this regard. miR-30a-5p acts as a tumour suppressor by inhibiting cell proliferation and invasion [61–63]. In blood cells, miR-30a takes part in the regulation of erythrocyte maturation [64]. However, the precise mechanism that leads to the observed elevated levels in RIF patients' blood in our study remains obscure.

As an additional finding in our study, we identified 18 novel miRNAs. Recently, a publication involving nearly 500 small RNA libraries from different mammalian primary cells analysed whether previously annotated and unannotated short RNA sequences serve as valid miRNAs [65]. Several sequences proposed as novel unannotated miRNAs in our study matched the genomic coordinates of proposed precursor sequences (including chr2_1900 and chr2_4401) and some met the multiple high-confidence criteria set for miRNAs in the aforementioned publication [65] (Table S5). This adds

further confidence that the reported novel sequences in our study are genuine miRNAs. The most interesting finding to emerge from novel miRNA analysis was sequence chr2_4401 that showed higher expression levels in the MS vs. ES phase endometrium. Sequence alignment of chr2_4401 to several miR-200 family members demonstrates that their seed regions are identical. Therefore, we conclude that chr2_4401 plays similar important roles in MS endometrial functions as described above for the miR-200 family.

5. Conclusions

Our study approach, involving 230 samples, provides additional knowledge about the complex regulation of endometrial receptivity for successful embryo implantation. Our study results highlight the involvement of miR-30 and miR-200 family members in endometrial receptivity development, and miR-424-5p as dysregulated in RIF. The current study also identified several miRNAs not previously known to be involved in endometrial receptivity. We discovered several novel miRNAs, among which validated miRNA chr2_4401 is of most interest in MS endometrial functions. The results of our study provide new aspects of miRNA functions in MS endometrial processes in health and disease and provide means for further studies for identifying molecular biomarkers of fertility/infertility from endometrial and/or blood samples.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4425/9/12/574/s1>, Table S1: (a) miRNAs detected in EST endometrial samples. (b) miRNAs detected in ESP endometrial samples. (c) miRNAs detected in EST blood samples. (d) miRNAs detected in ESP blood samples. Table S2: (a) Differentially expressed miRNAs between MS and ES endometrial samples from fertile women. (b) Differentially expressed miRNAs from MS endometrial samples between RIF and fertile women. Table S3: (a) Predicted target genes for differentially expressed miRNAs in MS vs. ES phase endometrium from fertile women. (b) Predicted target genes for differentially expressed miRNAs in MS endometrium from fertile vs. infertile women with RIF. Table S4: (a) miR-424-5p target genes involved in different canonical pathways in MS phase endometrium. (b) Novel miRNA chr2_4401 target genes involved in different canonical pathways in MS phase endometrium. (c) Canonical pathways involved in MS endometrial functions (MS vs. ES endometrium from fertile women). (d) Canonical pathways involved in mid-secretory endometrial functions from infertile RIF women. Table S5: Novel miRNAs detected from endometrial and blood samples. Table S6: (a) All predicted targets for novel chr2_4401 miRNA (sequence: gaacacugaaguuauagcg) according to miRDB. (b) Predicted targets for novel miRNA chr2_4401 down-regulated in MS endometrium.

Author Contributions: K.R. performed experiments and data analysis, and wrote the manuscript; S.A. performed data analysis, and wrote the manuscript; M.S. participated in designing the study, and performed experiments; M.P. contributed to data analysis and interpretation, and critically revised the manuscript; J.F.M.-B. performed experiments; F.M.C. performed experiments; F.V. performed experiments; C.S. participated in designing the study; A.S. participated in designing the study, and critically revised the manuscript; A.V.-M. participated in designing the study, performed experiments and data analysis, and critically revised the manuscript. All authors contributed to drafting the manuscript and have approved the final version of the article.

Funding: This research was funded by Estonian Ministry of Education and Research (grant IUT34-16); Enterprise Estonia (grant EU48695); the EU-FP7 Eurostars program (grant NOTED, EU41564); the EU-FP7 Marie Curie Industry-Academia Partnerships and Pathways (IAPP, grant SARM, EU324509); Horizon 2020 innovation programme (WIDENLIFE, 692065); grant from the University of Granada (Incorporación de jóvenes doctores) and grant from the Spanish Ministry of Economy, Industry and Competitiveness–MINECO–(RYC-2016-21199 and grant ENDORE SAF2017-87526).

Acknowledgments: We are thankful to the personnel at the Nova Vita Clinic and Elite Clinic in Estonia as well as IVI Clinic Valencia for their involvement in sample collection, to Katrin Kepp and Pilar Alamá for the recruitment of healthy women to the study, and to Sergio Cabanillas for the recruitment of RIF patients. We are grateful to all study participants. We acknowledge Tatjana Jatsenko for her help with figures. We thank Parameswaran Grace Lalitkumar from Karolinska Institutet, Sweden for assisting with IPA analyses.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Simon, A.; Laufer, N. Repeated implantation failure: Clinical approach. *Fertil. Steril.* **2012**, *97*, 1039–1043. [[CrossRef](#)] [[PubMed](#)]

2. Macklon, N.S.; Stouffer, R.L.; Giudice, L.C.; Fauser, B.C. The science behind 25 years of ovarian stimulation for in vitro fertilization. *Endocr. Rev.* **2006**, *27*, 170–207. [[CrossRef](#)] [[PubMed](#)]
3. Altmäe, S.; Esteban, F.J.; Stavreus-Evers, A.; Simon, C.; Giudice, L.; Lessey, B.A.; Horcajadas, J.A.; Macklon, N.S.; D’Hooghe, T.; Campoy, C.; et al. Guidelines for the design, analysis and interpretation of ‘omics’ data: Focus on human endometrium. *Hum. Reprod. Update* **2014**, *20*, 12–28. [[CrossRef](#)] [[PubMed](#)]
4. Altmäe, S.; Koel, M.; Vosa, U.; Adler, P.; Suhorutsenko, M.; Laisk-Podar, T.; Kukushkina, V.; Saare, M.; Velthut-Meikas, A.; Krjutskov, K.; et al. Meta-signature of human endometrial receptivity: A meta-analysis and validation study of transcriptomic biomarkers. *Sci. Rep.* **2017**, *7*, 10077. [[CrossRef](#)] [[PubMed](#)]
5. Koler, M.; Achache, H.; Tsafir, A.; Smith, Y.; Revel, A.; Reich, R. Disrupted gene pattern in patients with repeated in vitro fertilization (IVF) failure. *Hum. Reprod.* **2009**, *24*, 2541–2548. [[CrossRef](#)] [[PubMed](#)]
6. Altmäe, S.; Martinez-Conejero, J.A.; Salumets, A.; Simon, C.; Horcajadas, J.A.; Stavreus-Evers, A. Endometrial gene expression analysis at the time of embryo implantation in women with unexplained infertility. *Mol. Hum. Reprod.* **2010**, *16*, 178–187. [[CrossRef](#)] [[PubMed](#)]
7. Altmäe, S.; Tamm-Rosenstein, K.; Esteban, F.J.; Simm, J.; Kolberg, L.; Peterson, H.; Metsis, M.; Haldre, K.; Horcajadas, J.A.; Salumets, A.; et al. Endometrial transcriptome analysis indicates superiority of natural over artificial cycles in recurrent implantation failure patients undergoing frozen embryo transfer. *Reprod. Biomed. Online* **2016**, *32*, 597–613. [[CrossRef](#)] [[PubMed](#)]
8. Ledee, N.; Munaut, C.; Aubert, J.; Serazin, V.; Rahmati, M.; Chaouat, G.; Sandra, O.; Foidart, J.M. Specific and extensive endometrial deregulation is present before conception in IVF/ICSI repeated implantation failures (IF) or recurrent miscarriages. *J. Pathol.* **2011**, *225*, 554–564. [[CrossRef](#)] [[PubMed](#)]
9. Koot, Y.E.; van Hooff, S.R.; Boomsma, C.M.; van Leenen, D.; Koerkamp, M.J.G.; Goddijn, M.; Eijkemans, M.J.; Fauser, B.C.; Holstege, F.C.; Macklon, N.S. An endometrial gene expression signature accurately predicts recurrent implantation failure after IVF. *Sci. Rep.* **2016**, *6*, 19411. [[CrossRef](#)] [[PubMed](#)]
10. Fan, L.J.; Han, H.J.; Guan, J.; Zhang, X.W.; Cui, Q.H.; Shen, H.; Shi, C. Aberrantly expressed long noncoding RNAs in recurrent implantation failure: A microarray related study. *Syst. Biol. Reprod. Med.* **2017**, *63*, 269–278. [[CrossRef](#)] [[PubMed](#)]
11. Carthew, R.W.; Sontheimer, E.J. Origins and mechanisms of miRNAs and siRNAs. *Cell* **2009**, *136*, 642–655. [[CrossRef](#)] [[PubMed](#)]
12. 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. [[CrossRef](#)] [[PubMed](#)]
13. Lam, E.W.; Shah, K.; Brosens, J.J. The diversity of sex steroid action: The role of micro-RNAs and FOXO transcription factors in cycling endometrium and cancer. *J. Endocrinol.* **2012**, *212*, 13–25. [[CrossRef](#)] [[PubMed](#)]
14. Liu, W.; Niu, Z.; Li, Q.; Pang, R.T.; Chiu, P.C.; Yeung, W.S. MicroRNA and embryo implantation. *Am. J. Reprod. Immunol.* **2016**, *75*, 263–271. [[CrossRef](#)] [[PubMed](#)]
15. Altmäe, S.; Martinez-Conejero, J.A.; Esteban, F.J.; Ruiz-Alonso, M.; Stavreus-Evers, A.; Horcajadas, J.A.; Salumets, A. MicroRNAs miR-30b, miR-30d, and miR-494 regulate human endometrial receptivity. *Reprod. Sci.* **2013**, *20*, 308–317. [[CrossRef](#)] [[PubMed](#)]
16. Vilella, F.; Moreno-Moya, J.M.; Balaguer, N.; Grasso, A.; Herrero, M.; Martinez, S.; Marcilla, A.; Simon, C. Hsa-miR-30d, secreted by the human endometrium, is taken up by the pre-implantation embryo and might modify its transcriptome. *Development* **2015**, *142*, 3210–3221. [[CrossRef](#)] [[PubMed](#)]
17. Moreno-Moya, J.M.; Vilella, F.; Martinez, S.; Pellicer, A.; Simon, C. The transcriptomic and proteomic effects of ectopic overexpression of miR-30d in human endometrial epithelial cells. *Mol. Hum. Reprod.* **2014**, *20*, 550–566. [[CrossRef](#)] [[PubMed](#)]
18. Choi, Y.; Kim, H.R.; Lim, E.J.; Park, M.; Yoon, J.A.; Kim, Y.S.; Kim, E.K.; Shin, J.E.; Kim, J.H.; Kwon, H.; et al. Integrative analyses of uterine transcriptome and microRNAome reveal compromised LIF-STAT3 signaling and progesterone response in the endometrium of patients with recurrent/repeated implantation failure (RIF). *PLoS ONE* **2016**, *11*, e0157696. [[CrossRef](#)] [[PubMed](#)]
19. Li, R.; Qiao, J.; Wang, L.; Li, L.; Zhen, X.; Liu, P.; Zheng, X. microRNA Array and microarray evaluation of endometrial receptivity in patients with high serum progesterone levels on the day of hCG administration. *Reprod. Biol. Endocrinol.* **2011**, *9*, 29. [[CrossRef](#)] [[PubMed](#)]
20. Revel, A.; Achache, H.; Stevens, J.; Smith, Y.; Reich, R. microRNAs are associated with human embryo implantation defects. *Hum. Reprod.* **2011**, *26*, 2830–2840. [[CrossRef](#)] [[PubMed](#)]

21. Shi, C.; Shen, H.; Fan, L.J.; Guan, J.; Zheng, X.B.; Chen, X.; Liang, R.; Zhang, X.W.; Cui, Q.H.; Sun, K.K.; et al. Endometrial microRNA signature during the window of implantation changed in patients with repeated implantation failure. *Chin. Med. J. (Engl.)* **2017**, *130*, 566–573. [[PubMed](#)]
22. Sha, A.G.; Liu, J.L.; Jiang, X.M.; Ren, J.Z.; Ma, C.H.; Lei, W.; Su, R.W.; Yang, Z.M. Genome-wide identification of micro-ribonucleic acids associated with human endometrial receptivity in natural and stimulated cycles by deep sequencing. *Fertil. Steril.* **2011**, *96*, 150–155. [[CrossRef](#)] [[PubMed](#)]
23. Sigurgeirsson, B.; Amark, H.; Jemt, A.; Ujvari, D.; Westgren, M.; Lundberg, J.; Gidlof, S. Comprehensive RNA sequencing of healthy human endometrium at two time points of the menstrual cycle. *Biol. Reprod.* **2017**, *96*, 24–33. [[PubMed](#)]
24. Suhorutshenko, M.; Kukushkina, V.; Velthut-Meikas, A.; Altmae, S.; Peters, M.; Magi, R.; Krjutskov, K.; Koel, M.; Codoner, F.M.; Martinez-Blanch, J.F.; et al. Endometrial receptivity revisited: Endometrial transcriptome adjusted for tissue cellular heterogeneity. *Hum. Reprod.* **2018**, *33*, 2074–2086. [[CrossRef](#)] [[PubMed](#)]
25. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
26. Friedlander, M.R.; Mackowiak, S.D.; Li, N.; Chen, W.; Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucl. Acids Res.* **2012**, *40*, 37–52. [[CrossRef](#)] [[PubMed](#)]
27. Griffiths-Jones, S. The microRNA registry. *Nucl. Acids Res.* **2004**, *32*, D109–D111. [[CrossRef](#)] [[PubMed](#)]
28. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [[CrossRef](#)] [[PubMed](#)]
29. McCarthy, D.J.; Chen, Y.; Smyth, G.K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucl. Acids Res.* **2012**, *40*, 4288–4297. [[CrossRef](#)] [[PubMed](#)]
30. Robinson, M.D.; Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genom. Biol.* **2010**, *11*, R25. [[CrossRef](#)] [[PubMed](#)]
31. Friedlander, M.R.; Chen, W.; Adamidi, C.; Maaskola, J.; Einspanier, R.; Knespel, S.; Rajewsky, N. Discovering microRNAs from deep sequencing data using miRDeep. *Nat. Biotechnol.* **2008**, *26*, 407–415. [[CrossRef](#)] [[PubMed](#)]
32. Wong, N.; Wang, X. miRDB: An online resource for microRNA target prediction and functional annotations. *Nucl. Acids Res.* **2015**, *43*, D146–D152. [[CrossRef](#)] [[PubMed](#)]
33. Kim, D.; Pertea, G.; Trapnell, C.; Pimentel, H.; Kelley, R.; Salzberg, S.L. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genom. Biol.* **2013**, *14*, R36. [[CrossRef](#)] [[PubMed](#)]
34. Anders, S.; Pyl, P.T.; Huber, W. HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015**, *31*, 166–169. [[CrossRef](#)] [[PubMed](#)]
35. Walter, W.; Sanchez-Cabo, F.; Ricote, M. GOrilla: An R package for visually combining expression data with functional analysis. *Bioinformatics* **2015**, *31*, 2912–2914. [[CrossRef](#)] [[PubMed](#)]
36. Saare, M.; Rekker, K.; Laisk-Podar, T.; Soritsa, D.; Roost, A.M.; Simm, J.; Velthut-Meikas, A.; Samuel, K.; Metsalu, T.; Karro, H.; et al. High-throughput sequencing approach uncovers the miRNome of peritoneal endometriotic lesions and adjacent healthy tissues. *PLoS ONE* **2014**, *9*, e112630. [[CrossRef](#)] [[PubMed](#)]
37. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_t}$ method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)] [[PubMed](#)]
38. Kuokkanen, S.; Chen, B.; Ojalvo, L.; Benard, L.; Santoro, N.; Pollard, J.W. Genomic profiling of microRNAs and messenger RNAs reveals hormonal regulation in microRNA expression in human endometrium. *Biol. Reprod.* **2010**, *82*, 791–801. [[CrossRef](#)] [[PubMed](#)]
39. Panda, H.; Pelakh, L.; Chuang, T.D.; Luo, X.; Bukulmez, O.; Chegini, N. Endometrial miR-200c is altered during transformation into cancerous states and targets the expression of *ZEBs*, *VEGFA*, *FLT1*, *IKK β* , *KLF9*, and *FBLN5*. *Reprod. Sci.* **2012**, *19*, 786–796. [[CrossRef](#)] [[PubMed](#)]
40. Sontakke, S.D.; Mohammed, B.T.; McNeilly, A.S.; Donadeu, F.X. Characterization of microRNAs differentially expressed during bovine follicle development. *Reproduction* **2014**, *148*, 271–283. [[CrossRef](#)] [[PubMed](#)]
41. Aghajanova, L.; Altmae, S.; Bjuresten, K.; Hovatta, O.; Landgren, B.M.; Stavreus-Evers, A. Disturbances in the LIF pathway in the endometrium among women with unexplained infertility. *Fertil. Steril.* **2009**, *91*, 2602–2610. [[CrossRef](#)] [[PubMed](#)]

42. Pawar, S.; Starosvetsky, E.; Orvis, G.D.; Behringer, R.R.; Bagchi, I.C.; Bagchi, M.K. STAT3 regulates uterine epithelial remodeling and epithelial-stromal crosstalk during implantation. *Mol. Endocrinol.* **2013**, *27*, 1996–2012. [[CrossRef](#)] [[PubMed](#)]
43. Altmäe, S.; Reimand, J.; Hovatta, O.; Zhang, P.; Kere, J.; Laisk, T.; Saare, M.; Peters, M.; Vilo, J.; Stavreus-Evers, A.; et al. Research resource: Interactome of human embryo implantation: Identification of gene expression pathways, regulation, and integrated regulatory networks. *Mol. Endocrinol.* **2012**, *26*, 203–217. [[CrossRef](#)] [[PubMed](#)]
44. Aghajanova, L. Update on the role of leukemia inhibitory factor in assisted reproduction. *Curr. Opin. Obstet. Gynecol.* **2010**, *22*, 213–219. [[CrossRef](#)] [[PubMed](#)]
45. Gonzalez, R.R.; Simon, C.; Caballero-Campo, P.; Norman, R.; Chardonnens, D.; Devoto, L.; Bischof, P. Leptin and reproduction. *Hum. Reprod. Update* **2000**, *6*, 290–300. [[CrossRef](#)] [[PubMed](#)]
46. Margetic, S.; Gazzola, C.; Pegg, G.G.; Hill, R.A. Leptin: A review of its peripheral actions and interactions. *Int. J. Obes. Relat. Metab. Disord.* **2002**, *26*, 1407–1433. [[CrossRef](#)] [[PubMed](#)]
47. Wang, X.-M.; Hong, J.; Zhang, W.-X.; Li, Y. The effects of growth hormone on clinical outcomes after frozen-thawed embryo transfer. *Int. J. Gynaecol. Obstet.* **2016**, *133*, 347–350.
48. Altmäe, S.; Mendoza-Tesarik, R.; Mendoza, C.; Mendoza, N.; Cucinelli, F.; Tesarik, J. Effect of growth hormone on uterine receptivity in women with repeated implantation failure in an oocyte donation program: A randomized controlled trial. *J. Endocr. Soc.* **2017**, *2*, 96–105. [[CrossRef](#)] [[PubMed](#)]
49. Sonderegger, S.; Pollheimer, J.; Knofler, M. Wnt signalling in implantation, decidualisation and placental differentiation—Review. *Placenta* **2010**, *31*, 839–847. [[CrossRef](#)] [[PubMed](#)]
50. Evans, G.E.; Phillipson, G.T.M.; Sykes, P.H.; McNoe, L.A.; Print, C.G.; Evans, J.J. Does the endometrial gene expression of fertile women vary within and between cycles? *Hum. Reprod.* **2018**, *33*, 452–463. [[CrossRef](#)] [[PubMed](#)]
51. Rekker, K.; Saare, M.; Roost, A.M.; Salumets, A.; Peters, M. Circulating microRNA profile throughout the menstrual cycle. *PLoS ONE* **2013**, *8*, e81166. [[CrossRef](#)] [[PubMed](#)]
52. Johnson, G.A.; Burghardt, R.C.; Bazer, F.W. Osteopontin: A leading candidate adhesion molecule for implantation in pigs and sheep. *J. Anim. Sci. Biotechnol.* **2014**, *5*, 56. [[CrossRef](#)] [[PubMed](#)]
53. Kang, Y.J.; Forbes, K.; Carver, J.; Aplin, J.D. The role of the osteopontin-integrin $\alpha\beta3$ interaction at implantation: Functional analysis using three different in vitro models. *Hum. Reprod.* **2014**, *29*, 739–749. [[CrossRef](#)] [[PubMed](#)]
54. Casals, G.; Ordi, J.; Creus, M.; Fabregues, F.; Carmona, F.; Casamitjana, R.; Balasch, J. Osteopontin and $\alpha\beta3$ integrin as markers of endometrial receptivity: The effect of different hormone therapies. *Reprod. Biomed. Online* **2010**, *21*, 349–359. [[CrossRef](#)] [[PubMed](#)]
55. Haller-Kikkatalo, K.; Altmäe, S.; Tagoma, A.; Uibo, R.; Salumets, A. Autoimmune activation toward embryo implantation is rare in immune-privileged human endometrium. *Semin. Reprod. Med.* **2014**, *32*, 376–384. [[CrossRef](#)] [[PubMed](#)]
56. Pathare, A.D.S.; Zaveri, K.; Hinduja, I. Downregulation of genes related to immune and inflammatory response in IVF implantation failure cases under controlled ovarian stimulation. *Am. J. Reprod. Immunol.* **2017**, *78*. [[CrossRef](#)] [[PubMed](#)]
57. Gamper, N.; Fillon, S.; Feng, Y.; Friedrich, B.; Lang, P.A.; Henke, G.; Huber, S.M.; Kobayashi, T.; Cohen, P.; Lang, F. K^+ channel activation by all three isoforms of serum- and glucocorticoid-dependent protein kinase SGK. *Pflugers Arch.* **2002**, *445*, 60–66. [[PubMed](#)]
58. Ruan, Y.C.; Chen, H.; Chan, H.C. Ion channels in the endometrium: Regulation of endometrial receptivity and embryo implantation. *Hum. Reprod. Update* **2014**, *20*, 517–529. [[CrossRef](#)] [[PubMed](#)]
59. Shupp, A.; Casimiro, M.C.; Pestell, R.G. Biological functions of CDK5 and potential CDK5 targeted clinical treatments. *Oncotarget* **2017**, *8*, 17373–17382. [[CrossRef](#)] [[PubMed](#)]
60. Seshadri, S.; Sunkara, S.K. Natural killer cells in female infertility and recurrent miscarriage: A systematic review and meta-analysis. *Hum. Reprod. Update* **2014**, *20*, 429–438. [[CrossRef](#)] [[PubMed](#)]
61. Ortega, M.; Bhatnagar, H.; Lin, A.P.; Wang, L.; Aster, J.C.; Sill, H.; Aguiar, R.C. A microRNA-mediated regulatory loop modulates NOTCH and MYC oncogenic signals in B- and T-cell malignancies. *Leukemia* **2015**, *29*, 968–976. [[CrossRef](#)] [[PubMed](#)]
62. Zhang, S.; Liu, Q.; Zhang, Q.; Liu, L. microRNA-30a-5p suppresses proliferation, invasion and tumor growth of hepatocellular cancer cells via targeting FOXA1. *Oncol. Lett.* **2017**, *14*, 5018–5026. [[CrossRef](#)] [[PubMed](#)]


63. Liu, X.; Ji, Q.; Zhang, C.; Liu, Y.; Liu, N.; Sui, H.; Zhou, L.; Wang, S.; Li, Q. miR-30a acts as a tumor suppressor by double-targeting COX-2 and BCL9 in *H. pylori* gastric cancer models. *Sci. Rep.* **2017**, *7*, 7113. [[CrossRef](#)] [[PubMed](#)]
64. Rouzbeh, S.; Kobari, L.; Cambot, M.; Mazurier, C.; Hebert, N.; Faussat, A.M.; Durand, C.; Douay, L.; Lapillonne, H. Molecular signature of erythroblast enucleation in human embryonic stem cells. *Stem Cells* **2015**, *33*, 2431–2441. [[CrossRef](#)] [[PubMed](#)]
65. de Rie, D.; Abugessaisa, I.; Alam, T.; Arner, E.; Arner, P.; Ashoor, H.; Astrom, G.; Babina, M.; Bertin, N.; Burroughs, A.M.; et al. An integrated expression atlas of miRNAs and their promoters in human and mouse. *Nat. Biotechnol.* **2017**, *35*, 872–878. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

Systems Analysis of Transcriptomic and Proteomic Profiles Identifies Novel Regulation of Fibrotic Programs by miRNAs in Pulmonary Fibrosis Fibroblasts

Steven Mullenbrock ^{1,†}, Fei Liu ^{1,‡}, Suzanne Szak ¹, Xiaoping Hronowski ¹, Benbo Gao ¹, Peter Juhasz ¹, Chao Sun ¹, Mei Liu ¹, Helen McLaughlin ¹, Qiurong Xiao ¹, Carol Feghali-Bostwick ² and Timothy S. Zheng ^{1,*} 

¹ Biogen, Cambridge, MA 02142, USA; steven.mullenbrock@cellsignal.com (S.M.); fliu@admirx.com (F.L.); suzanne.szak@biogen.com (S.S.); xiaoping.hronowski@biogen.com (X.H.); benbo.gao@biogen.com (B.G.); peter.juhasz@biogen.com (P.J.); chao.sun@biogen.com (C.S.); mei.liu@biogen.com (M.L.); helen.mclaughlin@biogen.com (H.M.); qiurong.xiao@biogen.com (Q.X.)

² Division of Rheumatology and Immunology, Department of Medicine, The Medical University of South Carolina, Charleston, SC 29425, USA; feghalib@musc.edu

* Correspondence: timothy.zheng@boehringer-ingenheim.com

† Current address: Cell Signaling Technology, Danvers, MA 01923, USA.

‡ Current address: Admirx Inc., Cambridge, MA 02139, USA.

Received: 13 October 2018; Accepted: 23 November 2018; Published: 29 November 2018

Abstract: Fibroblasts/myofibroblasts are the key effector cells responsible for excessive extracellular matrix (ECM) deposition and fibrosis progression in both idiopathic pulmonary fibrosis (IPF) and systemic sclerosis (SSc) patient lungs, thus it is critical to understand the transcriptomic and proteomic programs underlying their fibrogenic activity. We conducted the first integrative analysis of the fibrotic programming in these cells at the levels of gene and microRNA (miRNA) expression, as well as deposited ECM protein to gain insights into how fibrotic transcriptional programs culminate in aberrant ECM protein production/deposition. We identified messenger RNA (mRNA), miRNA, and deposited matrisome protein signatures for IPF and SSc fibroblasts obtained from lung transplants using next-generation sequencing and mass spectrometry. SSc and IPF fibroblast transcriptional signatures were remarkably similar, with enrichment of WNT, TGF- β , and ECM genes. miRNA-seq identified differentially regulated miRNAs, including downregulation of miR-29b-3p, miR-138-5p and miR-146b-5p in disease fibroblasts and transfection of their mimics decreased expression of distinct sets of fibrotic signature genes as assessed using a Nanostring fibrosis panel. Finally, proteomic analyses uncovered a distinct “fibrotic” matrisome profile deposited by IPF and SSc fibroblasts compared to controls that highlights the dysregulated ECM production underlying their fibrogenic activities. Our comprehensive analyses of mRNA, miRNA, and matrisome proteomic profiles in IPF and SSc lung fibroblasts revealed robust fibrotic signatures at both the gene and protein expression levels and identified novel fibrogenesis-associated miRNAs whose aberrant downregulation in disease fibroblasts likely contributes to their fibrotic and ECM gene expression.

Keywords: interstitial lung disease; idiopathic pulmonary fibrosis; systemic sclerosis; myofibroblast; gene expression; proteomics

1. Introduction

Interstitial lung disease (ILD) associated with idiopathic pulmonary fibrosis (IPF) and systemic sclerosis (SSc) has a critical impact on a patient’s quality of life and is the predominant cause of

mortality in these diseases. Although the pathogenesis of pulmonary fibrosis in IPF and SSc remains incompletely understood, it is generally accepted that they stem from different root causes, with clinical and genetic evidence supporting epithelial injury/dysfunction and vasculopathy/inflammation as the underlying pathogenic triggers for IPF and SSc, respectively [1]. Despite their distinct origins, the resulting persistent tissue injury events converge on pathological fibroblast/myofibroblast activation, culminating in excessive extracellular matrix (ECM) deposition and ultimately progressive loss of lung function in both IPF and SSc [2].

To understand mechanisms underlying pulmonary fibrosis, several groups have undertaken transcriptomic analyses of both tissue and fibroblasts/myofibroblasts derived from fibrotic lung of IPF and SSc patients [3–9]. These studies found that a limited number of genes, pathways, and functions are altered in pulmonary fibrosis, including TGF- β and WNT, as well as altered expression of ECM genes such as collagens, crosslinking enzymes, TIMPs, and MMPs, many of which have been shown to be functionally relevant in fibrogenesis by subsequent *in vitro* and *in vivo* studies. Importantly, key fibrosis-associated pathways (e.g., ECM, WNT, and TGF- β) identified from transcriptomic analyses of fibrotic lung tissue were also captured in the gene signatures of isolated fibroblasts, suggesting that the molecular programming in pulmonary fibrosis is driven in large part by fibroblasts/myofibroblasts.

While these studies have shed insights into mechanisms underlying fibrogenic activation of fibroblasts/myofibroblasts, their fibrotic programs remain incompletely understood. For example, most early studies were performed using microarrays, which are less able to accurately detect low abundance transcripts and are limited to interrogating the expression of the transcripts present on each array platform, which precludes measuring microRNAs (miRNAs) in many cases [3–9]. To our knowledge, no global miRNA studies have been reported for either IPF or SSc lung fibroblasts and how miRNAs regulate their fibrotic program remains to be characterized. In addition, excessive ECM protein deposition by fibroblasts/myofibroblasts is directly responsible for IPF and SSc disease pathology, yet surprisingly little has been done to characterize the aberrant matrixome protein profile of these disease fibroblasts, which is not only affected by transcriptional changes, but is also subjected to post-transcriptional regulation [10].

To thoroughly interrogate mechanisms by which pathological activation of fibroblasts/myofibroblasts is regulated in IPF and SSc, we performed genome-wide analyses of both messenger RNA (mRNA) and miRNA in these cells, and characterized their ECM deposition properties by proteomic analysis. Altogether, this current study is the first integrative analysis of fibrotic gene and protein signatures, providing novel insights into the multitude of regulatory mechanisms governing the fibrogenic potential of fibroblasts/myofibroblasts in pulmonary fibrosis.

2. Materials and Methods

2.1. Cell Culture

Primary fibroblasts were isolated from lung tissues of normal donors whose lungs were not used for transplantation and SSc or IPF patients who underwent lung transplantation at the University of Pittsburgh Medical Center under a protocol approved by the institution's Institutional Review Board. Isolation and subsequent culture of lung fibroblasts was previously described [4].

2.2. RNA-seq and miRNA-seq Analysis

RNA-seq libraries were prepared using the Illumina (San Diego, CA, USA) TruSeq RNA Sample kit with poly-T selection and sequenced using a HiSeq (75-bp paired-end reads). Reads were mapped using Tophat (version 2.0.8), transcripts were assembled using Cufflinks (version 2.2.1), and differential expression was calculated using CuffDiff.

miRNA-seq libraries were prepared using the Illumina TruSeq Small RNA Sample Kit and sequenced using an Illumina miSeq (51-bp single-end reads). Bowtie (version 0.12.5) was used to perform a stepwise alignment of fastq files to Illumina databases. Differentially expressed miRNAs were identified using Bioconductor's limma package.

2.3. Real-Time Reverse Transcription-Polymerase Chain Reaction

RNA was reverse transcribed using the High Capacity cDNA Reverse Transcription Kit with RNase inhibitor (Life Technologies/Thermo Fisher Scientific (Waltham, MA, USA) #4374966) or TaqMan MicroRNA Reverse Transcription Kit (Life Technologies/Thermo Fisher Scientific (Waltham, MA, USA) #4366596) in conjunction with appropriate miRNA reverse transcription primers (Table S1). Real-time quantitative PCR (see Table S1 for TaqMan probes) was run on a Life Technologies QuantStudio 12K Flex.

2.4. Gene Ontology and Signature Analysis

Gene ontology (GO) analysis was conducted using DAVID Bioinformatics Database [11] and gene signature analysis was conducted using NextBio curated studies [12] and pre-ranked gene set enrichment analysis (GSEA) [13]. Ingenuity Pathway Analysis (IPA) was utilized to interrogate pathways upstream of IPF and SSc differentially expressed gene sets.

2.5. miRNA Mimic Transfection and Nanostring Gene Expression Analysis

Primary lung fibroblasts were reverse transfected with a miRNA mimic or negative control mimic (2.5 nM final concentration) using Lipofectamine RNAiMAX (Invitrogen/Thermo Fisher Scientific, Waltham, MA, USA) for 48 h, after which gene expression was assayed with the Nanostring (Seattle, WA, USA) platform using a custom codeset of common “fibrosis” genes.

2.6. Extracellular matrix Proteomic Profiling

Extracellular matrix proteins deposited by patient-derived fibroblasts after 4 weeks of culture were enriched by a sequential extraction of cellular and extracellular proteins as described previously [14,15]. The extracted soluble and insoluble ECM samples were further processed for proteomic analysis using Thermo Fisher Scientific’s (Waltham, MA, USA) TMT10plex label reagent. The labeled samples were pooled together and fractionated into three fractions. LC-MS of the fractions were acquired using a Thermo Fisher Scientific QE HF mass spectrometer. Peptide identification and quantification were performed using Maxquant [16] searched against the human Swiss-Prot reference database (<https://www.uniprot.org/>). Protein levels among fibroblast groups were compared by ANOVA test followed by Tukey’s HSD test.

Additional details for Materials and Methods are provided in Supplementary Materials.

3. Results

3.1. Transcriptional Profiling Identified Similar Dysregulated Gene Expression Programs in Idiopathic Pulmonary Fibrosis and Systemic Sclerosis Lung Fibroblasts

Lung fibroblasts isolated from IPF and SSc patients undergoing lung transplant and unused healthy donor lung were grown under similar culture conditions at low passage number (passage 2-3) prior to performing RNA-seq and miRNA-seq. As reflected by the low forced vital capacity % (FVC) (FVC < 60%) and the patients’ requirement for a transplant, this study provides a snapshot of fibroblasts from patients with severe, end-stage lung disease (Tables S2 and S3).

RNA-seq analysis of the 30 primary lung fibroblasts (n = 10 for each group) identified 297 differentially expressed genes (DEGs) across all cohort comparisons ($\geq \pm 1.5$ -fold, false discovery rate (FDR) $q < 0.05$) (Figure 1A and Table S4). We confirmed differential expression of several genes by qPCR (Figure 1B) and expression measured by qPCR and RNA-seq were highly correlated (Figure S1).

The majority of DEGs were observed between disease and control fibroblasts, with 168 DEGs for the IPF vs. Normal and 176 DEGs for the SSc vs. Normal comparison. The IPF and SSc fibroblast gene signatures are very similar, as over 40% of the DEGs overlap between them ($p < 0.001$), and for those that do not overlap their expression changes generally trend in a similar direction (Figure 1A,C). This is further supported by the principle component analysis (PCA) of the patient samples using the

297 DEGs across all comparisons, which illustrates that the IPF and SSc patients cluster close to one another along the first principle component (Figure 1D).

Interestingly we did identify expression differences between IPF and SSc fibroblasts (68 DEGs IPF vs. SSc, $\geq \pm 1.5$ -fold, FDR $q < 0.05$, Figure S2), although this seemed to be driven by a small subset of SSc patients that were among those with notes of pulmonary hypertension, as shown by their clear separation along the second principle component of the PCA plot (Figure 1D), although it should be noted that other SSc patients with pulmonary hypertension did not separate out similarly.

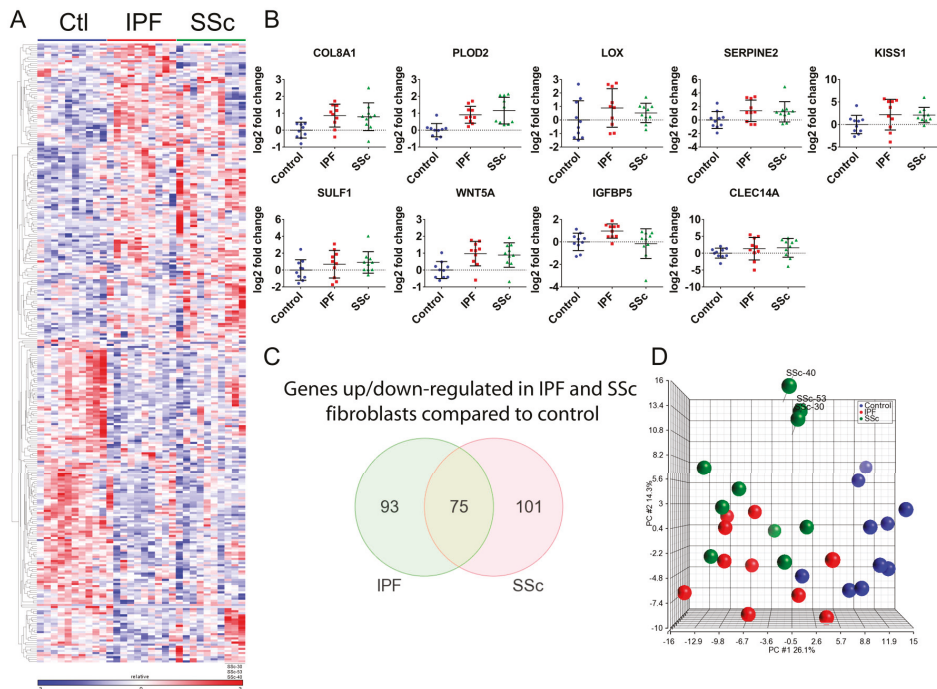


Figure 1. RNA-seq identifies 297 differentially expressed genes (DEGs) across idiopathic pulmonary fibrosis (IPF), systemic sclerosis (SSc), and healthy control primary lung fibroblast comparisons and demonstrates similar disease signatures between IPF and SSc lung fibroblasts. (A) Heatmap depicting 297 differentially expressed genes as determined from RNA-seq analysis using Cuffdiff (fold-change $\geq +1.5$ -fold or ≤ -1.5 -fold, $q < 0.05$) across all comparisons (IPF vs. Control, SSc vs. Control, SSc vs. IPF). (B) Quantitative polymerase chain reaction (qPCR) analysis validated differential expression of several genes involved in profibrotic pathways. Gene expression analysis using real-time qPCR on the 30 patient fibroblast samples was conducted as described in Materials and Methods and normalized to *GAPDH*. Data is plotted as a \log_2 fold-change relative to the mean of the healthy control samples. (C) Venn diagram demonstrating overlap of statistically significant differentially expressed genes between IPF and SSc fibroblasts. (D) Principle component analysis on the 30 patient fibroblasts was run using the 297 genes that were significantly different across all patient group comparisons. This allows for the visualization of how similar/different the patient groups are from one another based on disease genes as well as IPF or SSc-specific genes. Healthy control fibroblasts= Ctl.

3.2. Idiopathic Pulmonary Fibrosis and Systemic Sclerosis Lung Fibroblast Disease Signatures are Associated with Pro-Fibrotic Pathways and Extracellular Matrix

The highly similar gene expression profiles of IPF and SSc lung fibroblasts likely represent a fibrotic disease signature that reflects aberrant activation of upstream signaling that sustains these fibrotic programs and that is directly involved with the pathological function of fibroblasts/myofibroblasts

in ILD. Utilizing computational analyses such as GO enrichment analysis, IPA, and gene signature analysis using rank-based directional enrichment tools such as NextBio [12] and GSEA [13], we characterized the IPF and SSc fibroblast signatures to determine how they may relate to both upstream signaling pathways and potential downstream functions.

These tools revealed that IPF and SSc fibroblast signatures are associated with the activation of several profibrotic signaling pathways such as WNT (Figure 2B), TGF- β (Figure 2C, Figure S3), NOTCH1 (Figure 2D), and HIF1A (Figure 2D), as well as inhibition of the anti-fibrotic PPARG pathway (Figure 2D).

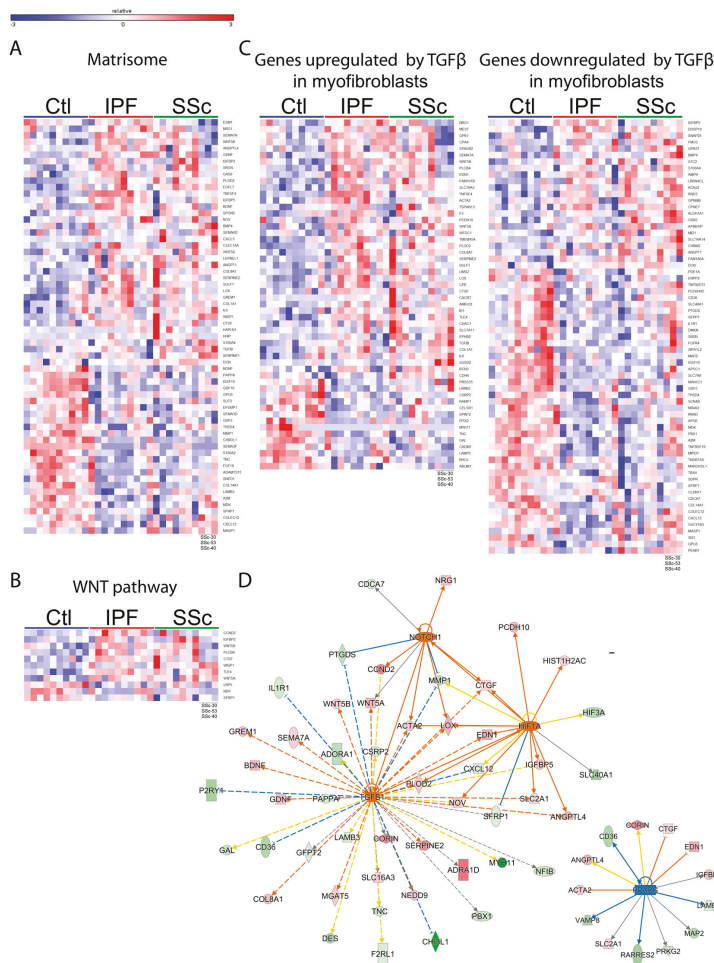


Figure 2. The dysregulated gene expression program in disease fibroblasts is composed of altered matrisome genes and associated with signaling pathways reflective of pathological fibroblast activation. Shown are heatmaps depicting expression of genes differentially expressed in IPF or SSc fibroblasts that overlapped significantly with various gene/pathway signatures such as matrisome (A), WNT (B), or TGF- β (C). (D) IPA upstream regulator analysis was used to predict pathway/transcription factor activation state upstream of the IPF or SSc DEGs. The pathways shown here are based on data from IPF DEGs (SSc DEGs had similar results). Shown is the predicted activation/repression of TGF- β , NOTCH1, HIF1A, and PPARG with lines connecting to genes differentially expressed in IPF fibroblasts that are downstream of these pathways. Healthy control fibroblasts= Ctl.

To further probe potential downstream functions of genes altered in IPF and SSc fibroblasts, we used GO analysis and found that the disease fibroblast gene signatures are enriched for genes associated with “cell proliferation” (IPF and SSc upregulated genes), “muscle contraction” (IPF upregulated genes), “response to wounding” (IPF downregulated genes), and “metallopeptidase activity” (IPF downregulated genes) (Figure 3). In addition, GO terms associated with ECM were the most significant and frequently observed terms for the upregulated and downregulated gene sets for both IPF and SSc. Consistent with this, we also found that the IPF and SSc fibroblast signatures are significantly enriched for components of the in silico matrisome derived by Naba et al. [17], with 64 of the 269 DEGs being associated with the matrisome (Figure 2A, $p < 0.001$).

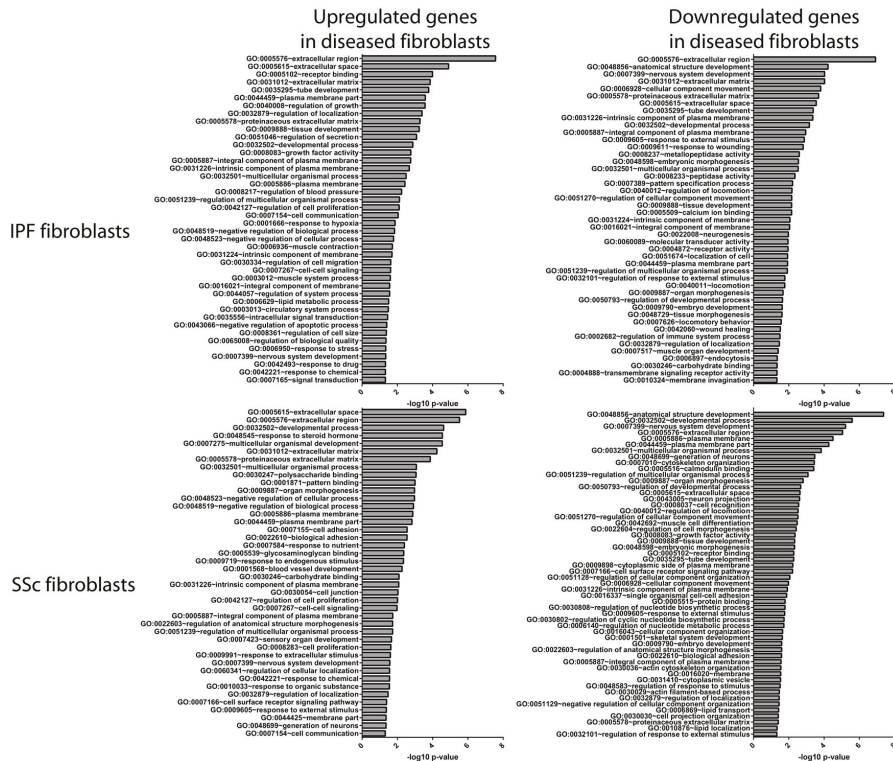


Figure 3. IPF and SSc fibroblast signatures are primarily enriched for genes associated with the extracellular matrix. Gene ontology (GO) analysis was conducted to identify over-represented GO terms for the genes upregulated in IPF or SSc fibroblasts (left panel) and downregulated in IPF or SSc fibroblasts (right panel). Shown are the significantly enriched GO terms ($p \leq 0.05$, $\geq 5\%$ of genes had to be classified by a GO term), with highly similar GO terms being collapsed using REVIGO [18].

3.3. Global microRNA Profiling Identified Similar Fibrotic microRNA Signatures in Idiopathic Pulmonary Fibrosis and Systemic Sclerosis Lung Fibroblasts

In order to get a more complete picture of the transcriptional profiles of IPF and SSc lung fibroblasts and to identify potentially novel mechanisms by which their expression programs are regulated, we analyzed their miRNA expression using miRNA-seq.

miRNA-seq revealed miRNA expression differences between normal and disease fibroblasts that were relatively moderate (generally <2-fold), which is typical for miRNA expression data. Because of these moderate changes, we utilized a less stringent expression cutoff to minimize the possibility of false negatives. Compared to control fibroblasts, IPF fibroblasts exhibited 3 upregulated and

16 downregulated miRNAs, whereas there were 12 upregulated and 12 downregulated miRNAs for SSc fibroblasts (± 1.35 -fold, $p < 0.1$, Figure 4A, Table S5). Analogous to the mRNA expression data, the IPF and SSc miRNA signatures were very similar to one another, with ~half of the differentially expressed miRNAs overlapping between them (Figure 4C). Interestingly, as was observed for mRNA expression, there were also minor differences in miRNA expression between SSc and IPF fibroblasts, with 6 miRNAs having higher and 1 miRNA having lower expression in SSc fibroblasts. To increase our confidence in these modest differences observed, we performed a second miRNA-seq and the technical replicate demonstrated high reproducibility of the data, and qPCR analysis further verified differential expression of several miRNAs (Figure 4B).

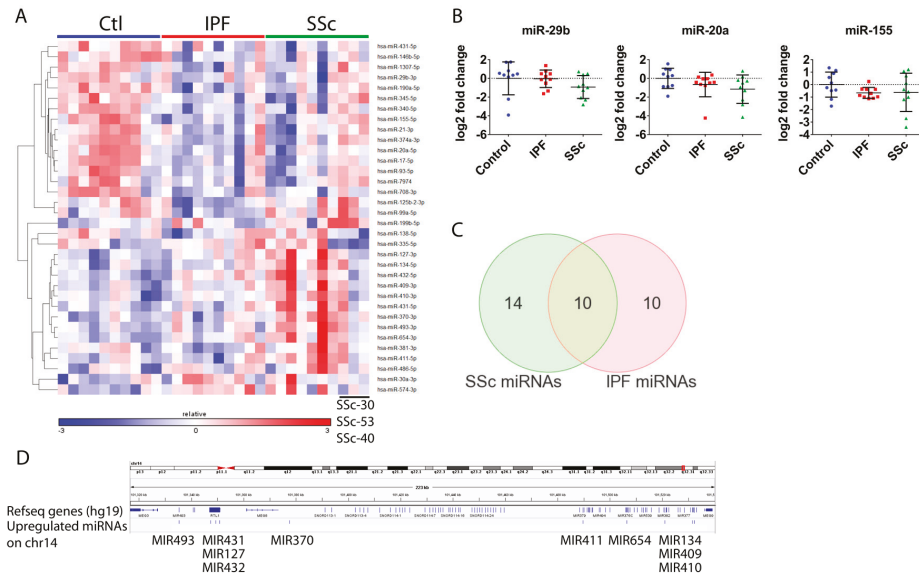


Figure 4. miRNA-seq identified similar miRNA signatures for IPF and SSc lung fibroblasts. (A) Heatmap depicting differentially expressed miRNAs across all patient group comparisons (fold-change $\geq +1.35$ -fold or ≤ -1.35 -fold, $p < 0.1$). (B) miRNA expression analysis using real-time qPCR on the 30 patient fibroblast samples was conducted as described in Materials and Methods for miR-20a, miR-155 and miR-29b in order to demonstrate similar expression changes as observed by miRNA-seq. (C) Venn diagram depicting overlap between miRNAs that were differentially expressed in SSc and IPF fibroblasts. (D) Genome view (hg19) of the Chr14q32 region where a cluster of 10 miRNAs were observed to be upregulated in SSc fibroblasts. Annotated Refseq genes are shown with the location of the indicated upregulated miRNAs marked directly below.

The resulting miRNA signature contained many dysregulated miRNAs with either established links to fibrosis (e.g., miR-29b and the miR-17~92 cluster), or affecting pathways and processes relevant to IPF and SSc fibroblasts (e.g., TGF- β) (Table S5). Additionally, a number of the miRNAs upregulated in SSc fibroblasts have previously been reported to be upregulated in IPF lung tissue (several miRNAs in the Chr14q32 miRNA cluster, Figure 4D).

3.4. Differentially Expressed microRNAs in Idiopathic Pulmonary Fibrosis and Systemic Sclerosis Fibroblasts Regulate Expression of Fibrosis-Associated Genes

To characterize the functional relevance of these dysregulated miRNAs to fibroblast/myofibroblast pathology, we examined whether modulating several miRNAs (miR-29b-3p, miR-138-5p, and miR-146b-5p) in disease fibroblasts would affect expression of ECM and other profibrotic genes. These miRNAs

were chosen either because of their extensive associations with fibrotic disease/pathways/ECM (miR-29b-3p) or because their role in fibrosis is relatively uncharacterized, yet they were among the most highly downregulated miRNAs in disease fibroblasts (miR-138-5p, miR-146b-5p).

Transfection of miR-29b-3p, miR-146b-5p, or miR-138-5p mimics into IPF and SSc lung fibroblasts all had significant effects on ECM and fibrosis gene expression using a “fibrosis” nanostring panel, with 175 genes being modulated by at least one of the miRNAs (Figure 5). As the genes in this panel are generally pro-fibrotic, these miRNA mimics resulted mostly in their downregulation in both IPF and SSc fibroblasts.

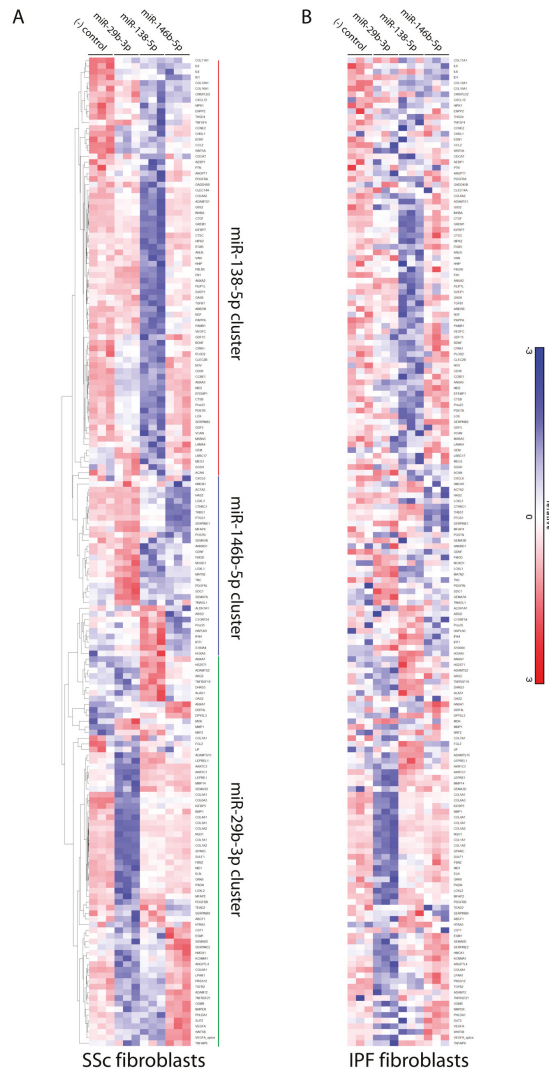


Figure 5. miR-29b-3p, miR-138-5p, and miR-146b-5p mimics downregulate expression of pro-fibrotic genes. miRNA mimics for miR-29b-3p, miR-138-5p, or miR-146b-5p were transfected into primary SSc (A) or IPF (B) lung fibroblasts as described in Materials and Methods for 48 hours prior to collecting RNA. Gene expression was measured using Nanostring analysis of a panel of 500+ “fibrosis” genes. Shown is a heatmap depicting expression of genes whose expression was significantly affected ($\geq +1.5$ -fold or ≤ -1.5 -fold, $p \leq 0.05$) by at least one of the miRNA mimics in either the SSc or IPF fibroblasts.

As we have gene and miRNA expression from parallel samples, we next aimed to determine if these miRNAs could contribute to the dysregulated fibrotic signature we observed in IPF and SSc fibroblasts. The fibrosis nanostring panel includes 46 genes that were also upregulated in the IPF and SSc fibroblast gene signatures. Of these, 30/46 (65%) were affected by at least one of the miRNA mimics in either SSc or IPF fibroblasts (Figure S4), and in nearly all cases the mimic resulted in downregulation.

Interestingly, each miRNA preferentially affected distinct clusters of genes (Figure 5 and Figure S4). As expected for miR-29b-3p, a number of pro-fibrotic and ECM genes including *COL1A1*, a well-studied target of miR-29, were downregulated. miR-138-5p also reversed expression of a large subset of the disease fibroblast signature, including key pro-fibrotic players such as *LOX*, *CTGF*, and *GREM1*. miR-146b-5p had a more subtle effect on downstream gene expression, however it was the only mimic that significantly reduced *ACTA2* levels in SSc fibroblasts.

3.5. Proteomic Profiling Characterized a Distinct Fibrotic Matrisome Deposited by Idiopathic Pulmonary Fibrosis and Systemic Sclerosis Lung Fibroblasts

Aberrant secretion of matrisome and matrisome-associated proteins is both the primary means and end outcome by which fibroblasts drive disease pathology in IPF and SSc ILD. This was reflected in our mRNA profiling data, where ECM/matrisome components were the most prevalent within the fibrotic gene signatures, and in our miRNA-seq data, in which several miRNAs affected matrisome gene expression. As there are additional mechanisms beyond transcriptional regulation that affect ECM production/deposition, we characterized directly the matrisome deposited by IPF, SSc, and normal fibroblasts at the protein level.

Extracellular matrix deposited by the 30 patient fibroblasts ($n = 10$ for each group) was collected and guanidine-soluble and insoluble fractions were subjected to mass spectrometric proteomic profiling separately, as the insoluble fraction is thought to contain more highly cross-linked ECM that could be more relevant in the context of fibrotic disease. In total, 277 matrisome proteins were detected in at least one of the samples. Importantly, using the detected matrisome proteins in the soluble fraction, hierarchical clustering of the fibroblasts resulted in a distinct clustering of IPF and SSc fibroblasts away from controls, strongly suggesting that disease fibroblasts secrete a distinct fibrotic matrisome (Figure S5A). While similar separation was not observed for the insoluble fraction (Figure S5B), the protein recovery for the insoluble fraction was highly variable between different samples, which may have confounded our analysis.

Across all fibroblast group comparisons, the majority of differences were observed between disease vs. normal fibroblasts, with SSc vs. Normal yielding 26 and 18 differentially expressed proteins (DEPs) for the soluble and insoluble fractions respectively and IPF vs. Normal yielding 28 and 6 DEPs for the soluble and insoluble fractions respectively (Figure S5C). Analogous to what we observed at the transcriptional level, the IPF and SSc matrisomes tended to be more similar than different as demonstrated by the observed DEP overlap (Figure S5D) and by the observation that IPF and SSc fibroblasts generally clustered together (Figure S5A and Figure 6). Interestingly we did identify minor differences between SSc and IPF matrisome protein expression, with 8 and 2 DEPs being observed for the soluble and insoluble fractions respectively. Hierarchical clustering using DEPs across all cohort comparisons resulted in a cluster of ~half of the disease fibroblasts when using either the soluble (Figure 6A) or insoluble (Figure 6B) fraction data, demonstrating that disease fibroblasts secrete a distinct matrisome compared to normal fibroblasts.

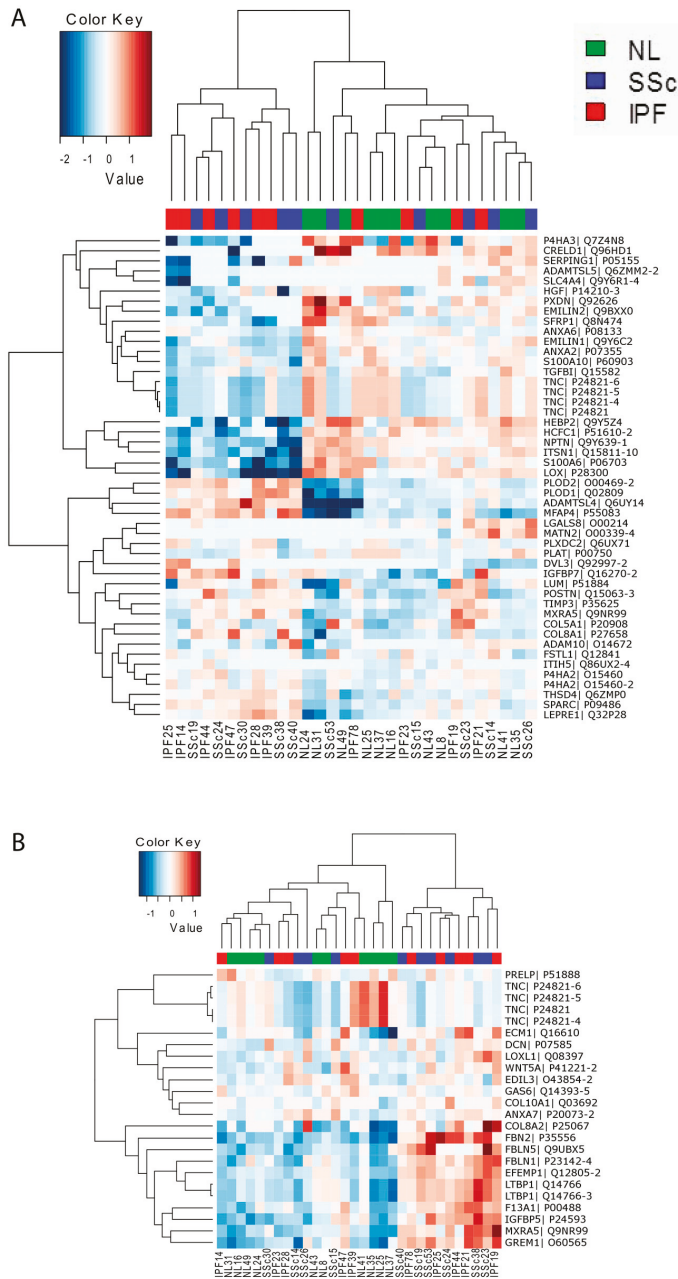


Figure 6. Hierarchical clustering using differentially expressed matrisome proteins is able to delineate normal-like and disease-like fibroblast groups. Shown are heatmaps of z-scored normalized mass spectrometry protein intensity values of matrisome proteins that were differentially expressed across all patient comparisons (fold-change $\geq +1.2$ -fold or ≤ -1.2 -fold, $p \leq 0.05$) and then clustered. Shown is the data from the soluble fraction (A) and from the insoluble fraction (B).

4. Discussion

Fibroblasts/myofibroblasts are key effector cells that directly contribute to the pro-fibrotic milieu, aberrant ECM deposition, increased tissue stiffening, disruption of tissue architecture, and ultimately impaired organ function in pulmonary fibrosis [2]. Understanding the transcriptomic and proteomic profiles of these disease fibroblasts provides insights into their dysregulated fibrotic program at multiple levels and sheds light on potential therapeutic strategies. To our knowledge, the current study is the first attempt at integrating data from mRNA, miRNA, and secreted matrisome of IPF and SSc derived lung fibroblasts and our results revealed a number of novel mechanistic insights into the fibrotic programming of these effector cells underlying pulmonary fibrosis and its complex regulation.

Consistent with previous microarray-based studies, our RNA-seq clearly identified similar “fibrotic gene signatures” for IPF and SSc fibroblasts that reflect a signature of activated myofibroblasts, as illustrated by the correlation for genes upregulated in activated hepatic stellate cells (HSCs) with both the IPF and SSc DEGs (Figure S3). Further computational analysis revealed that these signatures are associated with known fibrotic pathways (WNT, TGF- β , HIF1A, NOTCH1, and PPARG) and effector functions (ECM) of activated fibroblasts/myofibroblasts.

Although the fibrotic transcriptomes from our and previous transcriptomic analyses of fibrotic lung fibroblasts are indicative of pathological myofibroblast activation, our current study also yielded interesting novel findings. First, the exact gene makeup of the reported signatures varied considerably among studies [3–7], which could result from study-specific differences related to patient profiles, source of tissue, culture methods, and profiling methods. However, disease stage is likely a significant driver as we observed stark differences compared to the study reported by Lindahl et al. [6], which used fibroblasts from earlier stage patients, as opposed to end-stage patients in our study. While their study identified a broader set of pathway signatures that were both inflammatory and fibrotic in nature, ECM changes were the most predominant feature in our disease signatures. This suggests that inflammatory changes, such as downregulation of the interferon signature, may be important for progression during early but not at later stages of disease when ECM changes predominate. Second, while previous studies have observed only minimal expression differences between IPF and SSc fibroblasts, our study successfully identified 68 genes that were differentially expressed between them, including inflammation genes (*TNFRSF21*, *CXCL5*, *IL8*) and genes associated with the GO function “oxidoreductase activity” (Figure S2). Interestingly, these differences appear to be driven by a subset of SSc patients (SSc-53, SSc-40, SSc-30). While the precise reason for this is not understood, it could be reflective of concomitant pathological changes frequently associated with SSc, such as pulmonary hypertension and inflammation.

The general similarity between IPF and SSc transcriptomes with subtle differences was also mirrored in the miRNA and matrisome data. Interestingly, for the three SSc patients noted above, we also observed that their miRNA expression tended to differ from the other SSc patients (Figure 4), raising the intriguing possibility that SSc-associated expression changes, such as those that might reflect pathological changes other than fibrosis (e.g., hypertension), could be governed in part at the miRNA level. At the protein level, the secreted matrisome also exhibited a few differences between IPF and SSc. However, unlike the mRNA and miRNA expression data, those three SSc patients did not appear to exhibit distinct matrix protein profiles compared to the other SSc fibroblasts, possibly because the matrisome signature is more reflective of fibrosis and less likely to be indicative of other SSc-associated pathologies.

As the most salient effector mechanism underlying pulmonary fibrosis, ECM deposition and remodeling by IPF and SSc fibroblasts remain poorly understood. Although the ECM/matrisome transcriptomic signature was the most predominant signal altered in both IPF and SSc fibroblasts, matrix genes were not universally upregulated and in fact many were downregulated (Figure 2A). Thus our data indicate that the fibrotic matrisome does not just result from increased expression of ECM genes but it involves complex dysregulated expression patterns, including both increased and decreased mRNA expression of different collagen types (*COL1A1* upregulation vs. *COL14A1*

downregulation), downregulation of ECM-degradation enzymes (*MMP1* and *ADAMTS15*), and upregulation of ECM crosslinking/assembly enzymes (*LOX*, *LEPREL1* and *PLOD2*). Importantly, we confirmed aberrant dysregulation of ECM at the protein level as well by conducting the first proteomic analysis of IPF and SSc fibroblast-deposited matrix. We successfully identified an ECM protein signature shared between IPF and SSc that differentiates disease from healthy fibroblasts, which included several proteins implicated in fibrogenesis such as *PLOD2*, *LUM*, *POSTN*, *IGFBP5*, *GREM1*, and *SPARC*, as well as less characterized ECM proteins such as *MXRA5*, *LEPRE1*, *MFAP4*, and *FSTL1*. Upon comparing our RNA-seq and mass spectrometry results, we observed similar dysregulation at the protein and mRNA levels for such matrisome components as *WNT5A*, *GREM1*, *DCN*, *IGFBP5*, *COL8A1*, *PLOD2*, *SFRP1* and *TNC*. Surprisingly, these were the only shared dysregulated genes/proteins between the RNA-seq and mass spectrometry data suggesting that differences between protein and mRNA levels of matrisome components are likely due to the presence of a myriad of post-transcriptional regulatory mechanisms at the levels of protein translation, secretion, intracellular and extracellular assembly, as well as enzymatic crosslinking and degradation. However we cannot rule out that the observed differences between the RNA-seq and mass spectrometry results are due to differences in the length of time the fibroblasts were cultured for each experiment, (short-term for RNA-seq, long-term to accumulate sufficient deposited matrix for mass-spectrometry experiments) or that exposure of these cells to tissue culture plastic influenced their protein expression profile.

miRNAs can affect both mRNA levels and protein translation, and a handful of miRNAs have been identified as key players in lung fibrosis. Our current study is the first comprehensive characterization of the global miRNA profiles in fibrotic lung fibroblasts, which included the identification of miRNAs capable of regulating ECM gene expression. Our miRNA-seq analysis identified a number of aberrantly expressed miRNAs in IPF and SSc fibroblasts and several lines of evidence support the relevance of these miRNAs in their fibrotic programming. In particular, the miRNA signature includes many miRNAs previously linked to fibrosis through in vitro and in vivo studies (Table S5). Of particular interest, miR-29b-3p has reduced expression in multiple fibrosis models and human fibrotic disease and can inhibit fibrosis in several mouse models [19–32]. miR-138-5p and miR-146b-5p, which demonstrated some of the highest levels of downregulation in the disease fibroblasts, have been implicated in processes relevant to fibrosis. For example, miR-138-5p has been suggested to play a role in hypertrophic scar fibroblasts during abnormal wound healing [33], osteogenic differentiation and bone formation [34,35], and epithelial-mesenchymal transition [36]. In addition to miR-146b-5p's effect on TGF- β -signaling [37], its closely-related family member miR-146a inhibits TGF- β -mediated activation of dermal fibroblasts and HSCs, as well as renal fibrosis in the unilateral ureteral obstruction model [38–40]. Importantly, we also demonstrated experimentally that several “fibrotic miRNAs” identified in our study (miR-29b-3p, miR-138-5p and miR-146b-5p) regulate the expression of fibrotic/ECM genes. Their identification as important miRNAs modulating distinct aspects of the fibrotic transcriptomes in IPF and SSc lung fibroblasts is highly significant in our view, as this further supports a role for miR-29 as a master ECM regulator, and reveals novel roles for miR-138-5p and miR-146b-5p in pulmonary fibrosis. Since we only measured their effects on transcript levels, our data understates the potential contribution these miRNAs have in the pathology of disease fibroblasts as they could have additional effects on the deposited fibrotic matrisome signature via translation inhibition. It is important to note that these dysregulated miRNAs could be affecting fibrotic/ECM genes either directly or indirectly through other transcriptional regulators.

Equally intriguing is our finding that 10 miRNAs in the Chr14q32 locus appear to be a fibrotic “miRNA module” in SSc fibroblasts, similar to the previously described Chr14q32 region in IPF lung tissue [41]. Additional miRNA clusters, whereby miRNAs/genes exhibited similar differential gene expression and genomic localization, were identified for miR-17-5p and miR-20a-5p, miR-125b-2-3p and miR-99a-5p, and for miR-335-5p and *MEST*. This implies that miRNA modules are dysregulated in IPF and SSc fibroblasts and indeed we see significant expression level correlations for the genes within these clusters (data not shown).

Altogether, through an integrative approach, we successfully characterized distinct mRNA, miRNA, and deposited matrix protein signatures for IPF and SSc fibroblasts, and in doing so identified novel-regulation of fibrotic gene expression by aberrantly expressed miRNAs.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4425/9/12/588/s1>, Supplementary Text: Supplementary Materials and Methods, Figure S1: Gene expression values are highly correlative between RNA-seq and qPCR, Figure S2: RNA-seq identifies 68 DEGs between IPF and SSc lung fibroblasts, Figure S3: IPF and SSc fibroblast gene expression profiles reflect activated hepatic stellate cell (HSC) and TGF- β gene signatures, Figure S4: miR-29b-3p, miR-138-5p, and miR-146b-5p mimics downregulate expression of pro-fibrotic genes that are dysregulated in disease fibroblasts, Figure S5: Proteomic profiling of ECM deposited by IPF and SSc lung fibroblasts reveals a fibrotic matrisome signature that delineates them from normal fibroblasts, Table S1: List of Taqman probes and miRNA mimics used, Table S2: Patient information summary, Table S3: Individual patient information, Table S4: RNAseq expression data for DEGs across all cohort comparisons, Table S5: Differentially expressed miRNAs across all patient lung fibroblast comparisons.

Author Contributions: Conceptualization, S.M., T.S.Z., F.L., and C.-F.B. Investigation, S.M., F.L., X.H., M.L., H.M., and Q.X. Formal Analysis, S.M., F.L., S.S., B.G., C.S., and P.J. Resources, C.-F.B. Writing—Original Draft Preparation, S.M. and T.S.Z. Writing—Review & Editing, all authors.

Funding: This research was supported by Biogen, Cambridge, MA, USA.

Conflicts of Interest: All authors other than C.F.B. were employed by Biogen and may own Biogen stock.

References

- Herzog, E.L.; Mathur, A.; Tager, A.M.; Feghali-Bostwick, C.; Schneider, F.; Varga, J. Review: interstitial lung disease associated with systemic sclerosis and idiopathic pulmonary fibrosis: how similar and distinct? *Arthritis Rheumatol.* **2014**, *66*, 1967–1978. [[CrossRef](#)] [[PubMed](#)]
- Kendall, R.T.; Feghali-Bostwick, C.A. Fibroblasts in fibrosis: novel roles and mediators. *Front. Pharmacol.* **2014**, *5*, 123. [[CrossRef](#)] [[PubMed](#)]
- Emblom-Callahan, M.C.; Chhina, M.K.; Shlobin, O.A.; Ahmad, S.; Reese, E.S.; Iyer, E.P.; Cox, D.N.; Brenner, R.; Burton, N.A.; Grant, G.M.; et al. Genomic phenotype of non-cultured pulmonary fibroblasts in idiopathic pulmonary fibrosis. *Genomics* **2010**, *96*, 134–145. [[CrossRef](#)] [[PubMed](#)]
- Hsu, E.; Shi, H.; Jordan, R.M.; Lyons-Weiler, J.; Pilewski, J.M.; Feghali-Bostwick, C.A. Lung tissues in patients with systemic sclerosis have gene expression patterns unique to pulmonary fibrosis and pulmonary hypertension. *Arthritis Rheum.* **2011**, *63*, 783–794. [[CrossRef](#)] [[PubMed](#)]
- Kabuyama, Y.; Oshima, K.; Kitamura, T.; Homma, M.; Yamaki, J.; Munakata, M.; Homma, Y. Involvement of selenoprotein P in the regulation of redox balance and myofibroblast viability in idiopathic pulmonary fibrosis. *Genes Cells* **2007**, *12*, 1235–1244. [[CrossRef](#)] [[PubMed](#)]
- Lindahl, G.E.; Stock, C.J.; Shi-Wen, X.; Leoni, P.; Sestini, P.; Howat, S.L.; Bou-Gharios, G.; Nicholson, A.G.; Denton, C.P.; Grutters, J.C.; et al. Microarray profiling reveals suppressed interferon stimulated gene program in fibroblasts from scleroderma-associated interstitial lung disease. *Respir. Res.* **2013**, *14*, 80. [[CrossRef](#)] [[PubMed](#)]
- Peng, R.; Sridhar, S.; Tyagi, G.; Phillips, J.E.; Garrido, R.; Harris, P.; Burns, L.; Renteria, L.; Woods, J.; Chen, L.; et al. Bleomycin induces molecular changes directly relevant to idiopathic pulmonary fibrosis: a model for “active” disease. *PLoS ONE* **2013**, *8*, e59348. [[CrossRef](#)] [[PubMed](#)]
- Kass, D.J.; Kaminski, N. Evolving genomic approaches to idiopathic pulmonary fibrosis: moving beyond genes. *Clin. Transl. Sci.* **2011**, *4*, 372–379. [[CrossRef](#)] [[PubMed](#)]
- Christmann, R.B.; Sampaio-Barros, P.; Stifano, G.; Borges, C.L.; de Carvalho, C.R.; Kairalla, R.; Parra, E.R.; Spira, A.; Simms, R.; Capellozzi, V.L.; et al. Association of Interferon- and transforming growth factor beta-regulated genes and macrophage activation with systemic sclerosis-related progressive lung fibrosis. *Arthritis Rheumatol.* **2014**, *66*, 714–725. [[CrossRef](#)] [[PubMed](#)]
- Parker, M.W.; Rossi, D.; Peterson, M.; Smith, K.; Sikström, K.; White, E.S.; Connett, J.E.; Henke, C.A.; Larsson, O.; Bitterman, P.B. Fibrotic extracellular matrix activates a profibrotic positive feedback loop. *J. Clin. Invest.* **2014**, *124*, 1622–1635. [[CrossRef](#)] [[PubMed](#)]
- Huang da, W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4*, 44–57. [[CrossRef](#)] [[PubMed](#)]

12. Kupershmidt, I.; Su, Q.J.; Grewal, A.; Sundaresh, S.; Halperin, I.; Flynn, J.; Shekar, M.; Wang, H.; Park, J.; Cui, W.; et al. Ontology-based meta-analysis of global collections of high-throughput public data. *PLoS ONE* **2010**, *5*. [[CrossRef](#)] [[PubMed](#)]
13. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [[CrossRef](#)] [[PubMed](#)]
14. Didangelos, A.; Yin, X.; Mandal, K.; Baumert, M.; Jahangiri, M.; Mayr, M. Proteomics characterization of extracellular space components in the human aorta. *Mol. Cell Proteomics* **2010**, *9*, 2048–2062. [[CrossRef](#)] [[PubMed](#)]
15. Decaris, M.L.; Gatmaitan, M.; FlorCruz, S.; Luo, F.; Li, K.; Holmes, W.E.; Hellerstein, M.K.; Turner, S.M.; Emson, C.L. Proteomic analysis of altered extracellular matrix turnover in bleomycin-induced pulmonary fibrosis. *Mol. Cell Proteomics* **2014**, *13*, 1741–1752. [[CrossRef](#)] [[PubMed](#)]
16. Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26*, 1367–1372. [[CrossRef](#)] [[PubMed](#)]
17. Naba, A.; Clauser, K.R.; Hoersch, S.; Liu, H.; Carr, S.A.; Hynes, R.O. The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Mol. Cell Proteomics* **2012**, *11*, M111 014647. [[CrossRef](#)] [[PubMed](#)]
18. Supek, F.; Bosnjak, M.; Skunca, N.; Smuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* **2011**, *6*, e21800. [[CrossRef](#)] [[PubMed](#)]
19. He, Y.; Huang, C.; Lin, X.; Li, J. MicroRNA-29 family, a crucial therapeutic target for fibrosis diseases. *Biochimie* **2013**, *95*, 1355–1359. [[CrossRef](#)] [[PubMed](#)]
20. Montgomery, R.L.; Yu, G.; Latimer, P.A.; Stack, C.; Robinson, K.; Dalby, C.M.; Kaminski, N.; van Rooij, E. MicroRNA mimicry blocks pulmonary fibrosis. *EMBO Mol. Med.* **2014**, *6*, 1347–1356. [[CrossRef](#)] [[PubMed](#)]
21. Xiao, J.; Meng, X.M.; Huang, X.R.; Chung, A.C.; Feng, Y.L.; Hui, D.S.; Yu, C.M.; Sung, J.J.; Lan, H.Y. miR-29 inhibits bleomycin-induced pulmonary fibrosis in mice. *Mol. Ther.* **2012**, *20*, 1251–1260. [[CrossRef](#)] [[PubMed](#)]
22. Cushing, L.; Kuang, P.P.; Qian, J.; Shao, F.; Wu, J.; Little, F.; Thannickal, V.J.; Cardoso, W.V.; Lu, J. miR-29 is a major regulator of genes associated with pulmonary fibrosis. *Am. J. Respir. Cell Mol. Biol.* **2011**, *45*, 287–294. [[CrossRef](#)] [[PubMed](#)]
23. Qin, W.; Chung, A.C.; Huang, X.R.; Meng, X.M.; Hui, D.S.; Yu, C.M.; Sung, J.J.; Lan, H.Y. TGF-beta/Smad3 signaling promotes renal fibrosis by inhibiting miR-29. *J. Am. Soc. Nephrol.* **2011**, *22*, 1462–1474. [[CrossRef](#)] [[PubMed](#)]
24. Zhang, Y.; Huang, X.R.; Wei, L.H.; Chung, A.C.; Yu, C.M.; Lan, H.Y. miR-29b as a therapeutic agent for angiotensin II-induced cardiac fibrosis by targeting TGF-beta/Smad3 signaling. *Mol. Ther.* **2014**, *22*, 974–985. [[CrossRef](#)] [[PubMed](#)]
25. van Rooij, E.; Sutherland, L.B.; Thatcher, J.E.; DiMaio, J.M.; Naseem, R.H.; Marshall, W.S.; Hill, J.A.; Olson, E.N. Dysregulation of microRNAs after myocardial infarction reveals a role of miR-29 in cardiac fibrosis. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 13027–13032. [[CrossRef](#)] [[PubMed](#)]
26. Roderburg, C.; Urban, G.W.; Bettermann, K.; Vucur, M.; Zimmermann, H.; Schmidt, S.; Janssen, J.; Koppe, C.; Knolle, P.; Castoldi, M.; et al. Micro-RNA profiling reveals a role for miR-29 in human and murine liver fibrosis. *Hepatology* **2011**, *53*, 209–218. [[CrossRef](#)] [[PubMed](#)]
27. Sekiya, Y.; Ogawa, T.; Yoshizato, K.; Ikeda, K.; Kawada, N. Suppression of hepatic stellate cell activation by microRNA-29b. *Biochem. Biophys. Res. Commun.* **2011**, *412*, 74–79. [[CrossRef](#)] [[PubMed](#)]
28. Zhang, Y.; Wu, L.; Wang, Y.; Zhang, M.; Li, L.; Zhu, D.; Li, X.; Gu, H.; Zhang, C.Y.; Zen, K. Protective role of estrogen-induced miRNA-29 expression in carbon tetrachloride-induced mouse liver injury. *J. Biol. Chem.* **2012**, *287*, 14851–14862. [[CrossRef](#)] [[PubMed](#)]
29. Maurer, B.; Stanczyk, J.; Jungel, A.; Akhmetshina, A.; Trenkmann, M.; Brock, M.; Kowal-Bielecka, O.; Gay, R.E.; Michel, B.A.; Distler, J.H.; et al. MicroRNA-29, a key regulator of collagen expression in systemic sclerosis. *Arthritis Rheum.* **2010**, *62*, 1733–1743. [[CrossRef](#)] [[PubMed](#)]
30. Cushing, L.; Kuang, P.; Lu, J. The role of miR-29 in pulmonary fibrosis. *Biochem. Cell Biol.* **2015**, *93*, 109–118. [[CrossRef](#)] [[PubMed](#)]

31. Pandit, K.V.; Milosevic, J. MicroRNA regulatory networks in idiopathic pulmonary fibrosis. *Biochem. Cell Biol.* **2015**, *93*, 129–137. [[CrossRef](#)] [[PubMed](#)]
32. Pandit, K.V.; Corcoran, D.; Yousef, H.; Yarlagadda, M.; Tzouveleakis, A.; Gibson, K.F.; Konishi, K.; Yousem, S.A.; Singh, M.; Handley, D.; et al. Inhibition and role of let-7d in idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **2010**, *182*, 220–229. [[CrossRef](#)] [[PubMed](#)]
33. Xiao, Y.Y.; Fan, P.J.; Lei, S.R.; Qi, M.; Yang, X.H. MiR-138/peroxisome proliferator-activated receptor beta signaling regulates human hypertrophic scar fibroblast proliferation and movement in vitro. *J. Dermatol.* **2015**, *42*, 485–495. [[CrossRef](#)] [[PubMed](#)]
34. Qu, B.; Xia, X.; Wu, H.H.; Tu, C.Q.; Pan, X.M. PDGF-regulated miRNA-138 inhibits the osteogenic differentiation of mesenchymal stem cells. *Biochem. Biophys. Res. Commun.* **2014**, *448*, 241–247. [[CrossRef](#)] [[PubMed](#)]
35. Eskildsen, T.; Taipaleenmaki, H.; Stenvang, J.; Abdallah, B.M.; Ditzel, N.; Nossent, A.Y.; Bak, M.; Kauppinen, S.; Kassem, M. MicroRNA-138 regulates osteogenic differentiation of human stromal (mesenchymal) stem cells in vivo. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6139–6144. [[CrossRef](#)] [[PubMed](#)]
36. Liu, X.; Wang, C.; Chen, Z.; Jin, Y.; Wang, Y.; Kolokythas, A.; Dai, Y.; Zhou, X. MicroRNA-138 suppresses epithelial-mesenchymal transition in squamous cell carcinoma cell lines. *Biochem. J.* **2011**, *440*, 23–31. [[CrossRef](#)] [[PubMed](#)]
37. Geraldo, M.V.; Yamashita, A.S.; Kimura, E.T. MicroRNA miR-146b-5p regulates signal transduction of TGF-beta by repressing SMAD4 in thyroid cancer. *Oncogene* **2012**, *31*, 1910–1922. [[CrossRef](#)] [[PubMed](#)]
38. He, Y.; Huang, C.; Sun, X.; Long, X.R.; Lv, X.W.; Li, J. MicroRNA-146a modulates TGF-beta1-induced hepatic stellate cell proliferation by targeting SMAD4. *Cell Signal* **2012**, *24*, 1923–1930. [[CrossRef](#)] [[PubMed](#)]
39. Liu, Z.; Lu, C.L.; Cui, L.P.; Hu, Y.L.; Yu, Q.; Jiang, Y.; Ma, T.; Jiao, D.K.; Wang, D.; Jia, C.Y. MicroRNA-146a modulates TGF-beta1-induced phenotypic differentiation in human dermal fibroblasts by targeting SMAD4. *Arch. Dermatol. Res.* **2012**, *304*, 195–202. [[CrossRef](#)] [[PubMed](#)]
40. Morishita, Y.; Imai, T.; Yoshizawa, H.; Watanabe, M.; Ishibashi, K.; Muto, S.; Nagata, D. Delivery of microRNA-146a with polyethylenimine nanoparticles inhibits renal fibrosis in vivo. *Int. J. Nanomedicine* **2015**, *10*, 3475–3488. [[CrossRef](#)] [[PubMed](#)]
41. Milosevic, J.; Pandit, K.; Magister, M.; Rabinovich, E.; Ellwanger, D.C.; Yu, G.; Vuga, L.J.; Weksler, B.; Benos, P.V.; Gibson, K.F.; et al. Profibrotic role of miR-154 in pulmonary fibrosis. *Am. J. Respir. Cell Mol. Biol.* **2012**, *47*, 879–887. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

An Exonic Switch Regulates Differential Accession of microRNAs to the Cd34 Transcript in Atherosclerosis Progression

Miguel Hueso ^{1,*}, Josep M. Cruzado ¹, Joan Torras ¹ and Estanis Navarro ^{2,*}

¹ Department of Nephrology, Hospital Universitari Bellvitge and Bellvitge Research Institute (IDIBELL), L'Hospitalet de Llobregat, 08907 Barcelona, Spain; jmcruzado@bellvitgehospital.cat (J.M.C.); 15268jta@comb.cat (J.T.)

² Independent Researcher, Esplugues de Llobregat, 08950 Barcelona, Spain

* Correspondence: mhueso@idibell.cat (M.H.); estanis.navarro@gmail.com (E.N.); Tel.: +34-932607602 (M.H.)

Received: 18 December 2018; Accepted: 14 January 2019; Published: 21 January 2019

Abstract: Background: CD34⁺ Endothelial Progenitor Cells (EPCs) play an important role in the recovery of injured endothelium and contribute to atherosclerosis (ATH) pathogenesis. Previously we described a potential atherogenic role for miR-125 that we aimed to confirm in this work. Methods: Microarray hybridization, TaqMan Low Density Array (TLDA) cards, qPCR, and immunohistochemistry (IHC) were used to analyze expression of the miRNAs, proteins and transcripts here studied. Results: Here we have demonstrated an increase of resident CD34-positive cells in the aortic tissue of human and mice during ATH progression, as well as the presence of clusters of CD34-positive cells in the intima and adventitia of human ATH aortas. We introduce miR-351, which share the seed sequence with miR-125, as a potential effector of CD34. We show a splicing event at an internal/cryptic splice site at exon 8 of the murine *Cd34* gene (exonic-switch) that would regulate the differential accession of miRNAs (including miR-125) to the coding region or to the 3'UTR of *Cd34*. Conclusions: We introduce new potential mediators of ATH progression (CD34 cell-clusters, miR-351), and propose a new mechanism of miRNA action, linked to a cryptic splicing site in the target-host gene, that would regulate the differential accession of miRNAs to their cognate binding sites.

Keywords: atherosclerosis; miR-125b; CD34 cell-clusters; CD34 cryptic splicing; ORF miRNA target; exon switch

1. Background

Atherosclerosis (ATH) is a chronic inflammatory disease initiated at the vascular endothelium which shows a complex pathogenesis [1]. Endothelium damage displays a focal distribution, particularly at sites of disturbed blood flow, suggesting the contribution of vascular repair mechanisms in ATH progression. Recent studies suggest that endothelial progenitor cells (EPCs) do have an important role in the recovery and repair of the injured endothelium [2], since mobilized EPCs are able to migrate to damaged sites where they would differentiate into mature endothelial cells (ECs) in situ [3]. EPCs come from three major sources; i.e., peripheral blood and bone marrow, the vascular wall, and resident monocytes and macrophages [4,5]. In patients with ATH it has been observed a decrease in the number of circulant EPCs [6], although it is not well known if this is the result from a decreased production or from an increased EPC cell death [7].

Bone marrow-derived EPCs express the progenitor receptor CD34 and are released into the peripheral blood for vascular repair/angiogenesis, since these have reparative potential of endothelial dysfunction [8]. CD34, a membrane glycoprotein, is a critical component of the group of surface

receptors that regulate migration and engraftment of progenitor cells to target tissues [9]. *CD34* is expressed as two variant forms after alternative splicing of the so-called “exon-X” at the *CD34* gene [10]. These two forms differ in their cytoplasmic domain, which is almost totally lacking in the variant-1, a fact that abrogates phosphorylation by protein kinase C (PKC) at two sites in the Cterminal cytoplasmic domain [11]. Interestingly, bioinformatic analyses have detected a cluster of miRNA binding sites at the putative 3'UTR of the human and murine *CD34* (Hueso, M., this work) that could be involved in the regulation of the expression balance among both isoforms.

MicroRNAs (miRNAs) are very small RNAs (over 22 nucleotides long) with critical roles in the regulation of gene expression and whose changes in expression have been related to the onset and progression of different diseases (see [12] for a recent review). As of March 2018, the human miRNAome included 1917 mature miRNAs [13]. Aberrant miRNA expression profiles have been described during the progression of ATH, cardiovascular disease (CVD) [14] or renal fibrosis [15]. Thus, plaque progression and rupture were linked to the expression of miR-23a-5p, miR-210 and miR-222 [16,17] while, on the contrary, miR-19b and miR-33a/miR-33b had protective roles on plaque stability [18,19]. Furthermore, plasma miR-144 and miR-33 levels were seen to be increased in coronary artery disease (CAD) patients [20,21], miR-155 was seen to inhibit transformation of macrophages into foam cells by targeting CEH expression [22], and miR-181b was found to be overexpressed in human atherosclerotic plaques and abdominal aortic aneurysms, where it downregulated expression of the tissue inhibitor of MMP-3, and elastin [23]. Lastly, miR-296 has been described as a positive regulator of ATH onset and progression by promoting neovascularization and favoring M1 macrophage polarization [24].

We have previously described the relationship of miR-125b expression with ATH progression. Aortic tissue from ApoE-deficient mice fed with a western diet showed a significant overexpression of miR-125b, but not of the highly related miR-125a. This overexpression was reversed by treating mice with a siRNA against the immune mediator CD40 [25]. MiR-125b belongs to a conserved family (composed by the highly related miR-125a, miR-125b-1 and miR-125b-2) with key roles in cellular differentiation, proliferation and apoptosis [26]. Furthermore, there is experimental evidence of the involvement of the miR-125 family with pathogenesis of ATH. In this sense, it has been described that up-regulation of miR-125a contributed to the differentiation of monocytes towards the inflammatory phenotype [27] through the repression of the TNFAIP3 a negative regulator of NF- κ B signaling [28], and that miR-125a was upregulated in oxLDL activated macrophages [29]. Lastly, members of the miR-125 family have been shown to target transcripts whose product could be also involved in mechanisms promoting ATH onset and progression, as STAT3 [30], FOXP3 [31], VEGF [32], MARK1 [33], and BCL2, BCL2L12 and MCL1 [34] among others.

We are interested in the mechanisms of ATH progression, and more specifically in those which involve genes that are implicated in the function of stem or progenitor cells, such as *CD34* or *Lin28a/b*. *CD34* was previously detected as a target of miR-125 in miRNA by bioinformatic analysis (M. Hueso et al., unpublished results) and here we have deepened the molecular relationship among miR-125 and *CD34* in the context of ATH progression. The most significant finding of this work is the description of a new molecular mechanism of miRNA action by which an splicing event at an internal/cryptic splice site of the murine *Cd34* gene would regulate the differential accession of a number of micro-RNAs (including miR-125) to the coding region or to the 3'UTR of the two isoforms of the *Cd34* transcript [10], likely disturbing the expression balance of both CD34 protein isoforms.

2. Methods

2.1. Ethics Statement

Here, we have used authorized autopsy material from the Pathology Department at the Hospital Universitari de Bellvitge (HUB), L'Hospitalet, Barcelona, Spain. Confidential information from patients

was protected following national normatives. This study was performed conforming the declaration of Helsinki, and with the approval by the Clinical Research Ethics Committee of HUB (PR163/13).

2.2. Patients and Samples

The characteristics of patients and samples studied in this work have been reported previously [25,35]. Briefly, abdominal aortas were obtained from deceased patients from the HUB, and included atherosclerotic plaques, incipient atherosclerotic plaques, and normal abdominal aortas without injury.

2.3. Mice

In this work we used tissue samples from 25 female ApoE^{-/-} mice (Homociguous *ApoE^{tm1Unc}* in the C57BL/6 background, Jackson Laboratory, Charles River, Wilmington, MA, USA) that were stored in a previous study [25]. Briefly, 5 mice were considered as controls and sacrificed at 8 weeks (8W). Other 20 mice were i.p. injected with 50 µg of siRNA against CD40 (siCD40), or of a scrambled siRNA (SC) as control, twice weekly. Mice were euthanized by isoflurane anesthesia after 14 or 24 weeks of treatment (SC/14w, *n* = 5; siCD40/14W, *n* = 5; SC/24W, *n* = 5; siCD40/24w, *n* = 5). Mice were fed with a western diet enriched in cholesterol (0.2% *w/w*), which provided 42% of calories as fat (TD.88137; Harlan-Tekland, Madison, WI, USA). We isolated the ascendant aortas from 10 ApoE^{-/-} mice treated with the anti-siCD40 siRNA and from 10 ApoE^{-/-} controls, 5 mice treated with the scrambled siRNA (SC) and 5 mice treated with the vehicle (veh) for 24 weeks. This study followed the EU guidelines on animal care and the protocols were approved by the Ethics Committee for Animal Research of the University of Barcelona-HUB. The effect of the anti-CD40 treatment had been previously confirmed by checking the downregulation in the expression of CD40 mRNA and the lowering in the counting of CD40-positive cells (see [25] and references therein).

2.4. Reactives

Biotinylated secondary antibodies and vectastain were from Vector (Burlingame, CA, USA). Expression of the isoforms of CD34 (Mn01310773 and Mn00519283) was measured with Taqman Gene Expression Assays (Life technologies, Thermo Fisher Scientific Inc., Waltham, MA, USA). For the anti-CD34 staining, we used an anti-CD34 antibody (RabMab EP373Y; ab81289) from Abcam (Cambridge, UK) following standard procedures (see [25] for details).

2.5. Histological and Immunohistochemical CD34 Analysis

Preparation of aortic tissue for protein/RNA expression analysis has been described in detail elsewhere [25,35]. Briefly, aortic sections (aortic arch, descending thoracic aorta, abdominal aorta above renal arteries and abdominal aorta below renal arteries) were split for RNA extraction or for IHC after fixing with 4% PFA. ATH lesions were assessed in sequential, 5 µm thick, sections of the aortic roots. stained with hematoxylin/eosin (H/E) or with a specific antibody by standard procedures. A semiquantitative evaluation of the immunostaining scored the total number of antigen-positive cells with regard of the number of total cells (nuclear H/E staining). Lastly, the ProgResCapturePro 2.7.7 software (Jenoptik, Jena, Germany) was used for morphometric image analysis.

2.6. qPCR and miRNA Expression Data

Primary murine miRNA expression data was extracted from the original Taqman Low Density Array card (TLDA) experiment on murine aortas, already described [25], by using the ExpressionSuite Software version 1.1 (Applied Biosystems-Thermo Fisher Scientific, Waltham, MA, USA). RNA extraction from aortas and cDNA conversion have been described previously [25]. Expression of human CD34 was measured by using Taqman probes in the conditions stated by the manufacturer. The $\Delta\text{Ct}/\Delta\Delta\text{Ct}$ method was used to compare results with the basal conditions samples

(8w mice). For mRNA expression normalization, the housekeeping GAPDH was used, while miRNA expression was normalized by comparison with RNU48 (human) or U6-snrRNA (mouse) also from the TLDA data.

2.7. Data Mining and Bioinformatic Analysis

The genomic structures of the miR-125a/miR-125b-1 and miR-125b-2 clusters were obtained from the Ensembl genome browser, release 94 [36]. MiRNA sequences were retrieved from the Mirbase [13]. The genomic structure of the murine *Cd34* gene was obtained by “Blasting” the complete murine genome (Annotation release 106) with the cDNA sequences for the two *Cd34* isoforms at the NCBI server. Prediction of *Cd34* miRNA binding sites was made at the “Targetscan” for mouse, release 7.2 [37].

2.8. Statistics

All data are expressed as mean \pm SD. The Kruskal-Wallis test (SPSS 20.0 software, SPSS Inc. Chicago, IL, USA) was used to compare miRNA-expression among basal (8W), control SC-siRNA and anti-CD40 siRNA-treated mice. The Mann-Whitney test was used to compare CD34 expression among siRNA control/treated mice A or in aortas from patients with or without CKD. A value of $p < 0.05$ was considered as statistically significant.

3. Results

We are interested in the mechanisms of ATH progression, and more specifically on those that depend on the function of different lineages of stem cells, such as hemopoietic progenitors, or endothelial progenitor cells, characterized by the expression of the CD34 marker. In this context we have recently described an animal model of ATH progression and reversion in which ApoE-deficient mice were fed with a fat-enriched chow to facilitate ATH development and treated for 14 or 24 weeks with an anti-CD40 siRNA to reverse disease progression, or with an irrelevant, scrambled sequence siRNA as control [25]. This disease model was characterized by the upregulation of miR-125b, a member of the miR-125 family, with ATH progression, which was reversed by the treatment with the anti-CD40 specific siRNA [25]. Interestingly, another highly related member of the family, miR-125a, which has been proposed by bioinformatic analysis to target CD34 did not show up in our previous analysis [25]. Here we address the relationship among the members of the miR-125 family and CD34, and describe a new regulatory mechanism by which a cryptic splice event at the CD34 gene could regulate the accessibility of miR-125a (and other miRNAs) to the 3'UTR or to the coding region of the CD34 transcript.

3.1. MiR-125: A Complex miRNA Family

miR-125b is encoded by two different genes in human and mice (miR-125b-1 and miR-125b-2), co-clustered with genes encoding members of the miR-99 (miR-99a and miR-99b, miR-100) and let-7 (let-7a-2, let-7c-1 and let-7e) families (see Figure 1A). Furthermore, miR-125b is also highly homologous to miR-125a with which it shares the seed sequence (Figure 1B). Bioinformatic analysis predicted specific promoters for human miR-125b-1/2, but expression of miR-125a seemed to be co-regulated with that of the members of its cluster, miR-99b and let-7e [38]. Given the high homology among miR-125a and miR-125b, and the fact that both shared the same seed sequence but had apparently different roles during ATH progression, we re-evaluated the original data from the TLDA experiment to study changes in miR-125a expression during ATH progression. Figure 1C shows the individual, absolute Ct values obtained for each one of the miR-125a, miR-125b and the housekeeping gene U6-snrRNA in the different experimental conditions tested. Furthermore, Table 1 shows the same data expressed as the Δ Ct values after normalization to U6 (miR-125-U6). Both clearly show that expression of miR-125a and miR-125b was very similar in the basal, initial condition at 8 weeks (9.43 ± 0.5 cycles vs. 10.47 ± 0.6 cycles; for miR-125a and miR-125b, respectively; $n = 2$), diverged after 24W of ATH

progression (6.7 ± 0.6 vs. 3.97 ± 0.6 cycles, respectively; $n = 3$) and returned to almost basal levels after treatment with the anti-CD40 siRNA (7.16 ± 1.6 vs. 6.26 ± 2.0 cycles, respectively; $n = 3$). Given the small differences in the expression of miR-125a and miR-125b, we considered that miR-125a could also have a role, though perhaps less relevant than miR-125b, in ATH progression.

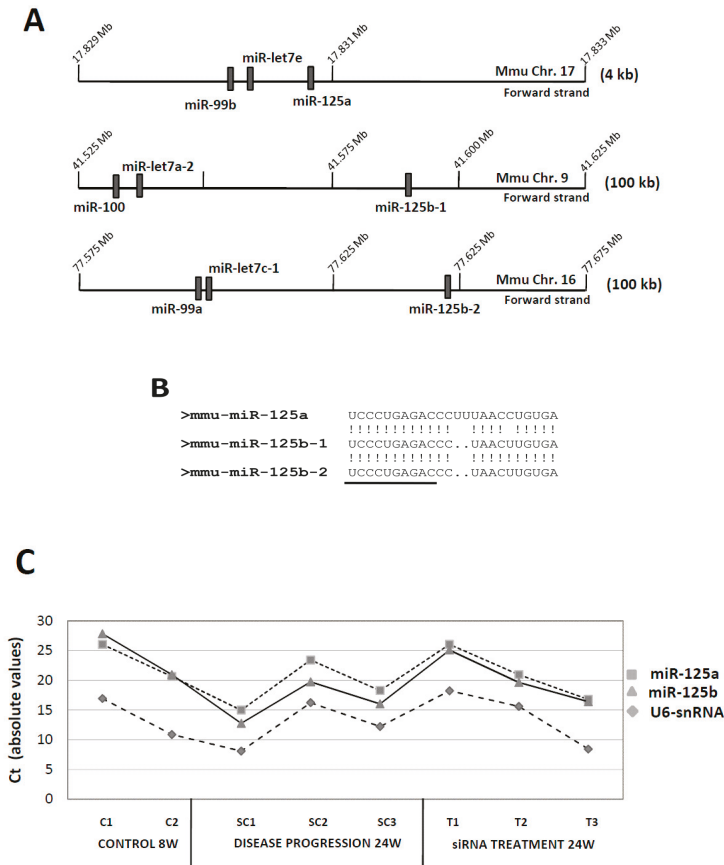


Figure 1. Structure and expression of the murine miR-125a/miR-125b-1/miR-125b-2 gene clusters in ATH progression. (A) Graphic diagram of the three miRNA gene clusters. miRNA gene locations (relative positions) were taken from the Ensembl genome browser [36]. Ensemble code for miR-125a was ENSMUSG00000065479, for miR-125b-1 was ENSMUSG00000093354, and for miR-125b-2 was ENSMUSG00000065472. Only miRNA genes and no other coding/non-coding genes are shown. Graphics (genomic regions and gene boxes) are not drawn to scale, and scales are different for each cluster (see numbering for absolute positions in the chromosome). Numbers to the right show the amount of genomic sequence covered by each graph. (B) Homology among miR-125a/miR-125b-1/miR-125b-2. Shown are the sequences of the mature miRNAs. (!) stands for a homologous nucleotide, while (.) stands for a deleted nucleotide. Underlined, the seed sequence. (C) Absolute expression (without normalization) of murine miR-125a, miR-125b and U6-snRNA (as normalizer) in control, basal samples (8W) in mice treated with the control SC-siRNA (24W, ATH progression) and in mice treated with the specific anti-CD40 siRNA (siRNA treatment, 24W). Shown are the Cts for each individual sample. Expression data was extracted from the original Taqman Low Density Array card (TLDA) experiment [25].

Table 1. Expression of the individual members of the murine miR-125a/miR-125b-1/miR-125b-2 gene clusters in ATH progression and after treatment with the specific anti-CD40 siRNA. The different miRNAs were tested by Taqman Low Density Array card (TLDA) in the different experimental conditions stated [25], and individual expression values were normalized to U6 expression (Δ Ct Gene-U6 snRNA) and expressed as mean \pm SD of the different samples. C8W, control mice at 8 weeks. SC24W, mice treated with the scrambled control siRNA for 24 weeks (control group for ATH progression, see Materials and Methods). T24W, mice treated with the anti-CD40 siRNA for 24 weeks (treatment group). (*) Only one sample could be analyzed.

	Cluster 99b/let7e/125a			miRNA Expression Cluster 100/let7a-2/125b-1			Cluster 99a/let7c-1/125b-2		
	99b	let7e	125a	100	let7a-2	125b-1	99a	let7c-1	125b-2
C8W (n = 2)	10.95 \pm 0.2	8.22 \pm 0.2	9.43 \pm 0.5	10.64 \pm 0.5	13.47 *	10.47 \pm 0.6	10.6 \pm 0.5	10.8 \pm 0.6	10.47 \pm 0.6
SC24W (n = 3)	4.51 \pm 1.3	3.35 \pm 1.8	6.7 \pm 0.6	5.57 \pm 0.8	7.20 \pm 1.8	3.97 \pm 0.6	6.32 \pm 0.3	3.03 \pm 2.2	3.97 \pm 0.6
T24W (n = 3)	6.27 \pm 2.8	4.99 \pm 2.8	7.16 \pm 1.6	6.15 \pm 1.8	7.5 \pm 2.8	6.26 \pm 2.0	6.98 \pm 2.1	5.74 \pm 4.1	6.26 \pm 2.0

3.2. CD34-Positive Cells Are Clustered in Human Aortic Lesions

We had previous evidences of the involvement of CD34 cells in the process of ATH progression (M. Hueso, unpublished results), so that we performed a more complete study of CD34-positive cells during ATH progression. We first analyzed human abdominal aortas from deceased patients by IHQ to detect CD34-positive cells. Figure 2 shows representative views of the human neointimae (Figure 2A,B) and adventitia (Figure 2D) from aortic sections with ATH plaques, as well as a representative section from a region with no plaques (Figure 2C), all of them stained with the anti-CD34 antibody. It can be clearly seen the presence of clusters of CD34-positive cells in the lesions (Figure 2A,B,D), which at a larger magnification showed a fibroblastoid morphology (arrows at Figure 2B). Interestingly, the non-lesion tissue showed a different distribution of the CD34-positive cells which were not detected in the intima, but scattered in muscle layer (Figure 2C for a representative field). We could not analyze the perivascular adipose tissue (PVAT) because it was lost during preparation of the human samples.

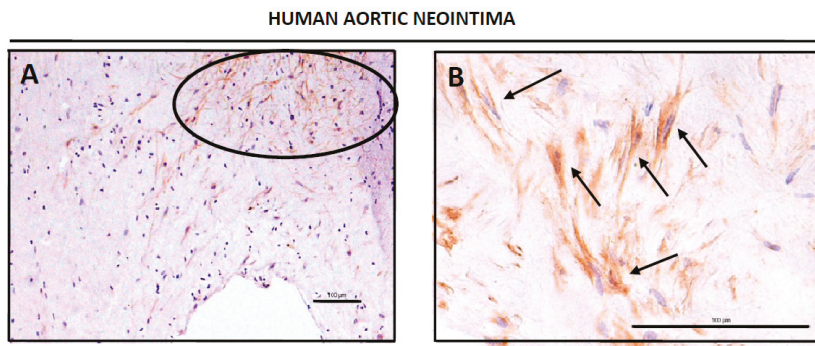


Figure 2. Cont.

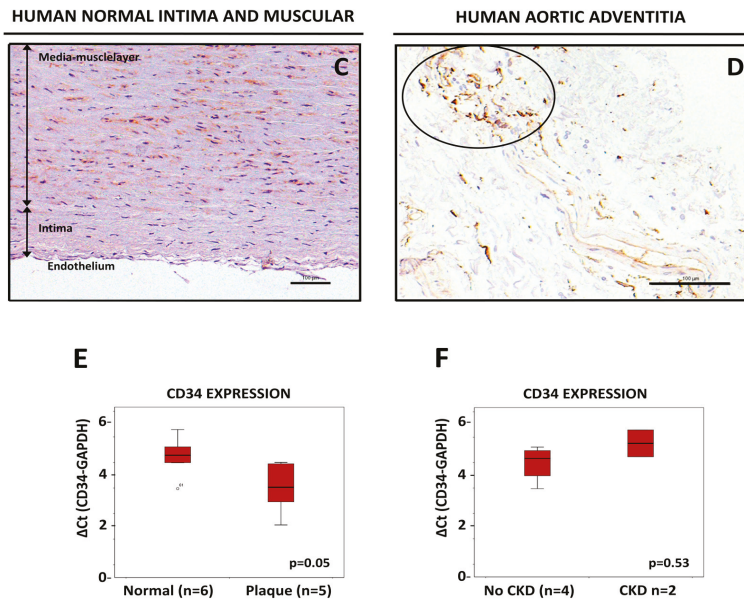


Figure 2. Detection of CD34-positive cells in human aortic tissue. Whole human arteries were isolated, sliced, prepared for IHC analysis and stained for CD34 as described in Material and Methods. (A,B) Shown are representative images of two different neointimas, stained for anti-CD34, at different magnifications (20 \times in 3A, 40 \times in 3B). The circle in (A) shows a cluster of CD34-positive cells, and arrows in (B) show \pm individual fibroblastoid cells positive for CD34 staining. Tissues were counterstained with hematoxylin-eosin. Bars are 100 μ m. (C) Representative image of a human aortic tissue from a normal section (non-lesion) showing the intima and the muscular layer (vertical arrows) and stained for anti-CD34. Tissue was counterstained with hematoxylin-eosin. Bar is 100 μ m. (D) Representative image of a human adventitia stained for anti-CD34. The circle in (D) shows a cluster of CD34-positive cells. Tissue was counterstained with hematoxylin-eosin. Bar is 100 μ m. (E) Expression of *CD34* mRNA measured by qPCR in human ATH plaques compared with normal abdominal aortas. (F) Expression of *CD34* mRNA in aortic tissue with normal vascular walls from CKD patients when compared with non-CKD patients.

We next dissected human ATH samples into plaque and normal tissue, which were tested by qPCR for mRNA levels of *CD34*. Figure 2E show that *CD34* mRNA levels in the plaque were significantly higher than in normal aortic tissue (Δ Ct = 3.5 ± 1.03 cycles in plaque vs. 4.7 ± 0.76 cycles in normal tissue; $p = 0.05$). Since ATH is a frequent complication of Chronic Kidney Disease, we also tested *CD34* mRNA expression in human normal aortas of CKD and non-CKD patients, but only a trend to a lower expression of *CD34* in CKD samples was found (Δ Ct = 4.47 ± 0.71 cycles in non-CKD vs. 5.26 ± 0.75 cycles in CKD; $p = 0.53$, see Figure 2F).

We took advantage of the mouse model of ApoE-deficient mice treated with an anti-CD40 siRNA to study the presence of CD34-positive cells in aortic tissues of animals subjected to different treatments. As in humans, we also found clusters of CD34-positive cells in the intima and adventitia of the murine samples (M. Hueso, unpublished result). Furthermore, we had the opportunity of studying the PVAT in these murine samples. As seen in Figure 3A,B individual CD34 cells were found in the perivascular aortic fat (PVAT) of mice treated with the control, scrambled siRNA (Figure 3A) or with the anti-CD40 siRNA (Figure 3B). When different fields were analyzed and counted, it arose that the number of CD34-positive cells was lower in the PVAT of mice treated with the CD40-siRNA when compared with SC-siRNA treated animals (from $14 \pm 9\%$ CD34+ cells per field in SC-animals vs. $12 \pm 9\%$ CD34+ cells per field in siCD40-treated animals, see Figure 3C), as well as in the adventitia ($18 \pm 6\%$ CD34+ cells

per field vs. $9 \pm 3\%$ respectively; $p < 0.008$, see Figure 3C) or in the intima ($11 \pm 1\%$ CD34+ cells per field vs. $3 \pm 6\%$ respectively; $p < 0.002$, see Figure 3C) suggesting that ATH lesions were associated with an increase in the number of tissue resident CD34 cells.

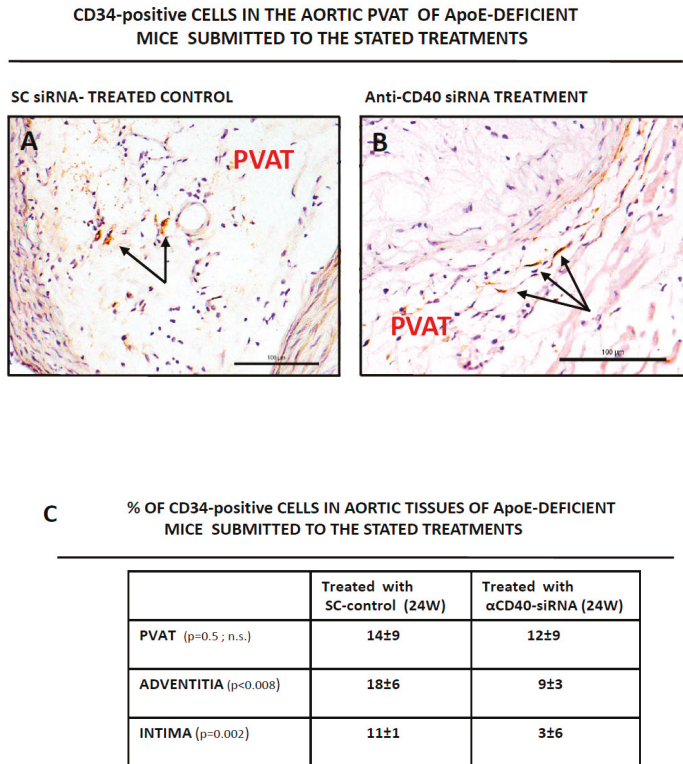


Figure 3. Detection of CD34-positive cells in the murine aortic perivascular tissue. Aortas from ApoE-deficient mice treated with a scrambled (SC) siRNA for 24 weeks as control, or with an anti-CD40 specific siRNA also for 24 weeks, were isolated, sliced, prepared for IHQ analysis and stained for CD34 as described in Material and Methods. (A,B) Show representative fields with individual CD34-positive cells (arrows) in the SC-siRNA control (A) or in the anti-CD40 specific siRNA (B). Tissues were counterstained with hematoxylin-eosin. Bars are 100 μ m. (C) Quantification of CD34-positive cells in the PVAT, intima and adventitia of aortas of ApoE-deficient mice submitted to the stated treatments. A total of 10 fields were examined by two different researchers and the result expressed as % of CD34-positive cells with regard of the total number of eosin-positive nuclei in the field.

3.3. An Exonic Switch Regulates Differential Accession of miR-125, and Other miRNAs, to the CDS or to the 3'UTR of the *Cd34* Transcript

We examined the relationship of *Cd34* with miRNAs, among them the miR-125a, during ATH progression. The murine *Cd34* gene is composed by 8 exons and is submitted to a complex splicing regulation. Exon 8 includes an internal/cryptic splice site (I/CSS, Figure 4A) and two in-frame stop-codons (TGA1 and TGA2). Activation of the usual-I/CSS splice sites lead to the expression of two different *Cd34* mRNAs. The transcript variant 1, TV1, links exon 7 to the entire exon 8 and incorporates a premature stop-codon (TGA1 at Figure 4B), which originates a short, truncated form of CD34 lacking almost the entire intracellular domain (Figure 4C). The second *Cd34* transcript, TV2, links exon 7 to the I/CSS (Figure 4A) and incorporates a partial exon 8 which lacks the premature stop-codon and terminates at the TGA2 (Figure 4B) to originate a “full-length” CD34 protein (Figure 4C). In this way,

the long transcript (including the entire exon 8) originates a short protein by activating a premature STOP codon, while the short transcript (incorporating a partial exon 8) originates the long CD34 protein (Figure 4B).

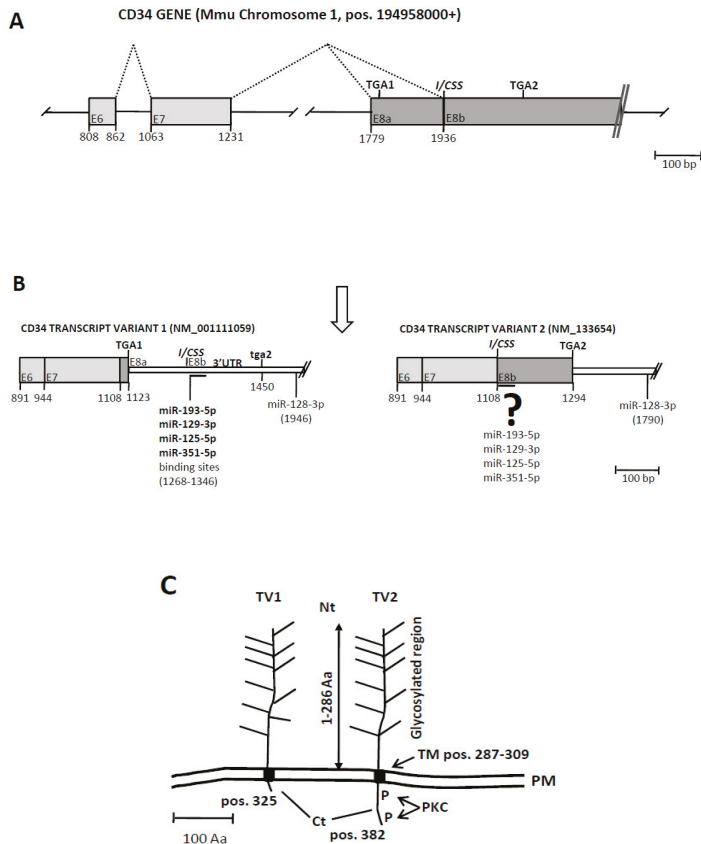
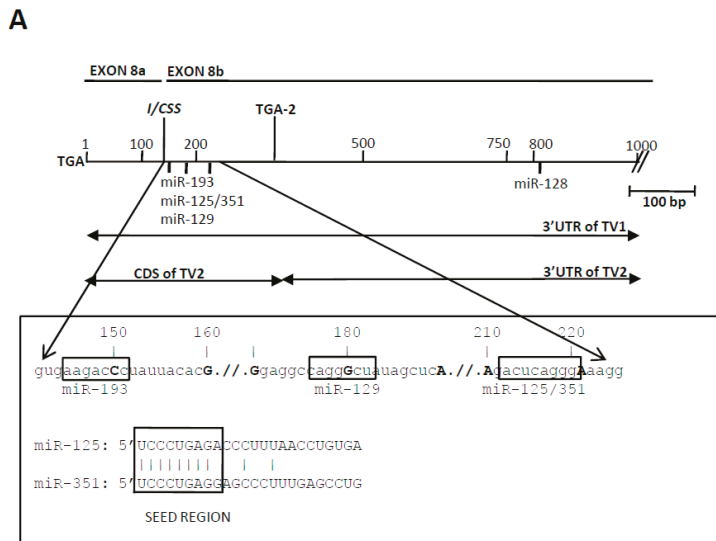


Figure 4. Structure and expression of the murine *Cd34* genetic loci and its interaction with miR-125 and other miRNAs. **(A)** Diagram of the murine *Cd34* gene. Exons are represented by colored boxes, intergenic regions by solid lines and splicing events by dotted lines. Shown are exons 6, 7 and 8 only. Former exon X is here shown as part “a” of exon 8 (see text). Shown are also the two stop codons (TGA1, TGA2) at exon 8, as well as the internal/cryptic splice site (I/CSS) at the virtual juncture of exon 8a/b. Numbering refers to the position in the NCBI Reference Sequence: NC_000067.6. Drawn to scale (1cm = 100bp). **(B)** Diagram showing the two *Cd34* transcripts originated by the internal/cryptic splice event at exon 8. Coding regions are expressed as colored boxes while the 3’UTRs are shown as white bars. Also shown are the two stop codons (the active one in capital letters), the internal/cryptic splice site and the miRNA binding sites predicted by the Targetscanmouse at the 3’UTR of *Cd34* transcript variant 1. Binding sites at the equivalent positions at the coding region of transcript variant 2 are shown under the question mark. Numbering refers to the Genbank accession numbers of the transcripts. Drawn to scale (1cm = 100bp). Ensembl code for *Cd34*-TV1: ENSMUST00000110815. Ensembl code for *Cd34*-TV2: ENSMUST0000016638. **(C)** Diagram showing the two different proteins encoded by CD34 transcripts variant 1 and 2 (TV1, TV2). Shown are the glycosylated Nterminal extracellular domain (Nt), the transmembrane domain (TM) and the different Cterminal domains (Ct) of both protein isoforms. Protein kinase C phosphorylation sites are labelled in TV2. Numbering refers to the Genbank accession numbers of the transcripts (see B). Drawn to scale (1cm = 100Aa). PM stands for plasma membrane.

Bioinformatic analysis with mouse “TargetsCan” of the *Cd34* transcript allowed the detection of binding sites for miR-193, miR-129, miR-125, and miR-351 clustered into 70 bp of sequence immediately downstream of the I/CSS. Interestingly, these binding sites would be located at the 3’UTR of *Cd34*-TV1 but at the coding region of *Cd34*-TV2 (Figures 4B and 5A), depending on the splice event that took place, so that a differential splicing would direct the binding of miR-125 (and the other miRNAs) to alternative locations in the *Cd34* transcript, the 3’UTR of *Cd34*-TV1 or the coding region of *Cd34*-TV2 (Figure 5A). Interestingly, two of the miRNAs that targeted the exon 8b of *Cd34*, miR-125 and miR-351 had an almost identical seed region (Figure 5A). Furthermore, another binding site for miR-128-3p was present at the far 3’UTR of both *Cd34* isoforms, but its location was not affected by the splicing event (see Figure 4B).



B

Expression of miRNAs targeting CD34-TV1 at different conditions					
	miR-193	miR-129	miR-125	miR-351	miR-128
C8W (n=3)	4.59± 2.10	17.88±0.48	9.43±0.5	10.07*	17.14*
SC24W (n=3)	2.77±1.83	9.7±0.48	6.07±0.6	14.28±2.02	12.04±1.66
siCD40-24W (n=3)	8.74±4.67	9.65±2.27	7.16±1.6	16.02±3.76	13.63±3.36

Figure 5. Detailed structure of the *Cd34* region harboring the predicted miRNA binding sites. (A) Diagram of 1 kb of sequence at the 3’end of *Cd34*. Shown is the position of the two stop codons (TGA1 and TGA2), the internal/cryptic splice site (I/CSS) and the position of the miRNA binding sites. Expanded is the sequence of the 70 bp, which concentrate the miRNA binding sites. Boxed, the seed sequences of the miRNAs and the comparison among those of miR-125 and miR-351. (B) Expression of the miRNAs targeting *Cd34* at different experimental conditions. Data shown after normalization to U6-snrRNA. Expression data extracted from the original Taqman Low Density Array card (TLDA) experiment [25]. (*) Only one sample could be analyzed.

Analysis of the expression of the *Cd34*-targeting miRNAs (Figure 5B) showed that miR-193, miR-129, miR-125 and miR-128 shared the same expression pattern characterized by an overexpression in SC24W (i.e., disease progression) that was reversed by the anti-CD40 treatment. On the other hand, miR-351 was the only one whose expression decreased in SC24W, with its values not recovering after si-CD40 treatment.

Lastly, we measured expression *Cd34* expression by qPCR in the SC24W and siCD40-24W samples from our disease model, but unfortunately we did not have (basal) control due to sample exhaustion. We could test expression of *Cd34*–TV1 but not of TV2 since there was not a specific commercial Taqman probe for this isoform, so instead we used a probe that amplified both isoforms. The results confirmed our previous results for CD34 mRNA and protein (as number of cells expressing CD34): in the sense that *Cd34*–TV1 and *Cd34*–(TV1 + TV2) were more expressed in disease progression SC24W (TV1: 6.24 ± 0.62 cycles; (TV1 + TV2): 4.73 ± 0.75 cycles) than in the treatment group siCD4/24W (TV1: 7.88 ± 0.37 cycles; (TV1 + TV2): 6.42 ± 0.50 cycles). Furthermore, the difference among *Cd34*–TV1 and *Cd34*–(TV1 + TV2) indicated that both, TV1 and TV2 were expressed.

4. Discussion

ATH progression has an hematopoietic component that includes expansion of the hematopoietic stem-cell compartment [39], and we are interested in the role of the stem-cell genes *Cd34* and *Lin28a/b* in the development of atherosclerotic lesions. Originally characterized as an hemopoietic progenitor cell marker [40], CD34 was subsequently detected in vascular endothelial cells [41] and neovascularized tissues [42]. CD34 has been shown to be important in the development of ATH lesions, mainly in neo/re-vascularization events, with CD34-positive cells being detected more frequently in inflammatory-erosive plaques, when compared with plaques of the lipid or degenerative-necrotic plaques [43]. It is thus evident that elucidating the mechanisms of *CD34* expression and CD34-cell differentiation would have a positive impact on CVD disease.

We have previously demonstrated the involvement of miR-125b in the progression of ATH [25], but our data was inconclusive about the role of the other highly related member of the family, miR-125a (Figure 1A,B). It is interesting to notice that *Cd34*, *Lin28a/b* and the miR-125 family shared a molecular link, since data mining of miRNA targets showed that *Cd34* and *Lin28a/b* were predicted targets of miR-125a/b. Furthermore, miR-125a was recently described as a regulator of hemopoietic progenitor cell stemness [44]. In this context, we have addressed the study of the relationship among *Cd34* and miR-125(a,b) with the aim to generate new tools for the study of ATH progression.

We firstly substantiated the involvement of miR-125a in ATH progression by reanalyzing our previous miRNA data to compare expression of the highly related miR-125a and miR-125b. As seen in Figure 1C, while expression of both (in absolute Ct values) was very similar in the control samples, miR-125a was slightly less expressed than miR-125b during ATH progression (SC1-SC3), thus justifying why miR-125a did not show up in our earlier analysis of ATH-progression related miRNA genes. Nevertheless, after treatment with the specific anti-CD40 siRNA both miRNAs showed similar expression levels. The difference in expression levels was confirmed after normalization with the U6-snrRNA control (Table 1), and could be likely due to the fact that miR-125b was encoded by two different genes (b-1 and b-2) while miR-125a was encoded by a single gene (Figure 1A).

We next studied CD34 expression in human ATH plaques, as well as in aortic tissues from the mice. CD34-positive cells could be detected as clusters of cells in the neointimas and adventitias of human (Figure 2) and mouse (Figure 3) diseased aortas. Furthermore, individual CD34-positive cells could also be detected in the murine perivascular adipose tissue (PVAT) (Figure 3). Interestingly, in aortic tissue from non diseased aortas the distribution of CD34-positive cells was totally different, being mostly detected as scattered cells in the media layer (Figure 2C). One possibility is that this differential distribution reflected distinct migratory or homing abilities of CD34-positive cells in the context of normal or ATH-diseased aortic tissue.

We next quantified CD34 expression/positive cells by two methods in our two disease models, by PCR (human, Figure 2E) and by directly counting CD34-positive cells (mouse, Figure 3C). Both approaches showed an increase in CD34 in ATH lesions. Increased levels of CD34 mRNA were detected in human ATH plaques when compared with normal aortic tissue (Figure 2E) and more cells were detected in SC24W mice (disease progression) than in siCD40 mice (Figure 3). This is very interesting because most works on CD34-positive PCs and ATH progression has been made with blood-circulating cells and there are scant data on lesion resident cells. Furthermore, a number of authors showed a direct correlation among newly formed blood vessels (as CD34-positive cells) at the neointima and ATH progression [45], that statin treated patients had reduced intraplaque angiogenesis [46], or that a catechin supplementation in the diet of apoE deficient mice resulted in a reduction of ATH lesions and a significant reduction in CD34 expression [47], in line with the results here shown. On the other hand ATH progression has been related with reduced levels of circulating EPCs [48], while a diminution in circulating CD34+/CD45(dim)/VEGFR2- or CD34+/CD45(dim)/CD133+/VEGFR2-cells was shown to highlight patients with coronary endothelial dysfunction [49].

Thus, it is clear that ATH progresses with a simultaneous increase in *CD34* expression and in the number of CD34-positive resident cells, a fact that would be apparently in full contradiction with a role for the miR-125 family in the regulation of *CD34* mRNA stability, since these are also overexpressed in ATH progression. Nevertheless, the end of the coding domain and 3'UTR region of the *CD34* gene is a very complex genomic region, and upon a detailed analysis two features arose that could help to understand the contradictory scenario. Firstly, this region is very poor in predicted miRNA binding sites (after Targetscan), with only five of them, four clustered in a short, 70 bp region. This must be a distinctive feature of the *CD34* gene since, on the other hand, only one miRNA (miR-665) has been experimentally confirmed as targeting *CD34* [50], although this was not predicted by Targetscan. Furthermore, two of these miRNAs, miR-125 and miR-351, shared the same seed sequence, but only this last showed an expression profile (downregulated in ATH progression) compatible with being a *CD34* effector (Figure 5B), thus casting serious doubts on the role of the miR-125 family in the regulation of *CD34* stability. It is thus possible that only miR-351, but not miR-125, be involved in *CD34* regulation, and that data-mining and bioinformatic analysis were misled by the similarities among their seed sequences. In this sense, miR-351 has been characterized as a developmental regulator that represses differentiation of mesenchymal stem cells (to osteoblasts) by targeting the vitamin D receptor [51], or suppresses angiogenesis [52], roles compatible with its downregulation during ATH progression in our animal model.

The second feature to consider when dealing with miRNAs targeting *Cd34* is the structure of its 3' end region. Originally, the *Cd34* gene was shown to be structured in 8 exons [53], and a subsequent work showed that an alternative splicing event at exons X and 8 generated two transcript variants which differed in the Cterminal domain of the protein, generating an isoform almost lacking this domain and another one with a full-length Cterminal domain [10]. This last one included two PKC phosphorylation cytoplasmic sites [54], suggesting that the truncated form was unable for signaling. Nevertheless, we have carefully reanalyzed the genomic sequence data of *Cd34* and concluded that the splicing event at exon 8 was more compatible with a splicing at an internal/cryptic splice site than with an alternative splicing as previously described, since the originally defined alternative exon X is in fact part of exon 8 and there is no intronic sequence among them.

This cryptic splicing (exon switch) is intriguing because it would have a compelling consequence on the miRNA binding sites of *Cd34*, since these would be located at the beginning of the 3'UTR sequence at the *Cd34*-TV1, but at the end of the coding region in *Cd34*-TV2 (Figures 4 and 5). This is very interesting because recent work suggested that miRNAs targeting coding regions would cause translational inhibition rather than degradation of the transcript (see [55] and references therein). In this way *Cd34*-TV2 could be also considered as a "miRNA sponge" that would decrease microRNA availability and relieve the repression of target RNAs through sequestering miRNAs away from

parental mRNA. Should this be the case, this could have an impact on the balance of CD34 isoform expression, as well as in the function of miR-125s since these could bind to sites in different transcript contexts (3'UTR and CDS) and likely compete with miR-351 for the binding site. This complex miRNA binding region at the 3' end of the *Cd34* transcript is worth studying “in vitro” with advanced tools of molecular biology to describe its real impact on the dynamics of the *Cd34* transcript and protein as well as on the cells expressing this membrane marker and their role in human disease.

In conclusion, here we have demonstrated an increase in tissue resident CD34-positive cells in the aortic tissue of human and mice during ATH progression. Secondly, although we have failed to provide a sound functional relationship among miR-125 and *Cd34*, we have detected miR-351 as a possible, alternative *Cd34* effector. Lastly, we describe a novel regulatory mechanism of miRNA function by which a cryptic splice event at the *CD34* gene could regulate the accessibility of miR-125a to the 3'UTR or to the coding region of the *CD34* transcript, likely producing different effects on its translation or its stability. Work is in progress to confirm this hypothesis.

Author Contributions: Formal analysis, M.H. and E.N.; Funding acquisition, M.H., J.M.C. and J.T.; Investigation, M.H., J.M.C., J.T. and E.N.; Methodology, M.H. and E.N.; Project administration, J.M.C. and J.T.; Resources, J.M.C. and J.T.; Supervision, M.H., J.M.C., J.T. and E.N.; Writing-original draft, M.H. and E.N.; Writing-review & editing, M.H., J.M.C., J.T. and E.N.

Funding: This research was funded by Instituto de Salud Carlos III Grant PI18/0110 (Co-funded by European Regional Development Fund. ERDF).

Acknowledgments: We thank the CERCA program/Generalitat de Catalunya, Spain, for institutional support.

Conflicts of Interest: The authors state no conflict of interest.

References

- Hueso, M.; Cruzado, J.M.; Torras, J.; Navarro, E. ALUminating the Path of Atherosclerosis Progression: Chaos Theory Suggests a Role for Alu Repeats in the Development of Atherosclerotic Vascular Disease. *Int. J. Mol. Sci.* **2018**, *19*, 1734. [[CrossRef](#)] [[PubMed](#)]
- Andreou, I.; Tousoulis, D.; Tentolouris, C.; Antoniadis, C.; Stefanadis, C. Potential role of endothelial progenitor cells in the pathophysiology of heart failure: Clinical implications and perspectives. *Atherosclerosis* **2006**, *189*, 247–254. [[CrossRef](#)]
- Urbich, C.; Dimmeler, S. Endothelial progenitor cells: Characterization and role in vascular biology. *Circ. Res.* **2004**, *95*, 343–353. [[CrossRef](#)]
- Du, F.; Zhou, J.; Gong, R.; Huang, X.; Pansuria, M.; Virtue, A.; Li, X.; Wang, H.; Yang, X.F. Endothelial progenitor cells in atherosclerosis. *Front. Biosci.* **2012**, *17*, 2327–2349. [[CrossRef](#)]
- Rohde, E.; Malischnik, C.; Thaler, D.; Maierhofer, T.; Linkesch, W.; Lanzer, G.; Guelly, C.; Strunk, D. Blood monocytes mimic endothelial progenitor cells. *Stem Cells* **2006**, *24*, 357–367. [[CrossRef](#)] [[PubMed](#)]
- Hill, J.M.; Zalos, G.; Halcox, J.P.; Schenke, W.H.; Waclawiw, M.A.; Quyyumi, A.A.; Finkel, T. Circulating endothelial progenitor cells, vascular function, and cardiovascular risk. *N. Engl. J. Med.* **2003**, *348*, 593–600. [[CrossRef](#)] [[PubMed](#)]
- Verma, S.; Kuliszewski, M.A.; Li, S.H.; Szmítko, P.E.; Zucco, L.; Wang, C.H.; Badiwala, M.V.; Mickle, D.A.; Weisel, R.D.; Fedak, P.W.; et al. C-reactive protein attenuates endothelial progenitor cell survival, differentiation, and function: Further evidence of a mechanistic link between C-reactive protein and cardiovascular disease. *Circulation* **2004**, *109*, 2058–2067. [[CrossRef](#)] [[PubMed](#)]
- Riesinger, L.; Saemisch, M.; Nickmann, M.; Methe, H. CD34(+) circulating cells display signs of immune activation in patients with acute coronary syndrome. *Heart Vessel.* **2018**. [[CrossRef](#)]
- Leung, K.T.; Chan, K.Y.; Ng, P.C.; Lau, T.K.; Chiu, W.M.; Tsang, K.S.; Li, C.K.; Kong, C.K.; Li, K. The tetraspanin CD9 regulates migration, adhesion, and homing of human cord blood CD34+ hematopoietic stem and progenitor cells. *Blood* **2011**, *117*, 1840–1850. [[CrossRef](#)]
- Suda, J.; Sudo, T.; Ito, M.; Ohno, N.; Yamaguchi, Y.; Suda, T. Two types of murine CD34 mRNA generated by alternative splicing. *Blood* **1992**, *79*, 2288–2295. [[PubMed](#)]
- Lanza, F.; Healy, L.; Sutherland, D.R. Structural and functional features of the CD34 antigen: An update. *J. Biol. Regul. Homeost. Agents* **2001**, *15*, 1–13.

12. Lekka, E.; Hall, J. Noncoding RNAs in disease. *FEBS Lett.* **2018**, *592*, 2884–2900. [[CrossRef](#)]
13. Kozomara, A.; Griffiths-Jones, S. miRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **2014**, *42*, D68–D73. [[CrossRef](#)] [[PubMed](#)]
14. Jackson, A.O.; Regine, M.A.; Subrata, C.; Long, S. Molecular mechanisms and genetic regulation in atherosclerosis. *Int. J. Cardiol. Heart Vasc.* **2018**, *21*, 36–44. [[CrossRef](#)] [[PubMed](#)]
15. Chung, A.C.; Lan, H.Y. MicroRNAs in renal fibrosis. *Front. Physiol.* **2015**, *6*, 50. [[CrossRef](#)] [[PubMed](#)]
16. Yang, S.; Ye, Z.M.; Chen, S.; Luo, X.Y.; Chen, S.L.; Mao, L.; Li, Y.; Jin, H.; Yu, C.; Xiang, F.X.; et al. MicroRNA-23a-5p promotes atherosclerotic plaque progression and vulnerability by repressing ATP-binding cassette transporter A1/G1 in macrophages. *J. Mol. Cell. Cardiol.* **2018**, *123*, 139–149. [[CrossRef](#)] [[PubMed](#)]
17. Rink, C.; Khanna, S. MicroRNA in ischemic stroke etiology and pathology. *Physiol. Genom.* **2011**, *43*, 521–528. [[CrossRef](#)] [[PubMed](#)]
18. Li, S.; Geng, Q.; Chen, H.; Zhang, J.; Cao, C.; Zhang, F.; Song, J.; Liu, C.; Liang, W. The potential inhibitory effects of miR19b on vulnerable plaque formation via the suppression of STAT3 transcriptional activity. *Int. J. Mol. Med.* **2018**, *41*, 859–867. [[CrossRef](#)]
19. Price, N.L.; Rotllan, N.; Canfran-Duque, A.; Zhang, X.; Pati, P.; Arias, N.; Moen, J.; Mayr, M.; Ford, D.A.; Baldan, A.; et al. Genetic Dissection of the Impact of miR-33a and miR-33b during the Progression of Atherosclerosis. *Cell Rep.* **2017**, *21*, 1317–1330. [[CrossRef](#)]
20. Chen, B.; Luo, L.; Wei, X.; Gong, D.; Jin, L. Altered Plasma miR-144 as a Novel Biomarker for Coronary Artery Disease. *Ann. Clin. Lab. Sci.* **2018**, *48*, 440–445.
21. Reddy, L.L.; Shah, S.A.; Ponde, C.K.; Rajani, R.M.; Ashavaid, T.F. Circulating miRNA-33: A potential biomarker in patients with Coronary Artery Disease (CAD). *Biomarkers* **2018**. [[CrossRef](#)] [[PubMed](#)]
22. Zhang, F.; Zhao, J.; Sun, D.; Wei, N. MiR-155 inhibits transformation of macrophages into foam cells via regulating CEH expression. *Biomed. Pharmacother.* **2018**, *104*, 645–651. [[CrossRef](#)] [[PubMed](#)]
23. Di Gregoli, K.; Mohamad Anuar, N.N.; Bianco, R.; White, S.J.; Newby, A.C.; George, S.J.; Johnson, J.L. MicroRNA-181b Controls Atherosclerosis and Aneurysms Through Regulation of TIMP-3 and Elastin. *Circ. Res.* **2017**, *120*, 49–65. [[CrossRef](#)] [[PubMed](#)]
24. Li, H.; Ouyang, X.P.; Jiang, T.; Zheng, X.L.; He, P.P.; Zhao, G.J. MicroRNA-296: A promising target in the pathogenesis of atherosclerosis? *Mol. Med.* **2018**, *24*, 12. [[CrossRef](#)] [[PubMed](#)]
25. Hueso, M.; De Ramon, L.; Navarro, E.; Ripoll, E.; Cruzado, J.M.; Grinyo, J.M.; Torras, J. Silencing of CD40 in vivo reduces progression of experimental atherogenesis through an NF-kappaB/miR-125b axis and reveals new potential mediators in the pathogenesis of atherosclerosis. *Atherosclerosis* **2016**, *255*, 80–89. [[CrossRef](#)] [[PubMed](#)]
26. Sun, Y.M.; Lin, K.Y.; Chen, Y.Q. Diverse functions of miR-125 family in different cell contexts. *J. Hematol. Oncol.* **2013**, *6*, 6. [[CrossRef](#)] [[PubMed](#)]
27. Graff, J.W.; Dickson, A.M.; Clay, G.; McCaffrey, A.P.; Wilson, M.E. Identifying functional microRNAs in macrophages with polarized phenotypes. *J. Biol. Chem.* **2012**, *287*, 21816–21825. [[CrossRef](#)]
28. Kim, S.W.; Ramasamy, K.; Bouamar, H.; Lin, A.P.; Jiang, D.; Aguiar, R.C. MicroRNAs miR-125a and miR-125b constitutively activate the NF-kappaB pathway by targeting the tumor necrosis factor alpha-induced protein 3 (TNFAIP3, A20). *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 7865–7870. [[CrossRef](#)]
29. Chen, T.; Huang, Z.; Wang, L.; Wang, Y.; Wu, F.; Meng, S.; Wang, C. MicroRNA-125a-5p partly regulates the inflammatory response, lipid uptake, and ORP9 expression in oxLDL-stimulated monocyte/macrophages. *Cardiovasc. Res.* **2009**, *83*, 131–139. [[CrossRef](#)]
30. Xiu, L.; King, Q.; Mao, J.; Sun, H.; Teng, W.; Shan, Z. miRNA-125b-5p Suppresses Hypothyroidism Development by Targeting Signal Transducer and Activator of Transcription 3. *Med. Sci. Monit.* **2018**, *24*, 5041–5049. [[CrossRef](#)]
31. Wang, S.; Wu, J.; Ren, J.; Vlantis, A.C.; Li, M.Y.; Liu, S.Y.W.; Ng, E.K.W.; Chan, A.B.W.; Luo, D.C.; Liu, Z.; et al. MicroRNA-125b Interacts with Foxp3 to Induce Autophagy in Thyroid Cancer. *Mol. Ther.* **2018**, *26*, 2295–2303. [[CrossRef](#)] [[PubMed](#)]
32. Wu, Q.B.; Chen, J.; Zhu, J.W.; Yin, X.; You, H.Y.; Lin, Y.R.; Zhu, H.Q. MicroRNA125 inhibits RKO colorectal cancer cell growth by targeting VEGF. *Int. J. Mol. Med.* **2018**, *42*, 665–673. [[CrossRef](#)] [[PubMed](#)]
33. Natalia, M.A.; Alejandro, G.T.; Virginia, T.J.; Alvarez-Salas, L.M. MARK1 is a Novel Target for miR-125a-5p: Implications for Cell Migration in Cervical Tumor Cells. *MicroRNA* **2018**, *7*, 54–61. [[CrossRef](#)] [[PubMed](#)]

34. Tong, Z.; Liu, N.; Lin, L.; Guo, X.; Yang, D.; Zhang, Q. miR-125a-5p inhibits cell proliferation and induces apoptosis in colon cancer via targeting BCL2, BCL2L12 and MCL1. *Biomed. Pharmacother.* **2015**, *75*, 129–136. [[CrossRef](#)] [[PubMed](#)]
35. Hueso, M.; Torras, J.; Carrera, M.; Vidal, A.; Navarro, E.; Grinyo, J. Chronic Kidney Disease is associated with an increase of Intimal Dendritic cells in a comparative autopsy study. *J. Inflamm.* **2015**, *12*, 26. [[CrossRef](#)] [[PubMed](#)]
36. Cunningham, F.; Achuthan, P.; Akanni, W.; Allen, J.; Amode, M.R.; Armean, I.M.; Bennett, R.; Bhai, J.; Billis, K.; Boddu, S.; et al. Ensembl 2019. *Nucleic Acids Res.* **2018**. [[CrossRef](#)] [[PubMed](#)]
37. Agarwal, V.; Bell, G.W.; Nam, J.W.; Bartel, D.P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **2015**, *4*. [[CrossRef](#)] [[PubMed](#)]
38. Shaham, L.; Binder, V.; Gefen, N.; Borkhardt, A.; Izraeli, S. MiR-125 in normal and malignant hematopoiesis. *Leukemia* **2012**, *26*, 2011–2018. [[CrossRef](#)]
39. Van der Valk, F.M.; Kuijk, C.; Verweij, S.L.; Stiekema, L.C.A.; Kaiser, Y.; Zeerleder, S.; Nahrendorf, M.; Voermans, C.; Stroes, E.S.G. Increased haematopoietic activity in patients with atherosclerosis. *Eur. Heart J.* **2017**, *38*, 425–432. [[CrossRef](#)]
40. Sutherland, D.R.; Watt, S.M.; Dowden, G.; Karhi, K.; Baker, M.A.; Greaves, M.F.; Smart, J.E. Structural and partial amino acid sequence analysis of the human hemopoietic progenitor cell antigen CD34. *Leukemia* **1988**, *2*, 793–803.
41. Fina, L.; Molgaard, H.V.; Robertson, D.; Bradley, N.J.; Monaghan, P.; Delia, D.; Sutherland, D.R.; Baker, M.A.; Greaves, M.F. Expression of the CD34 gene in vascular endothelial cells. *Blood* **1990**, *75*, 2417–2426. [[PubMed](#)]
42. Schlingemann, R.O.; Rietveld, F.J.; de Waal, R.M.; Bradley, N.J.; Skene, A.I.; Davies, A.J.; Greaves, M.F.; Denekamp, J.; Ruiter, D.J. Leukocyte antigen CD34 is expressed by a subset of cultured endothelial cells and on endothelial abluminal microprocesses in the tumor stroma. *Lab. Invest.* **1990**, *62*, 690–696. [[PubMed](#)]
43. Murashov, I.S.; Volkov, A.M.; Kazanskaya, G.M.; Kliver, E.E.; Chernyavsky, A.M.; Nikityuk, D.B.; Lushnikova, E.L. Immunohistochemical Features of Different Types of Unstable Atherosclerotic Plaques of Coronary Arteries. *Bull. Exp. Biol. Med.* **2018**, *166*, 102–106. [[CrossRef](#)] [[PubMed](#)]
44. Wojtowicz, E.E.; Lechman, E.R.; Hermans, K.G.; Schoof, E.M.; Wienholds, E.; Isserlin, R.; van Veelen, P.A.; Broekhuis, M.J.; Janssen, G.M.; Trotman-Grant, A.; et al. Ectopic miR-125a Expression Induces Long-Term Repopulating Stem Cell Capacity in Mouse and Human Hematopoietic Progenitors. *Cell Stem Cell* **2016**, *19*, 383–396. [[CrossRef](#)] [[PubMed](#)]
45. Nakano, T.; Nakashima, Y.; Yonemitsu, Y.; Sumiyoshi, S.; Chen, Y.X.; Akishima, Y.; Ishii, T.; Iida, M.; Sueishi, K. Angiogenesis and lymphangiogenesis and expression of lymphangiogenic factors in the atherosclerotic intima of human coronary arteries. *Hum. Pathol.* **2005**, *36*, 330–340. [[CrossRef](#)] [[PubMed](#)]
46. Koutouzis, M.; Nomikos, A.; Nikolidakis, S.; Tzavara, V.; Andrikopoulos, V.; Nikolaou, N.; Barbatis, C.; Kyriakides, Z.S. Statin treated patients have reduced intraplaque angiogenesis in carotid endarterectomy specimens. *Atherosclerosis* **2007**, *192*, 457–463. [[CrossRef](#)] [[PubMed](#)]
47. Auclair, S.; Milenkovic, D.; Besson, C.; Chauvet, S.; Gueux, E.; Morand, C.; Mazur, A.; Scalbert, A. Catechin reduces atherosclerotic lesion development in apo E-deficient mice: A transcriptomic study. *Atherosclerosis* **2009**, *204*, e21–e27. [[CrossRef](#)]
48. Schmidt-Lucke, C.; Rossig, L.; Fichtlscherer, S.; Vasa, M.; Britten, M.; Kamper, U.; Dimmeler, S.; Zeiher, A.M. Reduced number of circulating endothelial progenitor cells predicts future cardiovascular events: Proof of concept for the clinical importance of endogenous vascular repair. *Circulation* **2005**, *111*, 2981–2987. [[CrossRef](#)]
49. Boilson, B.A.; Kiernan, T.J.; Harbuzariu, A.; Nelson, R.E.; Lerman, A.; Simari, R.D. Circulating CD34+ cell subsets in patients with coronary endothelial dysfunction. *Nat. Clin. Pract. Cardiovasc. Med.* **2008**, *5*, 489–496. [[CrossRef](#)]
50. Fan, J.; Li, H.; Nie, X.; Yin, Z.; Zhao, Y.; Zhang, X.; Yuan, S.; Li, Y.; Chen, C.; Wang, D.W. MiR-665 aggravates heart failure via suppressing CD34-mediated coronary microvessel angiogenesis. *Aging* **2018**, *10*, 2459–2479. [[CrossRef](#)]
51. Hou, Q.; Huang, Y.; Luo, Y.; Wang, B.; Liu, Y.; Deng, R.; Zhang, S.; Liu, F.; Chen, D. MiR-351 negatively regulates osteoblast differentiation of MSCs induced by (+)-cholesten-3-one through targeting VDR. *Am. J. Transl. Res.* **2017**, *9*, 4963–4973. [[PubMed](#)]

52. Zhang, Y.; Liu, Y.; Zhang, H.; Wang, M.; Zhang, J. Mmu-miR-351 attenuates the survival of cardiac arterial endothelial cells through targeting STAT3 in the atherosclerotic mice. *Biochem. Biophys. Res. Commun.* **2015**, *468*, 300–305. [[CrossRef](#)] [[PubMed](#)]
53. Brown, J.; Greaves, M.F.; Molgaard, H.V. The gene encoding the stem cell antigen, CD34, is conserved in mouse and expressed in haemopoietic progenitor cell lines, brain, and embryonic fibroblasts. *Int. Immunol.* **1991**, *3*, 175–184. [[CrossRef](#)] [[PubMed](#)]
54. Fackler, M.J.; Civin, C.I.; Sutherland, D.R.; Baker, M.A.; May, W.S. Activated protein kinase C directly phosphorylates the CD34 antigen on hematopoietic cells. *J. Biol. Chem.* **1990**, *265*, 11056–11061. [[PubMed](#)]
55. Brummer, A.; Hausser, J. MicroRNA binding sites in the coding region of mRNAs: Extending the repertoire of post-transcriptional gene regulation. *BioEssays* **2014**, *36*, 617–626. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

Single Nucleotide Polymorphisms in *MIR143* Contribute to Protection against Non-Hodgkin Lymphoma (NHL) in Caucasian Populations

Gabrielle Bradshaw, Larisa M. Haupt, Eunise M. Aquino, Rodney A. Lea, Heidi G. Sutherland and Lyn R. Griffiths *

Genomics Research Centre, School of Biomedical Sciences, Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, QLD 4001, Australia; gabrielle.bradshaw@hdr.qut.edu.au (G.B.); larisa.haupt@qut.edu.au (L.M.H.); eunise.aquino@connect.qut.edu.au (E.M.A.); rodney.a.lea@gmail.com (R.A.L.); heidi.sutherland@qut.edu.au (H.G.S.)

* Correspondence: lyn.griffiths@qut.edu.au; Tel.: +617-3138-6102

Received: 11 February 2019; Accepted: 22 February 2019; Published: 27 February 2019

Abstract: Recent studies show an association of microRNA (miRNA) polymorphisms (miRSNPs) in different cancer types, including non-Hodgkin lymphoma (NHL). The identification of miRSNPs that are associated with NHL susceptibility may provide biomarkers for early diagnosis and prognosis for patients who may not respond well to current treatment options, including the immunochemotherapy drug combination that includes rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone (R-CHOP). We developed a panel of miRSNPs for genotyping while using multiplex PCR and chip-based mass spectrometry analysis in an Australian NHL case-control population (300 cases, 140 controls). Statistical association with NHL susceptibility was performed while using Chi-square (χ^2) and logistic regression analysis. We identified three SNPs in *MIR143* that are to be significantly associated with reduced risk of NHL: rs3733846 (odds ratio (OR) [95% confidence interval (CI)] = 0.54 [0.33–0.86], $p = 0.010$), rs41291957 (OR [95% CI] = 0.61 [0.39–0.94], $p = 0.024$), and rs17723799 (OR [95% CI] = 0.43 [0.26–0.71], $p = 0.0009$). One SNP, rs17723799, remained significant after correction for multiple testing ($p = 0.015$). Subsequently, we investigated an association between the rs17723799 genotype and phenotype by measuring target gene Hexokinase 2 (*HKII*) expression in cancer cell lines and controls. Our study is the first to report a correlation between miRSNPs in *MIR143* and a reduced risk of NHL in Caucasians, and it is supported by significant SNPs in high linkage disequilibrium (LD) in a large European NHL genome wide association study (GWAS) meta-analysis.

Keywords: biomarker; cancer; *MIR143*; miRNP; non-Hodgkin lymphoma; rs17723799; Hexokinase 2; HKII

1. Introduction

The global incidence of non-Hodgkin lymphoma (NHL) varies with the geographical region, with the highest incidence rates occurring in North America, Europe, and Australasia [1]. According to the American Cancer Society, non-Hodgkin lymphoma (NHL) is one of the most common cancers in the United States, with around 75,000 people being predicted to be diagnosed in 2018, resulting in around 20,000 deaths [2]. The European Cancer Information System predicts approximately 115,000 new NHL cases in 2018, resulting in 48,000 deaths [3]. Currently, in Australia, lymphoma is the sixth most commonly diagnosed cancer (most common is breast cancer, followed by prostate, colorectal, melanoma, and lung cancer) [4]. NHL is comprised of a large group of diverse lymphoid malignancies involving either the lymph nodes or extranodal sites and it is more prevalent in males living in

more developed countries over the age of 60 years [5]. The two most common NHL subtypes are diffuse large B-cell lymphoma (DLBCL), which is highly aggressive with a poor prognosis [6] and follicular lymphoma (FL), which is more indolent, but it can undergo a high-grade transformation into DLBCL [7]. Molecular mechanisms that are involved in NHL pathogenicity have been widely investigated; however, there is still a need to identify improved biomarkers for clinical diagnostic and therapeutic use.

Mature microRNAs (miRNAs) are short sequences of non-coding RNA that function to regulate the translational expression of their target genes through complementary binding to the 3'-untranslated regions (3'-UTRs) of the target mRNA, in what is referred to as the miRNA recognition element (MRE) [8,9]. The processing pathway and biogenesis of miRNAs has been well characterised [10–12], with several miRNAs being involved in targeting and regulating the expression of oncogenes and/or tumour suppressor genes [13]. Single nucleotide polymorphisms (SNPs) in miRNA biogenesis genes and miRNA genes themselves influence miRNA biogenesis, the processing of precursors to mature miRNAs, regulatory function, and stability of the miRNAs, and, if located in the mRNA 3'-UTR binding sites, can also disrupt miRNA binding and affect the expression of these target genes [14]. According to the miRNA SNP Disease Database (MSDD) [15] SNPs that occur in miRNA-related functional regions, such as mature miRNAs, promoter regions, pri- and pre-miRNAs, and target gene 3'UTR binding sites are collectively referred to as “miRSNPs” (miRNA polymorphisms). These polymorphisms represent a novel category of functional molecules that regulate gene expression [16]. SNPs in miRNA genes have been shown to affect miRNA biogenesis, with these SNPs potentially resulting in reduced and increased mature miRNA levels. It has been predicted that, if a SNP occurs in the miRNA stem region and decreases the stability of the hairpin structure, then it will reduce the amount of mature miRNA produced [16]; however, if it fails to decrease stability, it will tend to increase the amount of mature miRNA produced and lead to disease pathogenesis [16]. Furthermore, SNPs occurring in the non-coding regulatory regions of genes, such as the miRNA-binding sites in the 3'-UTRs of target genes, can cause mRNA instability changes, leading to disease pathogenesis. A number of studies have identified roles for miRNAs [17], as well as miRSNPs [18,19] in NHL. These studies have examined the suitability of miRSNPs as diagnostic and prognostic biomarkers for improved clinical treatment and survival, particularly as NHL is a heterogeneous disease with outcomes that vary from patient to patient. However, few studies have demonstrated the association of SNPs in miRNAs with NHL risk.

Serum lactate dehydrogenase A (LDH-A), which is an enzyme converting pyruvate to lactate, is commonly elevated in aggressive NHL cases, such as DLBCL and Burkitt's lymphoma, and it has been used as a prognostic indicator for overall survival [5]. Cancer cells are able to produce lactate in a high oxygen environment (aerobic glycolysis), causing a phenomenon of cancer metabolism, known as the Warburg effect. The unusual utilisation of glucose by cancer cells results in the production of lactate, which is normally only produced during anaerobic respiration [20]. Hexokinase 2 (HKII) is a tissue-specific isoenzyme that phosphorylates glucose to glucose-6-phosphate (G-6-P) at the start of the glycolysis pathway, and its upregulation contributes to cancer cell glycolysis and the Warburg effect [21–23]. Hexokinase 2 maintains the high rate of glucose catabolism that is required for the survival of tumour cells, allowing them to sustain a higher rate of proliferation and resistance to cell death signals [21,23]. A variety of tumours are characterised by upregulated HKII expression, making it an attractive therapeutic target [23–25]. Oncogenes, such as *MYC* and *RAS*, and tumour suppressor genes, such as *p53*, are known to regulate metabolic enzyme expression and they can cause cancer if they happen to undergo mutations [20]. In addition, hypoxia-inducible factor 1-alpha (HIF1A) and *MYC* proteins cooperate to regulate the expression of *HKII* and *PDK1* genes [26]. An investigation of the Warburg effect in Burkitt's lymphoma (BL) cells, where the HIF1A protein was highly expressed in EBV-positive BL cell lines, showed that the inhibition of *MYC* activity led to decreased expression of *MYC*-dependent genes and LDH-A activity, implicating *MYC* as the master regulator of aerobic glycolysis in these cells [26].

Most recently, Bhalla et al. investigated the role of hypoxia in DLBCL [27]. They demonstrated that the up-regulation of HIF1A resulted in repressed protein translation, however HKII was selectively translated by eIF4E1 to promote DLBCL growth in vitro and in vivo under hypoxic stress. Their findings suggest HKII as a key metabolic driver of the DLBCL phenotype. It has also been shown that acquired resistance in rituximab-resistant lymphoma cell lines (RRCL) was associated with the deregulation of glucose metabolism and an increase in the apoptotic threshold, leading to chemotherapy resistance, where RRCL expressed higher levels of HKII. Targeting HKII in these cells led to decreased resistance, implying that increased HKII levels in aggressive lymphoma causes chemotherapy resistance, while also identifying this as a potential therapeutic target [22]. Many HKII inhibitors have been effective in anti-cancer therapies, such as 3-bromopyruvate (3-BP), which was found to inhibit HKII, activate the mitochondrial cell death pathway, and deplete levels of ATP [22], and it was also shown to induce apoptosis in a breast cancer cell line (MDA-MB-231) [28].

The main aim of this study was to investigate the genetic association between miRSNPs that were previously implicated in tumorigenesis and/or NHL susceptibility and prognosis, based on a comprehensive review of recent literature [19,29,30]. We genotyped 39 miRSNPs using a multiplex PCR and matrix assisted laser desorption time-of-flight (MALDI-TOF) mass spectrometry (MS) MassARRAY® system in our Genomics Research Centre Genomics Lymphoma Population (GRC GLP-non-Hodgkin lymphoma) cohort. After basic association testing, three SNPs in *MIR143* were identified to be significantly associated with NHL, with one SNP, rs17723799, remaining significant after Bonferroni correction for multiple testing (p -value = 0.015). After logistic regression testing the same three SNPs in *MIR143* were significantly associated with reduced risk of NHL in the Additive model: rs3733846 (Odds ratio (OR) [95% confidence interval (CI)] = 0.54 [0.33–0.86], p = 0.010), rs41291957 (OR [95% CI] = 0.61 [0.39–0.94], p = 0.024), and rs17723799 (OR [95% CI] = 0.43 [0.26–0.71], p = 0.0009). As *HKII* is a known target gene for mature hsa-miR-143 (miR-143), our secondary aim examined HKII expression in four patient-derived NHL cell lines, as compared to a metastatic breast cancer (MDA-MB-231) and melanoma (MDA-MB-435) cell line, as well as two healthy control subjects to assess the potential functional link between miR-143 regulation and HKII levels in NHL.

2. Materials and Methods

2.1. Study Population

The GRC-GLP retrospective cohort consists of 300 NHL cases and 140 healthy controls. All of the samples are of Caucasian origin with Australian/British/European grandparents with no family history of a haematological malignancy. The cases were matched according to age- (within five years), sex-, and ethnicity with healthy cancer-free controls. Cases were collected between 2010 and 2014 from the Princess Alexandra Hospital in Brisbane, and the GRC clinic in Mermaid Waters on the Gold Coast. The case cohort mainly consists of FL (n = 95) and DLBCL (n = 88), with 79 cases being unclassified as NHL or “Other B-cell”. B-cell chronic lymphocytic lymphoma (CLL), cutaneous T-cell lymphoma, Mantle cell lymphoma (MCL), Splenic marginal zone lymphoma (SMZL), Mucosa-associated lymphoid tissue lymphoma (MALT), and Burkitt’s lymphoma (BL) make up the remaining subtypes, in the order of frequency from highest to lowest (Table 1). Patients and healthy volunteers were required to complete a personal questionnaire and provide written consent to participate in research. The cohort is comprised of 48% male and 52% female participants, with the mean age of cases 63.72 years (standard deviation (SD) = 12.95 years) and the mean age of controls 63.14 years (SD = 13.03 years). In addition, 35 new NHL cases were received in 2016 (collected in 2014), comprised of 71% male and 29% female, with an average age of 59.6 years (SD = 13.6 years). Ethics for the collection and the use of participant samples was approved by the Queensland University of Technology Research Ethics Committee (approval number 1400000125). All of the subjects gave written informed consent, in accordance with the Declaration of Helsinki.

Table 1. Retrospective Genomics Lymphoma Project-non-Hodgkin lymphoma (GLP-NHL) cohort comprising of different NHL subtypes and the number of cases for each subtype.

NHL Subtype	No. of Cases in the Cohort
FL	95
DLBCL	88
Other B-cell/NHL/unclassified	79
B-CLL	16
T-cell lymphoma	7
Mantle cell lymphoma (MCL)	6
Splenic marginal zone lymphoma (SMZL)	4
Mucosa-associated lymphoid tumour (MALT)	3
Burkitt's lymphoma (BL)	2
Total	300

GLP: Genomics lymphoma population; NHL: Non-Hodgkin lymphoma; FL: Follicular lymphoma; DLBCL: Diffuse large B cell lymphoma; CLL: Chronic lymphocytic lymphoma.

2.2. Genomic DNA Extraction

Genomic DNA (gDNA) was extracted from whole blood collected into EDTA tubes using an in-house salting-out method, as evaluated by Chacon et al. [31]. DNA concentration and purity was measured using the NanoDrop™ ND-1000 spectrophotometer (ThermoFisher Scientific Inc., Waltham, MA, USA) before dilution to 15–20 ng/μL and storage as stock gDNA at 4 °C.

2.3. miRSNP Selection and iPlex Primer Design

39 cancer-related microRNA-related SNPs (miRSNPs) were selected for genotyping by chip-based multiplex PCR and MALDI-TOF MS (Agena MassARRAY®, San Diego, CA, USA), following a comprehensive review of the recent literature related to miRNA biomarkers in tumourigenesis, including breast cancer and/or non-Hodgkin lymphoma [18]. Forward, reverse, and extension primers were designed for each SNP using the MassARRAY® AgenaCx design software. Primers were manufactured by Integrated DNA Technologies (IDT, Singapore) and primer sequences are available on request.

2.4. Primary Multiplex PCR

Genotyping was performed using the iPlex™ GOLD Reagent Kit (Agena), according to the manufacturer's instructions. Forward and reverse primers were pooled, with extension primers being pooled according to mass using the linear primer adjustment method. All of the PCR reactions were performed in 96-well reaction plates in a Veriti™ 96-well Thermal Cycler (ThermoFisher Scientific Inc., Waltham, MA, USA).

2.5. MALDI-TOF MS and Data Analysis

Prior to each run, the extension products were spotted on to the SpectroCHIP® (Agena) using the Agena MassARRAY® Nanodispenser (Agena), which was immediately loaded into the Agena MassARRAY® Analyser 4 for genotype detection of each SNP in the assay by the SpectroAcquire v4.0 software (Agena). Final data analysis was performed using the MassARRAY® Typer software v4.0 (Agena).

2.6. In-Vitro Culture of Cell Lines and Primary Lymphocytes

All of the commercial patient-derived NHL B-lymphoid (SU-DHL-4, Raji, Mino, Toledo), breast cancer (MDA-MB-231), and melanoma (MDA-MB-435) cell lines were previously acquired from the American Type Culture Collection (ATCC®, Manassas, VA, USA). The cell lines were authenticated by the GenePrint® 10 System, according to the manufacturer's instructions as a service provided by the School of Biomedical Sciences, QUT. For experiments, the cells were seeded into T-75 flasks

with 20 mL RPMI-1640 or DMEM growth media (Invitrogen, ThermoFisher), supplemented with 10% foetal bovine serum and 1% Penicillin-Streptomycin antibiotic. Cultures were incubated at 37 °C under 5% CO₂ conditions. The cells were harvested into TRIzol[®] and Runx protein-lysis buffer for RNA and protein extraction (see below) and then pelleted for DNA extraction at 90% confluence and 95% viability. Cell counts and viability were assessed via the Trypan Blue exclusion method on a TC10[™] automated cell counter (Bio-Rad, Hercules, CA, USA). Primary lymphocytes from healthy subjects were isolated from peripheral blood mononuclear cells (PBMCs) using the Ficoll-Histopaque centrifugation method (Sigma-Aldrich, St. Louis, MI, USA). Cells were washed with 1x PBS and then plated into T-75 culture flasks in RPMI-1640 with Phytohaemagglutinin (PHA) (Sigma-Aldrich) mitogen for up to 24 h to allow for monocytes to adhere to the flask. After a minimum of three hours, the lymphocytes in suspension were removed and plated into new flasks with PHA and interleukin 2 (IL-2) (Sigma-Aldrich). The lymphocytes were left to proliferate at 37 °C under 5% CO₂ conditions for five days. Live lymphocyte counts were performed on a haemocytometer and viability assessed by Trypan blue dye.

2.7. Validation of Genotyping by MassARRAY[®] and Genotyping of Cell Lines and Healthy Controls by Sanger Sequencing

The genotype plots on the MassARRAY[®] system were assessed and manually genotyped for scattered data not assigned by the analysis software. For statistically significant SNPs ($p < 0.05$), a subset of samples was validated by Sanger sequencing (SS) using primers (Table 2) to ensure the accuracy of genotype calls by the Agena MassARRAY[®] system. Briefly, amplified PCR products were treated with Exo-SAP for use in Big Dye Terminator v3.1 (ThermoFisher Scientific, Life Technologies) sequencing reactions, and analysed on the 3500 Genetic Analyser (ThermoFisher Scientific, Life Technologies). Sequencing data for each sample chromatogram was assessed using Chromas Lite 2.1.1 software). The genotypes obtained were found to be 100% concordant between the MassArray[®] and SS. gDNA was extracted from NHL cell line pellets using the GenePrint[®] 10 System, according to the manufacturer's instructions.

Table 2. Sanger sequencing primers for three analysed *MIR143* single nucleotide polymorphisms.

SNP	Forward Primer (5'–3')	Reverse Primer (5'–3')	Accession ID
rs3733846	TGTTTGCCATCTCCTCT	CCTTCCCATGGAGCTTTGT	NC_000005.1
rs41291957	CAGGAAACACAGTTGTGAGG ¹	AGGAGAAGGGGTGTTAGAGG ¹	NC_000005.1
rs17723799	TGGTCATCCAATCAGCCACC	GGAAGGGACCCTGTCAACTG	NC_000005.1

¹ Same sequence as MassARRAY[®] primer design (AgenaCx); SNP: Single Nucleotide Polymorphism.

2.8. RNA Extraction, cDNA Synthesis and q-PCR

Total RNA was extracted from cell lines and control primary lymphocyte homogenates using the Direct-zol[™] RNA MiniPrep extraction kit (Zymo Research, Irvine, CA, USA), according to the manufacturer's instructions. Total RNA was converted to complementary DNA (cDNA) while using the iScript[™] cDNA Synthesis Kit (Bio-Rad Laboratories), according to the manufacturer's instructions. Each q-PCR was performed in triplicate, with 120 ng cDNA being amplified with SYBR[®] Green PCR master mix (Promega, Madison, Wisconsin, USA) and *HKII* forward (5'-CCCGGAAAGCAACTGTTTG-3') and reverse (5'-ACCGGTGTTGAGAAGCTCTG-3') primers in a 10ul final reaction volume on the QuantStudio 7 instrument under the following reaction conditions: 50 °C for 2 min (×1 cycle), 95 °C for 3 min (×1 cycle), 95 °C for 3 s, and 60 °C for 30 s (×50 cycles). PCR efficiency for the *HKII* primer pair was 103%, with a single peak melt curve. *18S* was used as an endogenous control and gene expression was calculated using the relative quantification method ($\Delta\Delta$ Ct). A student's *t*-test was used to calculate the significant differences between *HKII* expression profiles in cell lines and healthy controls.

2.9. Protein Detection by Western Blot

Total protein was isolated from cell lysates using the RUNX protein-lysis buffer (20 mM HEPES, 25% glycerol, 1.5 mM MgCl₂, 420 mM NaCl, 0.5 mM DTT, 0.2 mM EDTA, 0.5% Igepal CA-630, 0.2 mM Na₃VO₄, 1 mM PMSF, and dH₂O containing a protease and phosphatase inhibitor cocktail). Protein concentration was measured using the QubitTM Protein Assay Kit (Invitrogen). For SDS-PAGE (Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis), 30 µg total protein was separated for 1.5 h at a constant 120 V in a mini-PROTEAN Tetra unit. The HKII protein (109 kDa) was detected using a mouse anti-HKII primary antibody (Abcam, ab104836) and an HRP-conjugated anti-mouse IgG secondary antibody (Cell Signaling, #7076). Sample loading was normalised using secondary HRP-conjugated anti-Beta-actin (Cell Signaling, #5125S). Enhanced chemiluminescence detection (ClarityTM, ECL, Bio-Rad) of HKII and Beta-actin was performed using the Fusion Spectra chemiluminescent system (Vilber Lourmat, Fisher Biotec) and optical density quantitation was assessed using Image J software [32].

2.10. Statistical Analysis

Chi-square (χ^2) analysis was used to determine the differences in genotype and allele frequencies between case and control samples, where a p -value < 0.05 was considered to be statistically significant. Genetic association and haplotype analysis of genotyping data was performed using the latest Plink v1.09 [33], RStudio [34], and HaploView [35] software. Quality control (QC) filters were applied to the full dataset consisting of 39 variants, 300 cases, and 140 controls. Following basic association testing, where correction for multiple testing was applied in the filtered dataset, a mixed model logistic regression test adjusting for the covariate ‘sex’ was performed in association with NHL with risk determined by OR and CI that were set at 95%.

2.11. NHL-GWAS Replication Dataset

Cerhan et al. [36] conducted a meta-analysis of three new NHL genome-wide association studies (GWAS): NCI, GELA, and MAYO_DLBCL, and one previous scan (SF), including around 3857 cases and 7665 healthy controls of European ancestry (InterLymph Consortium), to identify the genetic susceptibility loci for DLBCL, with further genotyping of nine SNPs in 1359 cases and 4557 healthy controls. For the first three groups, NCI, GELA, and MAYO_DLBCL, data was genotyped and for the scan, SF, data was imputed. We requested genotype summary statistics from the authors, for SNPs to be included in the GWAS that were in high linkage disequilibrium (LD) ($D' = 1.00$), with significantly associated SNPs in our SNP MassARRAY[®] panel, i.e., rs12659504 and rs878008, in order to investigate the replication of our findings.

3. Results

3.1. Genetic Association of MIR143 SNPs and NHL Risk

In this study, 39 cancer-related microRNA-related SNPs (miRSNPs) were selected for genotyping by chip-based multiplex PCR and MALDI-TOF MS in the GLP-NHL population of 300 cases and 140 healthy controls. Allele frequencies for each variant in the GLP-NHL case-control cohort were established and the minor allele frequencies (MAFs) were compared to those that were observed in European and British populations in the Hapmap 1000G database (Table 3). After the QC filters were applied, basic allelic association testing revealed two SNPs in *MIR143/145* (rs3733846 [OR = 0.56, $p = 0.012$], rs41291957 [OR = 0.56, $p = 0.008$]), and one SNP in the promoter region of *MIR143* (rs17723799 [OR = 0.42, $p = 0.0004$]) on chromosome 5 to be significantly associated with a reduced risk of NHL (Table 3). After correction for multiple testing, the rs17723799 SNP remained statistically significant using the Bonferroni ($p = 0.015$) adjustment method (Table 3). A logistic regression test adjusted for the covariate ‘sex’ revealed the same three SNPs in *MIR143* to be significantly associated with the reduced risk of NHL in the Additive model: rs3733846 (OR [95% CI] = 0.54 [0.33–0.86],

$p = 0.010$), rs41291957 (OR [95% CI] = 0.61 [0.39–0.94], $p = 0.024$), and rs17723799 (OR [95% CI] = 0.43 [0.26–0.71], $p = 0.0009$) (Table 4). Genotype and allele frequencies were determined for the three *MIR143* SNPs and were compared with 1000G and gnomAD MAF percentages (Table 5). The rs17723799 SNP was identified to have a significant difference in allele ($p = 0.013$) and genotype ($p = 0.039$) frequencies between the case and control sample cohorts (Table 5). A mixed model logistic regression test adjusted for the covariate ‘sex’ revealed the rs17723799 SNP to be significantly associated with a reduced risk of NHL in the Additive (OR [95% CI] = 0.43 [0.26–0.71], $p = 0.0009$), Dominant (OR [95% CI] = 0.54 [0.33–0.88], $p = 0.015$), Over-dominant (OR [95% CI] = 0.56 [0.33–0.92], $p = 0.024$), and Log-additive (OR [95% CI] = 0.58 [0.38–0.90], $p = 0.017$) models of inheritance (Table 6). The two other *MIR143* SNPs in LD with rs17723799, i.e., rs3733846, and rs41291957 (Table 7) were also shown to be associated with a significantly reduced risk of NHL after logistic regression testing in the Additive model only: rs3733846 (OR [95% CI] = 0.54 [0.33–0.86], $p = 0.010$) and rs41291957 (OR [95% CI] = 0.61 [0.39–0.94], $p = 0.024$). Analysis of the three SNPs was also performed in association with the two most common NHL subtypes that are present in the cohort, i.e., DLBCL and FL, however the results were non-significant for all three SNPs ($p > 0.05$). This could be due to the small number of cases representing these subtypes and the reduced statistical power. All three SNPs were validated by Sanger sequencing in 5–10 randomly selected cases and controls, with the genotype of all samples being confirmed.

Table 3. Basic association testing for 39 genotyped SNPs showing minor allele frequencies (MAFs), Hardy-Weinberg Equilibrium (HWE) score for controls, unadjusted odds ratios, and p -values and adjusted p -values after correction for multiple testing. A1, minor allele; A2, major allele; MAF, minor allele frequency; NCHROBS, number of chromosomes observed; 1000G, 1000 Genomes Database; HWE, Hardy-Weinberg Equilibrium; OR, odds ratio.

Chr	miRNA/Target Gene	SNP	A1	A2	MAF/NCHROBS	MAF 1000G	HWE UNAFF	OR	p -Value	Adjusted p -Value
1	<i>E2F2</i>	rs2075993	G	A	0.5/862	G = 0.3488/1747	0.2368	0.885	0.4535	1
1	<i>GEMIN3</i> 3'-UTR	rs197412	C	T	0.4092/870	C = 0.4744/2376	0.6013	0.8779	0.4301	1
2	hsa-miR-155-3p	rs4672612	A	G	0.338/858	A = 0.3878/1942	0.5476	1.095	0.6006	1
4	<i>TET2</i>	rs7670522	A*	C	0.4701/870	C = 0.3600/1803	1.000	1.089	0.6041	1
4	hsa-miR-4330/5100	rs2647257	T	A	0.408/848	T = 0.2386/1195	0.3732	0.9403	0.7094	1
5	hsa-miR-224-5p	rs12719481	G	A	0.2719/868	G = 0.3670/1838	1.000	1.072	0.7052	1
5	hsa-miR-143	rs3733846	G	A	0.1367/878	G = 0.2063/1033	1.000	0.5646	0.012	0.467
5	hsa-miR-143	rs17723799	T	C	0.1129/868	T = 0.1118/560	1.000	0.423	0.0004	0.015
5	hsa-miR-143	rs41291957	A	G	0.1412/878	A = 0.1214/608	1.000	0.5624	0.008	0.326
5	hsa-miR-145	rs353291	C	T	0.4237/826	C = 0.3608/1807	0.1473	1.204	0.2648	1
5	hsa-miR-146-a	rs2910164	C	G	0.2204/862	C = 0.2797/2881	0.5978	1.19	0.3908	1
5	hsa-miR-218-2	rs11134527	A	G	0.2189/868	A = 0.3462/1734	0.7982	1.061	0.7635	1
6	<i>XPO5</i>	rs11077	G	T	0.4255/872	G = 0.4036/2021	0.3862	0.9566	0.7869	1
6	<i>TAB2</i>	rs9485372	A	G	0.1965/850	A = 0.2408/1206	1.000	0.844	0.4084	1
6	<i>ESR1, C6orf97</i>	rs2046210	A	G	0.3353/850	A = 0.4121/2064	1.000	1.277	0.1659	1
8	<i>TP53</i>	rs896849	G	A	0.1501/866	G = 0.2183/1093	0.4915	0.9369	0.7695	1
8	<i>CASC21</i>	rs13281615	G	A	0.4417/840	G = 0.4912/2460	0.3853	0.7541	0.08388	1
8	<i>AGO2</i>	rs3864659	C	A	0.1023/860	C = 0.1436/719	0.6026	1.147	0.6275	1
8	<i>AGO2</i>	rs4961280	A	C	0.1835/872	A = 0.1490/746	0.7390	1.416	0.1075	1
9	hsa-miR-101-2	rs462480	G	T	0.3984/876	G = 0.4451/2229	0.7246	1.001	0.9948	1
10	hsa-miR-608	rs4919510	G	C	0.1979/874	G = 0.3638/1822	1.000	1.115	0.6034	1
10	hsa-miR-202	rs12355840	C	T	0.1368/848	C = 0.3189/1597	0.6598	0.494	0.1049	1
11	hsa-miR-210	rs1062099	C	G	0.1701/876	C = 0.1649/826	0.3067	1.29	0.2536	1
11	<i>LSP1</i>	rs3817198	C	T	0.3289/836	C = 0.2155/1079	0.7152	0.7085	0.04181	1
11	<i>TMEM45, BARX2</i>	rs7107217	A	C	0.4883/858	A = 0.4876/2442	0.5987	0.8992	0.514	1
12	<i>KRAS</i> 3'-UTR	rs61764370	C	A	0.0962/852	C = 0.0347/174	1.000	0.6795	0.1473	1
12	hsa-miR-196-a2	rs11614913	T	C	0.4205/880	T = 0.333/1666	0.226	0.928	0.6501	1
12	pre-miR-618	rs2682818	A	C	0.1465/874	A = 0.2424/1214	0.3061	1.115	0.6448	1
14	<i>HIF1A</i> 3'-UTR	rs2057482	T	C	0.1250/856	T = 0.2424/1214	0.6797	1.106	0.6831	1
14	<i>DICER1</i>	rs3742330	G	A	0.0878/854	G = 0.1382/692	0.2499	1.261	0.4423	1
14	<i>DICER1</i>	rs1057035	C	T	0.3709/852	C = 0.1723/863	0.854	0.8726	0.4174	1
16	<i>TOX3</i>	rs8051542	T	C	0.4322/856	T = 0.3133/1569	1.000	0.7698	0.1095	1

Table 3. Cont.

Chr	miRNA/ Target Gene	SNP	A1	A2	MAF/ NCHROBS	MAF 1000G	HWE UNAFF	OR	p-Value	Adjusted p-Value
16	TOX3	rs3803662	A	G	0.2207/852	A = 0.4403/2205	1.000	0.9065	0.6119	1
18	hsa-miR-143-5p	rs4987859	T	C	0.0631/856	T = 0.0477/239	0.597	0.744	0.3466	1
18	hsa-miR-27-a-5p	rs4987852	C	T	0.0667/854	C = 0.0190/95	0.4495	1.081	0.811	1
18	hsa-miR-27-a-5p	rs1016860	T	C	0.1175/868	T = 0.1166/584	1.000	0.8423	0.4808	1
21	hsa-miR-155 HG	rs987195	G	C	0.0917/840	G = 0.1472/737	0.361	0.6702	0.1265	1
21	hsa-miR-155	rs12482371	C	T	0.1632/858	C = 0.4151/2079	0.5662	0.7506	0.1748	1
X	hsa-miR-221/222	rs34678647	T	G	0.0375/667	T = 0.1423/537	0.1782	0.433	0.05456	1

Table 4. Logistic regression analysis of *MIR143* SNPs in association with NHL showing p-values after adjusting for covariate ‘sex’. A1, minor allele.

Chr	Gene	SNP	A1	Model	OR (CI 95%)	p-Value
5	<i>MIR143</i>	rs3733846	G	Additive	0.54 (0.34–0.87)	0.010
5	<i>MIR143</i>	rs41291957	A	Additive	0.61 (0.39–0.94)	0.024
5	<i>MIR143</i>	rs17723799	T	Additive	0.61 (0.26–0.71)	0.0009

Table 5. Allele and genotype frequencies of SNPs located in *MIR143* in the GLP-NHL case and control populations in comparison to allele frequencies obtained from HapMap 1000G and gnomAD databases. HWE, Hardy-Weinberg Equilibrium (HWE) test score.

rs17723799								
	Allele			Genotype			p-Value	HWE
	C (%)	T (%)	p-Value	C/C (%)	C/T (%)	T/T (%)		
Controls	234 (84.8)	42 (15.2)	0.013	99 (71.7)	36 (26.1)	3 (2.2)	0.039	1.000
Cases	536 (90.5)	56 (9.5)						
MAF	770 (88.7)	98 (11.3)						
1000G (%)	88.8	11.2						
gnomAD (%)	86.9	13.1						
rs3733846								
	Allele			Genotype			p-Value	HWE
	A (%)	G (%)	p-Value	A/A (%)	A/G (%)	G/G (%)		
Controls	231 (83)	47 (17)	0.057	96 (69)	39 (28)	4 (3)	0.167	1.000
Cases	527 (88)	73 (12)						
MAF (%)	758 (86.3)	120 (13.7)						
1000G (%)	79.4	20.6						
gnomAD (%)	84.4	15.6						
rs41291957								
	Allele			Genotype			p-Value	HWE
	G (%)	A (%)	p-Value	G/G (%)	G/A (%)	A/A (%)		
Controls	229 (82.4)	49 (17.6)	0.043	94 (67.6)	41 (29.5)	4 (2.9)	0.070	1.000
Cases	525 (87.5)	75 (12.5)						
MAF (%)	754 (85.9)	124 (14.1)						
1000G (%)	87.9	12.1						
gnomAD (%)	84	16						

Table 6. Mixed model logistic regression analysis adjusted for the covariate ‘sex’ for the rs17723799 SNP in *MIR143*.

Model	Genotype	Controls (%)	Cases (%)	χ^2	OR (95% CI)	p-Value
Allelic	T vs. C	39/459	36/178	12.84	-	0.0003
Additive	-	-	-		0.43 (0.26–0.71)	0.0009
Co-dominant	CC	97 (71.9)	243 (82.4)		1.00 (reference)	
	CT	35 (25.9)	48 (16.3)		0.55 (0.33–0.91)	
	TT	3 (2.2)	4 (1.4)		0.47 (0.10–2.23)	0.051
Dominant	CC	97 (71.9)	243 (82.4)		1.00 (reference)	
	CT–TT	38 (28.1)	52 (17.6)		0.54 (0.33–0.88)	0.015
Recessive	CC–CT	132 (97.8)	291 (98.6)		1.00 (reference)	
	TT	3 (2.2)	4 (1.4)		0.53 (0.11–2.52)	0.438
Over-dominant	CC–TT	100 (74.1)	247 (83.7)		1.00 (reference)	
	CT	35 (25.9)	48 (16.3)		0.56 (0.33–0.92)	0.024
Log-additive	-	135 (31.4)	295 (68.6)		0.58 (0.38–0.90)	0.017

Table 7. Linkage disequilibrium (D') values for SNPs on chromosome 5 in HapMap 1000G European (EUR) Central European (CEU)/Great Britain (GBR) populations.

SNP	Location (GRCh38)	rs3733846	rs12659504	rs878008	rs17723799	rs41291957
rs3733846	149,425,059	-	1.000/1.000	1.000/1.000	1.000/1.000	0.721/0.941
rs12659504	149,425,442	-	-	1.000/1.000	1.000/1.000	0.721/0.942
rs878008	149,425,488	-	-	-	1.000/1.000	0.721/0.942
rs17723799	149,427,514	-	-	-	-	0.999/0.999
rs41291957	149,428,827	-	-	-	-	-

3.2. Replication of Summary Statistics in a Large EUROPEAN NHL GWAS Meta-Analysis

As we did not have access to an in-house replication NHL population, we interrogated the complete list of SNPs (700,000+ loci) that are available on the HumanOmniExpress-12-v1-1-C chip (Illumina) studied in three NHL GWAS for the presence of SNPs found to be significantly associated with NHL in our study. Unfortunately, our three SNPs in *MIR143* were not included in the GWAS, however we were able to find two SNPs (rs12659504 and rs878008) on chromosome 5 in the GWAS in high LD ($r^2 = 1.00$, $D' = 1.00$) with rs3733846 and rs17723799 (Table 7). A high linkage disequilibrium (LD) score identified strong non-random association of nearby variants on the same chromosome. Additionally, a meta-analysis of three individual GWAS and one imputation scan performed by Cerhan et al. [37], which included 3855 cases and 7664 controls, showed a significantly reduced risk of NHL for both of these SNPs: rs12659504 (OR [95% CI] = 0.91 [0.84–0.99], $p = 0.033$) and rs878008 (OR [95% CI] = 0.92 [0.84–1.00], and $p = 0.041$) (InterLymph Consortium).

3.3. Haplotype Analysis of *MIR143* SNPs in LD on Chromosome 5

We performed haplotype analysis for the three SNPs (rs3733846, rs41291957, and rs17723799) that were located in *MIR143* on chromosome 5 found to be significantly associated with a reduced risk of NHL in our case-control cohort. The most commonly inherited haplotype was haplotype 3 (A-G-C), with a frequency of 0.82 (82%) in case samples and 0.79 (79%) in control samples (Table 9). A significantly reduced risk of NHL was observed with the inheritance of haplotype 4 (G-A-T) with all three minor alleles, which occurred at a frequency of 0.09 (9%): OR [95% CI] = 0.42 [0.18–1.00], $p = 0.049$) (Table 10). A linkage disequilibrium plot showed moderate linkage between the three SNPs when being assessed using Haploview 4.2 (Figure 1).

Table 8. Summary statistics for three European NHL-GWAS with an imputation scan (InterLymph Consortium) and meta-analysis (Cerhan et al. 2014) for two SNPs in high LD with rs17723799 on chromosome 5. EAF, effect allele frequency; OR, odds ratio, CI, confidence interval.

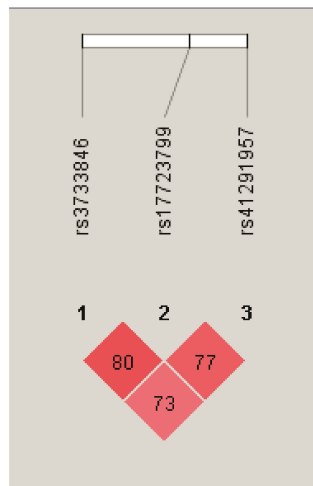
SNP	Chr	Location (GRCh38)	Group	Controls	Cases	Effect Allele	EAF Controls	EAF Cases	OR	CI (95%)	p-Value
<i>rs12659504</i>	5	149,425,442	NCI	6221	2661	G	0.1502	0.1359	0.93	0.84–1.02	0.120
<i>rs12659504</i>	5	149,425,442	GELA	525	548	G	0.1418	0.125	0.90	0.70–1.15	0.392
<i>rs12659504</i>	5	149,425,442	MAYO_DLCL	171	392	G	0.173	0.1569	0.79	0.54–1.14	0.211
<i>rs12659504</i>	5	149,425,442	SF	747	254	G	0.133	0.1205	0.89	0.66–1.21	0.456
<i>rs12659504</i>	5	149,425,442	Meta-analysis	7664	3855				0.91	0.84–0.99	0.033
<i>rs878008</i>	5	149,425,488	NCI	6221	2661	C	0.1501	0.1362	0.93	0.84–1.02	0.132
<i>rs878008</i>	5	149,425,488	GELA	524	548	C	0.1396	0.1242	0.91	0.71–1.16	0.445
<i>rs878008</i>	5	149,425,488	MAYO_DLCL	172	392	C	0.172	0.1548	0.78	0.53–1.13	0.188
<i>rs878008</i>	5	149,425,488	SF	747	253	C	0.1307	0.1199	0.90	0.66–1.23	0.516
<i>rs878008</i>	5	149,425,488	Meta-analysis	7664	3854				0.92	0.84–1.00	0.041

Table 9. Haplotype frequencies for analysed *MIR143* SNPs in GLP-NHL case and control populations.

Haplotype	rs3733846	rs41291957	rs17723799	Frequencies Cases	Frequencies Controls
1	A	A	C	0.03608	0.03255
2	A	A	T	0.00000	0.00000
3	A	G	C	0.81643	0.79232
4	A	G	T	0.02582	0.00517
5	G	A	C	0.02159	-
6	G	A	T	0.06733	0.14260
7	G	G	C	0.03092	0.02067
8	G	G	T	0.00182	0.00670

Table 10. Association of haplotypes for analysed *MIR143* SNPs in GLP-NHL case and control populations.

Haplotype	rs3733846	rs41291957	rs17723799	Haplotype Frequencies	OR (CI 95%)	p-Value
1	A	G	C	0.80883	0.90 (0.43–1.90)	0.7799
2	A	G	T	0.01898	6.34 (0.60–67.07)	0.1246
3	G	A	C	0.01461	Inf (Inf-Inf)	0.0000
4	G	A	T	0.09137	0.42 (0.18–1.00)	0.0495
5	G	G	C	0.02762	1.22 (0.38–3.88)	0.7370
rare	*	*	*	0.00359	0.12 (0.00–2.95)	0.1920

**Figure 1.** Linkage disequilibrium (LD) plot for analysed polymorphisms in the *MIR143* host gene using Haploview software version 4.2 (Daly Lab, Broad Institute, Cambridge, MA, USA).

3.4. Genotyping of rs17723799 (C>T) in Cell Lines and Healthy Control Samples

We genotyped four commercially sourced, immortalised NHL cell lines (SU-DHL-4, Raji, Mino and Toledo), a metastatic breast cancer cell line (MDA-MB-231), a melanoma cell line (MDA-MB-435), and two healthy cancer-free control subjects for the *MIR143* rs17723799 SNP. All of the NHL cell lines were homozygous for the wild type risk allele C (CC). The breast cancer cell line was homozygous wild type CC. The melanoma cell line was homozygous for the minor allele T (TT). Both the healthy control subjects were homozygous for the wild type C allele (CC). As the minor T allele is reported to be present at a frequency of 0.11 (11%) in the healthy European population (Table 5), these results were not surprising; however, it was interesting that all NHL cell lines were identified to be homozygous for the wild type allele C, which was observed as the risk allele for NHL in our genetic association analysis.

3.5. miR-143 Target Gene *HKII* Expression in Cancer Cell Lines and Healthy Control Lymphocytes

miR-143 inhibits the expression of its target gene *HKII* by binding to a conserved recognition motif that is located in the 3'-UTR of mRNA. We examined *HKII* expression at the mRNA and protein level in four NHL cell lines (SU-DHL-4, Raji, Mino, Toledo), a metastatic breast cancer cell line (MDA-MB-231), and a melanoma cell line (MDA-MB-435). In addition, two healthy control donors were included as a reference for basal *HKII* expression levels for statistical analysis.

3.5.1. *HKII* Gene Expression Is Increased in NHL Compared to Healthy Controls

q-PCR analysis of *HKII* mRNA transcript expression showed increased levels of *HKII* mRNA transcript in the four patient-derived NHL cell lines (SU-DHL-4, Raji, Mino, and Toledo) when compared to lymphocytes that are derived from two healthy control subjects, however the differences in expression levels were not statistically significant (Figure 2). Western blot and band quantitation analysis confirmed these findings, showing increased *HKII* protein levels in the NHL cell lines examined when compared to the controls (Figure 3).

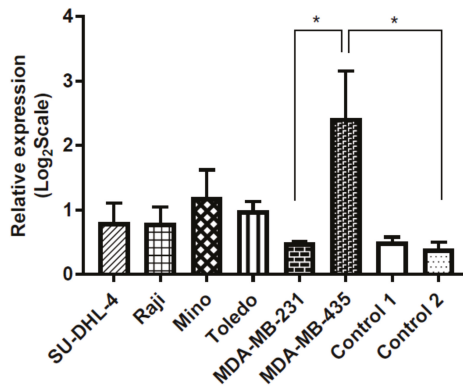


Figure 2. Hexokinase 2 (*HKII*) gene expression. q-PCR analysis of *HKII* mRNA transcript expression in NHL (SU-DHL-4, Raji, Mino, Toledo), breast cancer (MDA-MB-231) and melanoma (MDA-MB-435) cell lines as compared to healthy controls 1 and 2. *HKII* mRNA levels were increased in NHL cell lines compared to breast cancer cells and healthy controls, however expression levels were not significantly different. *HKII* mRNA levels were significantly increased in the melanoma (MDA-MB-435) cells when compared to breast cancer cells (MDA-MB-231) ($p = 0.044$), healthy control 1 ($p = 0.045$) and 2 ($p = 0.037$), but not significantly increased compared to the NHL cells. Relative expression normalized to *18S*, error bars = SEM and statistical significance calculated using paired Student's *t*-test and defined as: * $p < 0.05$.

3.5.2. Increased *HKII* Gene Expression May Be Associated with the MIR143 rs17723799 TT Genotype

q-PCR analysis demonstrated *HKII* mRNA transcript levels that were higher in the melanoma (MDA-MB-435) cells, correlating with the observed homozygous mutant rs17723799 TT genotype in these cells. *HKII* mRNA transcript levels were observed to be lower in all four NHL cell lines, and even lower in the breast cancer (MDA-MB-231) and healthy control cells, which correlated with the homozygous wild type rs17723799 CC genotype of these cells (Figure 2). A Student's *t*-test showed a significant difference in the relative *HKII* mRNA expression between melanoma cells and the breast cancer ($p = 0.044$), control subject 1 ($p = 0.045$), and control subject 2 ($p = 0.037$) cells. Western blot and band quantitation analysis demonstrated upregulated *HKII* protein expression in the melanoma cells, with lower expression being observed in the breast cancer cells and even lower in the four NHL cell lines examined (Figure 3). Surprisingly, the breast cancer cells expressed low *HKII* mRNA transcript, but high protein levels.

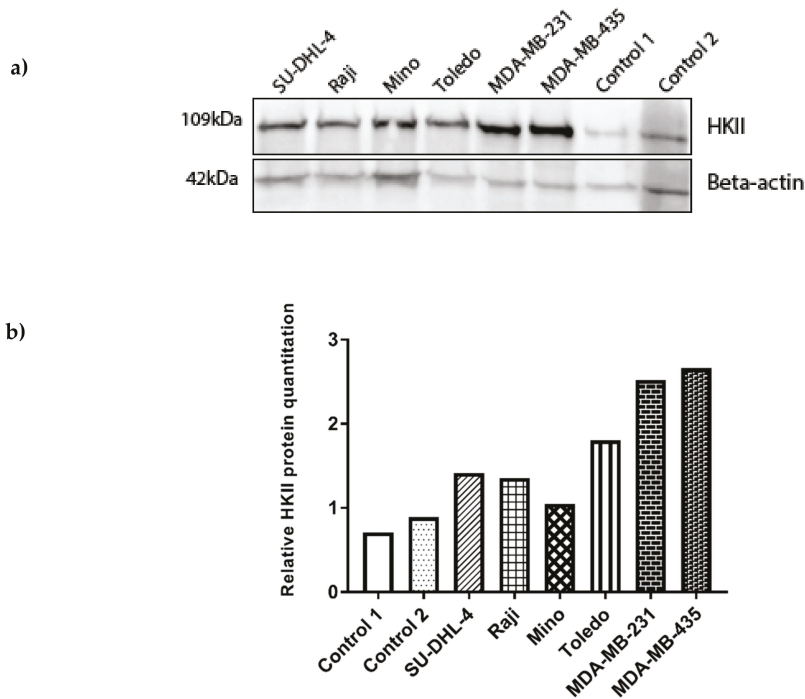


Figure 3. HKII protein detection. Western blot analysis of HKII protein (109kDa) expression in NHL (SU-DHL-4, Raji, Mino, Toledo), breast cancer (MDA-MB-231) and melanoma (MDA-MB-435) cell lines compared to healthy control subjects 1 and 2. 30 μ g total protein was loaded and samples were normalised against the Beta-actin (42kDa) loading control. (a) The HKII protein levels were increased in all four NHL cell lines compared to healthy control subjects 1 and 2, and decreased when compared to the breast cancer (MDA-MB-231) and melanoma (MDA-MB-435) cells. (b) HKII protein quantitation analysis compared the normalised intensity ratios between samples and showed the breast cancer (MDA-MB-231) and melanoma (MDA-MB-435) cells to have increased HKII levels as compared to NHL cells and healthy control subjects 1 and 2.

4. Discussion

Our study investigated 39 miRNAs previously implicated in human tumorigenesis that were identified from the literature, including those not previously reported to have an association with NHL susceptibility, but rather with disease prognosis, in Caucasians. Following genotyping by multiplex PCR and MALDI-TOF MS, subsequent genotypic association and haplotype analysis identified three SNPs on chromosome 5 (rs3433846, rs17723799, and rs41291957) in *MIR143* in moderately high LD, which together appear to confer protection against NHL in our GRC-GLP cohort. In particular, the rs17723799 SNP in the promoter region of *MIR143* remained significant at the $p < 0.05$ threshold after Bonferroni correction for multiple testing. We also provided evidence for replication in a large European NHL GWAS meta-analysis study.

One study identified the rs11614913 SNP (T > C) in hsa-miR-169-a2 (miR-196-a2) to be associated with NHL in a Chinese case-control cohort, where the presence of the risk C allele and CC genotype was shown to be more frequent in the case samples, conferring an increased risk (OR [95% CI] = 1.82 [1.16–2.85], $p = 0.009$) [36]. The authors confirmed that the variant genotype affected the expression of miR-196-a2, where the carriers of the CC genotype had significantly higher levels than those with only one C allele or the TT genotype. Interestingly, this SNP was also shown to increase risk of hepatocellular carcinoma in a Chinese cohort of 109 cases and 105 controls [38]; however, no functional analyses were

performed. The rs11614913 SNP was shown to be associated with decreased central nervous system (CNS) Acquired Immunodeficiency Syndrome (AIDS)-NHL in an American case-control cohort of 180 cases and 529 controls, where the CT genotype was more frequently observed in control samples and conferred protection (OR [95% CI] = 0.52 [0.27–0.99]). This study also showed an increased risk of systemic AIDS-NHL with the presence of the T allele in rs2057482, which creates a binding site for miR-196-a2 in the *HIF1A* 3'-UTR (OR [95% CI] = 1.73 [1.12–2.67]) [39]. We were not able to show a positive significant association for miR-196-a2 rs11614913 in our Australian/European cohort (OR [95% CI] = 0.93 [0.67–1.29], $p = 0.668$) (Table 3). This is most likely due to ethnicity differences in our cohort and the Chinese cohort of Li et al. [36]. Previous studies on this SNP have shown a significant association with cancer in predominantly Asian populations, but not in Caucasians, which may explain the lack of association in our cohort [40].

Other studies have shown the association of *MIR143/145* SNPs with reduced cancer risk in Chinese cohorts. Li et al. [41] conducted a case-control analysis of 12 SNPs in the promoter region of *MIR143/145* in 242 cases with colorectal cancer (CRC) and 283 healthy controls. In support of our findings in NHL, the rs3733846 (A > G) mutant genotype A/G was associated with a significantly reduced risk of CRC (OR [95% CI] = 0.57 [0.44–0.73], $p < 0.001$). A more recent study by Wu et al. [42] showed a functional association between the rs353293 SNP (G>A) AG/AA genotypes and a reduced risk of bladder cancer in 333 Chinese cases when compared to 536 healthy controls (OR [95% CI] = 0.64 [0.46–0.90], $p = 0.008$). In their study, *in vitro* luciferase reporter analysis was used to show a significantly reduced effect of the protective rs353293A allele as compared with the rs353293G allele on transcriptional activity ($p < 0.001$). The promoter transcriptional activity was identified to be reduced due to the polymorphism, which could also be the case for the rs17723799 SNP that we identified, however this would imply lower miR-143/145 levels and lower tumour suppressor activity, which appears to be in contradiction with their genotypic analysis data as well as ours, showing the protective effect of the polymorphism. The authors were not able to show whether serum miR-143/145 or the target gene expression levels were altered.

Epigenetic mechanisms, such as miRNAs, have been shown to regulate aerobic glycolysis in cancer cells through the targeting and down-regulation of glycolytic enzymes [14]. miR-143, an essential regulator of glycolysis [43], inhibits *HKII* expression by binding to a conserved recognition motif that is located in the 3'-UTR of *HKII* mRNA and has been observed to be inversely correlated with *HKII* levels in primary keratinocytes and in head and neck squamous cell carcinoma (HNSCC)-derived cell lines [44]. miR-143 and miR-155 have both been shown to regulate glycolysis by targeting *HKII* in breast cancer, with miR-155 being able to suppress miR-143 production through targeting the *C/EBPβ* transcription activator for miR-143 [45]. In addition, the transcriptional regulation of miR-145 was recently reviewed by Zeinali et al. [46]. The authors discuss miR-145 regulation by DNA-binding factors, including c-MYC, p53, forkhead transcription factors of the O class 1 and 3 (Foxo1 and Foxo3), and miR-143/145 level regulation by RREB1 in other cancers but not in lymphoma.

The significant *MIR143* SNP that was identified in this study (rs17723799) has been previously reported via *in-silico* analysis to affect aerobic glycolysis. The authors showed that this SNP overlaps with the transcription factor binding site in the *MIR143* promoter region (148783674–148788779, UCSC Genome Browser) and concluded that this SNP could potentially affect the regulation of miRNA biogenesis and alter the hsa-miR-143-3p and hsa-miR-143-5p levels, thereby affecting its target genes that directly or indirectly control glycolysis [14].

Following our genotypic analysis, *HKII* expression at the mRNA and protein level in NHL (SU-DHL-4, Raji, Mino, Toledo), breast cancer (MDA-MB-231), and melanoma (MDA-MB-435) cell lines was assessed to determine the potential functional link between *MIR143* regulation by a promoter polymorphism and *HKII* target gene expression in these cells, along with lymphocytes from two healthy control subjects. Our results showed higher *HKII* mRNA and *HKII* protein levels in the NHL cell lines when compared to healthy controls, however the levels were lower in the NHL cell lines compared to the breast cancer (protein only) and melanoma cell lines. Melanoma cells carrying the

MIR143 rs17723799 homozygous recessive TT genotype expressed a significantly increased level of *HKII* mRNA transcript and HKII protein in comparison to other cancer cells that were carrying the rs17723799 homozygous wild type CC genotype, indicating that this homozygous polymorphism may have an effect on the transcriptional regulation of *MIR143*. Interrogation of the GeneCards Database identified 38 promoters and enhancers for the *MIR143* gene with numerous transcription factor binding sites for each, implicating multiple mechanisms of regulation of this gene by multiple transcription factors in different tissue types. As we do not have access to an appropriate wild type melanoma cell line control, further studies on *MIR143* SNPs in melanoma may be beneficial. Our results also show an up-regulation of the HKII protein in the metastatic breast cancer (MDA-MB-231) cells, as reported by Palmieri et al. [24]. Interestingly, Western blot showed HKII levels were higher in the Toledo (DLBCL) cell line when compared to the other less aggressive NHL cell line subtypes.

Although not in NHL, but as a marker of non-small cell lung cancer (NSCLC), miR-143 expression in peripheral blood mononuclear cells (lymphocytes and monocytes) was significantly lower in NSCLC patients than in healthy individuals ($p < 0.0001$) [47]. Similarly, a study in renal cell carcinoma (RCC) identified the miR-143/145 cluster to be downregulated in RCC tissues when compared to adjacent non-cancerous tissues, with significantly higher HKII levels in RCC tissues as compared to non-cancerous tissues, confirming the tumour suppressive effect of the miR-143/145 cluster through targeting HKII [48]. One study in aggressive Burkitt's lymphoma (BL) showed that upregulated miR-143 by the regulation of PI3K/Akt prevented cell growth, implicating miR-143 as a tumour suppressor in NHL [49]. miR-143 has also been demonstrated to have an anti-tumour effect in leukaemia patients, where significantly lower miR-143 levels were observed in patients when compared to healthy controls, and the overexpression of miR-143 decreased DNA methyltransferase 3A (*DNMT3A*) mRNA and protein expression, thereby reducing cell proliferation, colony formation, and cell cycle progression, as well as increased apoptosis [50]. As there is a paucity of literature on *MIR143* and miR-143 expression and polymorphisms in NHL and its subtypes, we searched for significant eQTLs (expression quantitative trait loci) for the three analysed *MIR143* SNPs in the GTEx Portal [51]. GTEx analysis indicated increased *MIR143* host gene expression in skeletal muscle for heterozygous and homozygous mutant alleles in all three significant *MIR143* SNPs analysed (Figure S1). Although the relationship between miR-143/145 and its *MIR143* host gene is an open issue, the potentially increased *MIR143* host gene expression, as seen due to the polymorphisms in skeletal muscle, may enhance hsa-miR-143-3p and -5p processing, which poses a potential protective function for these SNPs in lymphocytes. The presence of the rs17723799 CT or TT genotype may increase *MIR143* expression and miR-143 transcription with the subsequent downregulation of *HKII* causing reduced glycolysis and lymphomagenesis. Although we did not measure mature miR-143 levels in the NHL cell lines or patient samples to correlate this with HKII expression, our genotypic analysis data indicate a possible protective functional effect of *MIR143* SNPs on miR-143 regulation, with a possible inhibitory effect on *HKII* expression in controls as compared to NHL cases. Further studies investigating *MIR143* and hsa-miR-143-3p/5p expression levels in different NHL subtypes and their association with HKII levels are therefore necessary. We have summarised the role of HKII and miR-143 in cancer in Figure 4a and a potential functional effect of the rs17723799 polymorphism in control cells when compared to NHL cells in Figure 4b.

A limitation of this study is the lack of expression data for mature hsa-miR-143-3p and hsa-miR-143-5p in the NHL cell lines or patient samples. We only had access to four NHL cell lines and we did not have access to RNA or protein from our retrospective patient cohort. Additionally, we were not able to compare HKII expression in NHL cell lines polymorphic for the rs17723799 SNP, as they were all wild type CC, and therefore further studies investigating miR-143 and HKII expression in NHL cell lines and/or patient-derived tumour samples with the CT or TT genotypes would be beneficial.

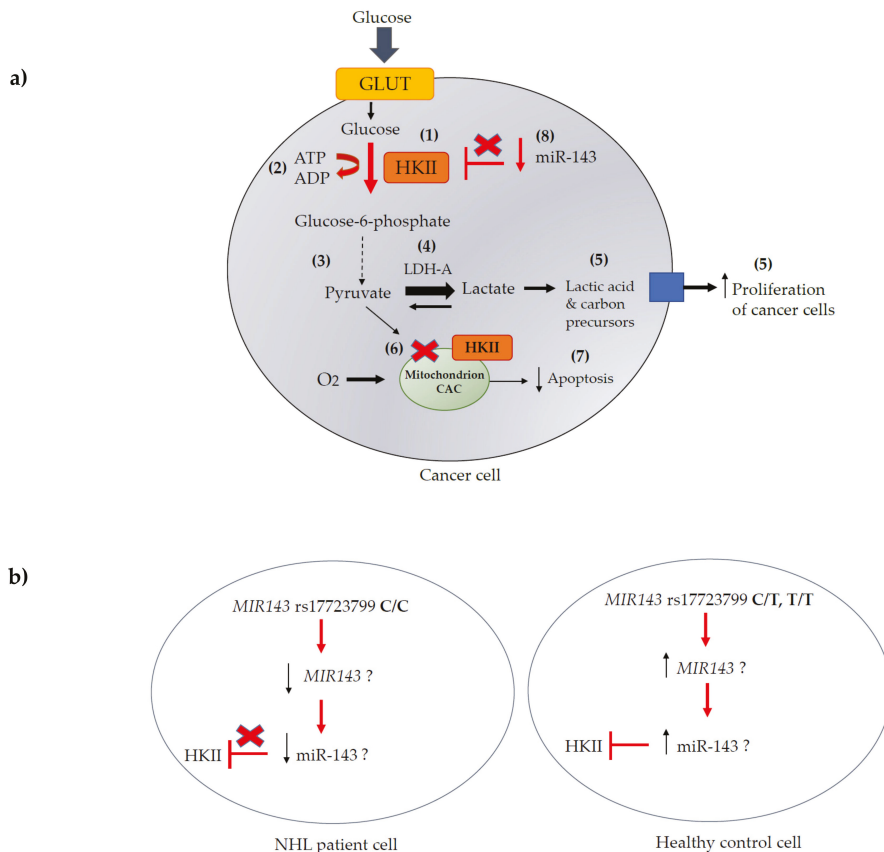


Figure 4. The role of HKII and miR-143 in cancer. (a) In a cancer cell, aerobic glycolysis occurs with an increased uptake of glucose and production of lactate even in the presence of oxygen (in normal cells glycolysis is anaerobic) and is known as the “Warburg effect”. Upregulation of Hexokinase 2 (*HKII*) in malignant tumours catalyses the conversion of glucose to glucose-6-phosphate (G-6-P) in the first step of the glycolysis pathway (1) with high consumption of ATP to ADP (2) to maintain the high rate of glycolysis required. G-6-P undergoes multiple conversions to pyruvate (3), which together with high lactate dehydrogenase A (LDH-A) activity in one direction (4) results in lactic acid production in the cytosol, incorporation of carbon precursors and increased cancer cell proliferation (5). Oxidation of pyruvate in the mitochondrion is reduced resulting in reduced citrate production for the citric acid cycle (CAC) (6). *HKII* bound to the mitochondrial membrane can also enhance cancer by increasing mitochondrial resistance to cell death signals helping to reduce apoptosis (7). Downregulation of miR-143 has been shown to directly reduce inhibition of *HKII* (8) causing increased glycolysis and tumour proliferation. (b) In a lymphoid cell, we propose that in the presence of the protective *MIR143* SNP rs17723799 genotype CT or TT, there may be increased *MIR143* expression and miR-143 transcription with increased targeting of *HKII* and reduced glycolysis via the Warburg effect causing decreased lymphomagenesis. This hypothesis is speculative, and further functional studies are required. ATP, adenosine triphosphate; ADP, adenosine diphosphate; CAC, citric acid cycle; GLUT, glucose transporter; LDH-A, lactate dehydrogenase A. (Adapted from Wikimedia [52] and Akins et al. (2018) [23]) [43].

5. Conclusions

This study is the first to report significant statistical correlation between SNPs in *MIR143* and reduced risk of NHL in Caucasians, and it is supported by the identification of significant SNPs in high LD in a large European NHL GWAS meta-analysis study. Our findings suggest that the three SNPs in LD in *MIR143* may be novel useful biomarkers to assess the risk of NHL in a clinical setting. Furthermore, this study gives some insight into *MIR143* transcriptional regulation by a protective promoter polymorphism in NHL. We do not suggest that there is a definitive functional association between *MIR143* SNPs and NHL onset, however further functional analyses are necessary to confirm this potential mechanism.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4425/10/3/185/s1>, Figure S1: Significant eQTLs for 3 analysed *MIR143* SNPs in 491 subjects. (a) rs17723799 (Homo Ref = 386, Het = 98, Homo Alt = 7), (b) rs3733846 (Homo Ref = 344, Het = 128, Homo Alt = 19) and (c) rs41291957 (Homo Ref = 369, Het = 114, Homo Alt = 8), showing increased *MIR143* host gene expression with heterozygous (Het) and homozygous alternate (Homo Alt) genotypes in skeletal muscle (GTEx Portal).

Author Contributions: The following authors contributed to this study: Conceptualisation, G.B., L.M.H. and L.R.G.; methodology, G.B., H.G.S. and L.M.H.; analysis, G.B. and E.M.A.; validation, R.A.L.; writing—original draft preparation, G.B.; writing—review and editing, G.B., L.M.H., H.G.S. and L.R.G.; supervision, H.G.S., L.M.H. and L.R.G.; funding acquisition, L.R.G.

Funding: Funding for this study has been generously provided by a Herbert family philanthropic donation, including PhD scholarship support for G.B. Funding has also been provided by the GRC Genomics Lymphoma Project (GLP) fund.

Acknowledgments: We would like to acknowledge Miles C. Benton for his technical support in the PLINK and Haploview analysis. We would also like to acknowledge the healthy control subjects for agreeing to donate their blood samples for the purposes of our research. We would like to thank the InterLymph Consortium for providing access to their GWAS summary statistics data for replication.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Skrabek, P.; Turner, D.; Seftel, M. Epidemiology of Non-Hodgkin Lymphoma. *Transfus. Apheresis Sci.* **2013**, *49*, 133–138. [[CrossRef](#)] [[PubMed](#)]
2. Cancer Facts & Figures 2018. Available online: <https://www.cancer.org/> (accessed on 15 November 2018).
3. Lymphoma Coalition Europe. Available online: <https://www.lymphomacoalition.org/europe> (accessed on 15 November 2018).
4. Australian Institute of Health and Welfare 2017. Cancer in Australia 2017 Cancer Series No. 101. Cat. No. CAN 100. Canberra: AIHW. Available online: www.aihw.gov.au (accessed on 23 September 2018).
5. Yadav, C.; Ahmad, A.; D'Souza, B.; Agarwal, A.; Nandini, M.; Ashok Prabhu, K.; D'Souza, V. Serum Lactate Dehydrogenase in Non-Hodgkin's Lymphoma: A Prognostic Indicator. *Indian J. Clin. Biochem.* **2016**, *31*, 240–242. [[CrossRef](#)] [[PubMed](#)]
6. Roehle, A.; Hoefig, K.P.; Repsilber, D.; Thorns, C.; Ziepert, M.; Wesche, K.O.; Thiere, M.; Loeffler, M.; Klapper, W.; Pfreundschuh, M.; et al. MicroRNA signatures characterize diffuse large B-cell lymphomas and follicular lymphomas. *Br. J. Haematol.* **2008**, *142*, 732–744. [[CrossRef](#)] [[PubMed](#)]
7. Lawrie, C.H.; Chi, J.; Taylor, S.; Tramonti, D.; Ballabio, E.; Palazzo, S.; Saunders, N.J.; Pezzella, F.; Boulwood, J.; Wainscoat, J.S.; et al. Expression of microRNAs in diffuse large B cell lymphoma is associated with immunophenotype, survival and transformation from follicular lymphoma. *J. Cell. Mol. Med.* **2003**, *13*, 1248–1260. [[CrossRef](#)] [[PubMed](#)]
8. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297. [[CrossRef](#)]
9. Shukla, G.C.; Singh, J.; Barik, S. MicroRNAs: Processing, Maturation, Target Recognition and Regulatory Functions. *Mol. Cell. Pharmacol.* **2011**, *3*, 83–92. [[PubMed](#)]
10. Salzman, D.W.; Weidhaas, J.B. SNPing cancer in the bud: MicroRNA and microRNA-target site polymorphisms as diagnostic and prognostic biomarkers in cancer. *Pharmacol. Ther.* **2013**, *137*, 55–63. [[CrossRef](#)] [[PubMed](#)]

11. Cullen, B.R. Transcription and Processing of Human microRNA Precursors. *Mol. Cell* **2004**, *16*, 861–865. [[CrossRef](#)] [[PubMed](#)]
12. Treiber, T.; Treiber, N.; Meister, G. Regulation of microRNA biogenesis and function. *Thromb. Haemost.* **2012**, *107*, 605. [[CrossRef](#)] [[PubMed](#)]
13. Vannini, I.; Fanini, F.; Fabbri, M. Emerging roles of microRNAs in cancer. *Curr. Opin. Genet. Dev.* **2018**, *48*, 128–133. [[CrossRef](#)] [[PubMed](#)]
14. Suresh, P.S.; Venkatesh, T.; Tsutsumi, R. In silico analysis of polymorphisms in microRNAs that target genes affecting aerobic glycolysis. *Ann. Transl. Med.* **2016**, *4*, 69. [[CrossRef](#)] [[PubMed](#)]
15. Yue, M.; Zhou, D.; Zhi, H.; Wang, P.; Zhang, Y.; Gao, Y.; Guo, M.; Li, X.; Wang, Y.; Zhang, Y.; et al. MSDD: A manually curated database of experimentally supported associations among miRNAs, SNPs and human diseases. *Nucleic Acids Res.* **2018**, *46*, D181–D185. [[CrossRef](#)] [[PubMed](#)]
16. Gong, J.; Tong, Y.; Zhang, H.-M.; Wang, K.; Hu, T.; Shan, G.; Sun, J.; Guo, A.-Y. Genome-wide identification of SNPs in microRNA genes and the SNP effects on microRNA target binding and biogenesis. *Hum. Mutat.* **2012**, *33*, 254–263. [[CrossRef](#)] [[PubMed](#)]
17. Zheng, B.; Xi, Z.; Liu, R.; Yin, W.; Sui, Z.; Ren, B.; Miller, H.; Gong, Q.; Liu, C. The Function of MicroRNAs in B-Cell Development, Lymphoma, and Their Potential in Clinical Practice. *Front. Immunol.* **2018**, *9*. [[CrossRef](#)] [[PubMed](#)]
18. Dzikiewicz-Krawczyk, A. MicroRNA polymorphisms as markers of risk, prognosis and treatment response in hematological malignancies. *Crit. Rev. Oncol. Hematol.* **2015**, *93*, 1–17. [[CrossRef](#)] [[PubMed](#)]
19. Bradshaw, G.; Sutherland, H.; Haupt, L.; Griffiths, L. Dysregulated MicroRNA Expression Profiles and Potential Cellular, Circulating and Polymorphic Biomarkers in Non-Hodgkin Lymphoma. *Genes* **2016**, *7*, 130. [[CrossRef](#)] [[PubMed](#)]
20. Vaitheesvaran, B.; Xu, J.; Yee, J.; Q.-Y., L.; Go, V.L.; Xiao, G.G.; Lee, W.N. The Warburg effect: A balance of flux analysis. *Metab. Off. J. Metab. Soc.* **2015**, *11*, 787–796. [[CrossRef](#)] [[PubMed](#)]
21. Wolf, A.; Agnihotri, S.; Micallef, J.; Mukherjee, J.; Sabha, N.; Cairns, R.; Hawkins, C.; Guha, A. Hexokinase 2 is a key mediator of aerobic glycolysis and promotes tumor growth in human glioblastoma multiforme. *J. Exp. Med.* **2011**, *208*, 313–326. [[CrossRef](#)] [[PubMed](#)]
22. Gu, J.J.; Singh, A.; Xue, K.; Mavis, C.; Barth, M.; Yanamadala, V.; Lenz, P.; Grau, M.; Lenz, G.; Czucman, M.S.; et al. Up-regulation of hexokinase II contributes to rituximab-chemotherapy resistance and is a clinically relevant target for therapeutic development. *Oncotarget* **2017**, 4020–4033. [[CrossRef](#)]
23. Akins, N.S.; Nielson, T.C.; Le, H.V. Inhibition of Glycolysis and Glutaminolysis: An Emerging Drug Discovery Approach to Combat Cancer. *Curr. Top. Med. Chem.* **2018**, *18*, 494–504. [[CrossRef](#)] [[PubMed](#)]
24. Palmieri, D.; Fitzgerald, D.; Shreeve, S.M.; Hua, E.; Bronder, J.L.; Weil, R.J.; Davis, S.; Stark, A.M.; Merino, M.J.; Kurek, R.; et al. Analyses of resected human brain metastases of breast cancer reveal the association between up-regulation of hexokinase 2 and poor prognosis. *Mol. Cancer Res.* **2009**, *7*, 1438–1445. [[CrossRef](#)] [[PubMed](#)]
25. Ros, S.; Schulze, A. Glycolysis Back in the Limelight: Systemic Targeting of HK2 Blocks Tumor Growth. *Cancer Discov.* **2013**, *3*, 1105. [[CrossRef](#)] [[PubMed](#)]
26. Mushtaq, M.; Darekar, S.; Klein, G.; Kashuba, E. Different Mechanisms of Regulation of the Warburg Effect in Lymphoblastoid and Burkitt Lymphoma Cells. *PLoS ONE* **2015**, *10*, e0136142. [[CrossRef](#)] [[PubMed](#)]
27. Bhalla, K.; Jaber, S.; Nahid, M.; Underwood, K.; Beheshti, A.; Landon, A.; Bhandary, B.; Bastian, P.; Evens, A.M.; Haley, J.; et al. Role of hypoxia in Diffuse Large B-cell Lymphoma: Metabolic repression and selective translation of HK2 facilitates development of DLBCL. *Sci. Rep.* **2018**, *8*, 744. [[CrossRef](#)] [[PubMed](#)]
28. Kwiatkowska, E.; Wojtala, M.; Gajewska, A.; Soszyński, M.; Bartosz, G.; Sadowska-Bartoszyńska, I. Effect of 3-bromopyruvate acid on the redox equilibrium in non-invasive MCF-7 and invasive MDA-MB-231 breast cancer cells. *J. Bioenergy Biomembr.* **2016**, *48*, 23–32. [[CrossRef](#)] [[PubMed](#)]
29. Chacon-Cortes, D.; Smith, R.A.; Haupt, L.M.; Lea, R.A.; Youl, P.H.; Griffiths, L.R. Genetic association analysis of miRNA SNPs implicates MIR145 in breast cancer susceptibility. *BMC Med. Genet.* **2015**, *16*, 107. [[CrossRef](#)] [[PubMed](#)]
30. Upadhyaya, A.; Smith, R.A.; Chacon-Cortes, D.; Revêchon, G.; Bellis, C.; Lea, R.A.; Haupt, L.M.; Chambers, S.K.; Youl, P.H.; Griffiths, L.R. Association of the microRNA-Single Nucleotide Polymorphism rs2910164 in miR146a with sporadic breast cancer susceptibility: A case control study. *Gene* **2016**, 576, 256–260. [[CrossRef](#)] [[PubMed](#)]

31. Chacon-Cortes, D.; Haupt, L.M.; Lea, R.A.; Griffiths, L.R. Comparison of genomic DNA extraction techniques from whole blood samples: A time, cost and quality evaluation study. *Mol. Biol. Rep.* **2012**, *39*, 5961–5966. [[CrossRef](#)] [[PubMed](#)]
32. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671–675. [[CrossRef](#)] [[PubMed](#)]
33. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)] [[PubMed](#)]
34. Team, R. *RStudio: Integrated Development for R*; RStudio, Inc.: Boston, MA, USA, 2015. Available online: <http://www.rstudio.com>.
35. Fry, B.; Maller, J.; Barrett, J.C.; Daly, M.J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **2004**, *21*, 263–265. [[CrossRef](#)]
36. Li, T.; Niu, L.; Wu, L.; Gao, X.; Li, M.; Liu, W.; Yang, L.; Liu, D. A functional polymorphism in microRNA-196a2 is associated with increased susceptibility to non-Hodgkin lymphoma. *Tumor Biol.* **2015**, *36*, 3279–3284. [[CrossRef](#)] [[PubMed](#)]
37. Cerhan, J.R.; Berndt, S.I.; Vijai, J.; Ghesquières, H.; McKay, J.; Wang, S.S.; Wang, Z.; Yeager, M.; Conde, L.; de Bakker, P.I.W.; et al. Genome-wide association study identifies multiple susceptibility loci for diffuse large B-cell lymphoma. *Nat. Genet.* **2014**, *46*, 1233–1238. [[CrossRef](#)] [[PubMed](#)]
38. Li, J.; Cheng, G.; Wang, S. A Single-Nucleotide Polymorphism of miR-196a2T>C rs11614913 Is Associated with Hepatocellular Carcinoma in the Chinese Population. *Genet. Test Mol. Biomark.* **2016**, *20*, 213–215. [[CrossRef](#)] [[PubMed](#)]
39. Peckham-Gregory, E.C.; Thapa, D.R.; Martinson, J.; Duggal, P.; Penugonda, S.; Bream, J.H.; Chang, P.-Y.; Dandekar, S.; Chang, S.-C.; Detels, R.; et al. MicroRNA-related polymorphisms and non-Hodgkin lymphoma susceptibility in the Multicenter AIDS Cohort Study. *Cancer Epidemiol.* **2016**, *45*, 47–57. [[CrossRef](#)] [[PubMed](#)]
40. Liu, Y.; He, A.; Liu, B.; Zhong, Y.; Liao, X.; Yang, J.; Chen, J.; Wu, J.; Mei, H. rs11614913 polymorphism in miRNA-196a2 and cancer risk: An updated meta-analysis. *Oncotargets Ther.* **2018**, *11*, 1121–1139. [[CrossRef](#)] [[PubMed](#)]
41. Li, L.; Pan, X.; Li, Z.; Bai, P.; Jin, H.; Wang, T.; Song, C.; Zhang, L.; Gao, L. Association between polymorphisms in the promoter region of miR-143/145 and risk of colorectal cancer. *Hum. Immunol.* **2013**, *74*, 993–997. [[CrossRef](#)] [[PubMed](#)]
42. Wu, J.; Huang, Q.; Meng, D.; Huang, M.; Li, C.; Qin, T. A Functional rs353293 Polymorphism in the Promoter of miR-143/145 Is Associated with a Reduced Risk of Bladder Cancer. *PLoS ONE* **2016**, *11*, e0159115. [[CrossRef](#)] [[PubMed](#)]
43. Fang, R.; Xiao, T.; Fang, Z.; Sun, Y.; Li, F.; Gao, Y.; Feng, Y.; Li, L.; Wang, Y.; Liu, X.; et al. MicroRNA-143 (miR-143) Regulates Cancer Glycolysis via Targeting Hexokinase 2 Gene. *J. Biol. Chem.* **2012**, *287*, 23227–23235. [[CrossRef](#)] [[PubMed](#)]
44. Peschiaroli, A.; Giacobbe, A.; Formosa, A.; Markert, E.K.; Bongiorno-Borbone, L.; Levine, A.J.; Candi, E.; D’Alessandro, A.; Zolla, L.; Finazzi Agrò, A.; et al. miR-143 regulates hexokinase 2 expression in cancer cells. *Oncogene* **2012**, *32*, 797. [[CrossRef](#)] [[PubMed](#)]
45. Jiang, S.; Zhang, L.F.; Zhang, H.W.; Hu, S.; Lu, M.H.; Liang, S.; Li, B.; Li, Y.; Li, D.; Wang, E.D.; et al. A novel miR-155/miR-143 cascade controls glycolysis by regulating hexokinase 2 in breast cancer cells. *Embo J.* **2012**, *31*, 1985–1998. [[CrossRef](#)] [[PubMed](#)]
46. Zeinali, T.; Mansoori, B.; Mohammadi, A.; Baradaran, B. Regulatory mechanisms of miR-145 expression and the importance of its function in cancer metastasis. *Biomed. Pharmacother.* **2019**, *109*, 195–207. [[CrossRef](#)] [[PubMed](#)]
47. Zeng, X.-L.; Zhang, S.-Y.; Zheng, J.-F.; Yuan, H.; Wang, Y. Altered miR-143 and miR-150 expressions in peripheral blood mononuclear cells for diagnosis of non-small cell lung cancer. *Chin. Med. J.* **2013**, *126*, 4510–4516. [[CrossRef](#)] [[PubMed](#)]
48. Yoshino, H.; Enokida, H.; Itesako, T.; Kojima, S.; Kinoshita, T.; Tatarano, S.; Chiyomaru, T.; Nakagawa, M.; Seki, N. Tumor-suppressive microRNA-143/145 cluster targets hexokinase-2 in renal cell carcinoma. *Cancer Sci.* **2013**, *104*, 1567–1574. [[CrossRef](#)] [[PubMed](#)]

49. Dos Santos Ferreira, A.C.; Robaina, M.C.; de Rezende, L.M.M.; Severino, P.; Klumb, C.E. Histone deacetylase inhibitor prevents cell growth in Burkitt's lymphoma by regulating PI3K/Akt pathways and leads to upregulation of miR-143, miR-145, and miR-101. *Ann. Hematol.* **2014**, *93*, 983–993. [[CrossRef](#)] [[PubMed](#)]
50. Shen, J.; Zhang, Y.; Fu, H.; Wu, D.; Zhou, H. Overexpression of microRNA-143 inhibits growth and induces apoptosis in human leukemia cells. *Oncol. Rep.* **2014**, *31*, 2035–2042. [[CrossRef](#)] [[PubMed](#)]
51. Carithers, L.J.; Ardlie, K.; Barcus, M.; Branton, P.A.; Britton, A.; Buia, S.A.; Compton, C.C.; DeLuca, D.S.; Peter-Demchok, J.; Gelfand, E.T.; et al. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv. Biobank.* **2015**, *13*, 311–319. [[CrossRef](#)] [[PubMed](#)]
52. Bcndoye. Comparison of LDH Activity in Normal and Cancerous Cell Metabolism. Available online: https://commons.wikimedia.org/wiki/File:LDH_activity-_normal_vs_canceous_cells.png (accessed on 9 November 2018).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Genes Editorial Office
E-mail: genes@mdpi.com
www.mdpi.com/journal/genes



MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34
Fax: +41 61 302 89 18

www.mdpi.com



ISBN 978-3-03921-715-1