MicroRNA and Cancer

Methods and Protocols

Edited by

Wei Wu

Institute for Biocomplexity and Informatics, Department of Biological Science, University of Calgary, Calgary, AB, Canada

Humana Press

Editor
Wei Wu
Institute for Biocomplexity and Informatics
Department of Biological Science
University of Calgary
Calgary, AB
Canada
wwu65@yahoo.com; wuwei@ucalgary.ca

ISSN 1064-3745 e-ISSN 1940-6029
ISBN 978-1-60761-862-1 e-ISBN 978-1-60761-863-8
DOI 10.1007/978-1-60761-863-8
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2010934277

© Springer Science+Business Media, LLC 2011
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Humana Press, c/o Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.
While the advice and information in this book are believed to be true and accurate at the date of going to press, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Humana Press is a part of Springer Science+Business Media (www.springer.com)

Preface

The discovery of microRNAs (miRNAs or miRs) heralded a new and an exciting era in biology and started a new chapter in human gene regulation. The miRNAs, a class of small endogenous noncoding RNAs (~22 nt), fine tune the gene expression at the posttranscriptional level through mainly binding 3′-UTR of mRNAs. They are involved in stem cell self-renewal, cellular development, differentiation, proliferation, and apoptosis.

Small miRNAs have big impacts in cancer development. Among the many miRNAs, a subset of miRNAs were identified as regulators of neoplastic transformation, tumor progression, invasion, and metastasis as well as tumor-initiating cells (cancer stem cells). The widespread deregulation of miRNomes in diverse cancers when compared to normal tissues have been unveiled. The oncomirs (oncogenic miRNAs), TSmiRs (tumor suppressive miRNAs), and MetastamiRs (miRNAs associated to cancer metastasis) comprise an important part of the cancer genome and confer pivotal diagnostic and prognostic significance. Moreover, cancer-associated miRNAs are proving worthwhile for developing effective cancer biomarkers for individualized medicine and potential therapeutic targets.

This book provides the latest and foremost miRNAs knowledge of cancer research applications from scientists all over the world. It is organized in two parts: the first part begins with a general introduction of miRNA biogenesis which is followed by chapters on the computational prediction of new microRNAs in cancer, the innovative technologies for modulating miRNAs of interests, and an overview of miRNA-based therapeutic approaches for cancer treatment. The second part of the book provides experimental and computational procedures in miRNA detection with diverse techniques, miRNA library construction, epigenetic regulation of miNRAs, microRNA::mRNA regulatory networks predicted with computational analysis in cancer cells or tissues, and the principle of designing the miRNA-mimics for miRNA activation. These chapters have been written for practical use in laboratories for graduate students, postdoctoral fellows, and scientists in cancer research. The contributors have shared their most valuable experiences in the translation of miRNA knowledge into cancer research.

MicroRNA is a fast growing field, and microRNA knowledge is a pivotal element of cancer biology. An individual miRNA interferes with a broad range of mRNAs; conversely, a single mRNA could be targeted by a variety of miRNAs. The complexity of miRNA::mRNA interaction is far-reaching and a bit beyond our understanding to date. This book is expected to provide the basic principles of experimental and computational methods for microRNA study in cancer research and provide a firm grounding for those who wish to develop further applications.

I am especially indebted to Prof. Shu Zheng and Dr. Suzanne D. Conzen for giving me the opportunity to gain substantial experience in cancer research. I would like to thank Dr. Gunglin Li for his kind support. Without their confidence and continuous support, many things would not have been possible. I also thank Prof. John Walker for his encouragement. Finally, I am grateful to all of the contributing authors for providing such high quality manuscripts. Additional thanks go to Zineb Elkadiri and the Humana Press for their excellent assistance.

Wei Wu
Calgary, AB, Canada

Contents

Preface . . . . . *v*
Contributors . . . . . *ix*

PART I REVIEWS

1 MicroRNA Biogenesis and Cancer 3
Julia Winter and Sven Diederichs

2 Computational Identification of miRNAs Involved in Cancer 23
Anastasis Oulas, Nestoras Karathanasis, and Panayiota Poirazi

3 The Principles of MiRNA-Masking Antisense Oligonucleotides Technology 43
Zhiguo Wang

4 The Concept of Multiple-Target Anti-miRNA Antisense Oligonucleotide Technology 51
Zhiguo Wang

5 Modulation of MicroRNAs for Potential Cancer Therapeutics 59
Wei Wu

PART II PROTOCOLS

6 Detection of MicroRNAs in Cultured Cells and Paraffin-Embedded Tissue Specimens by In Situ Hybridization 73
Ashim Gupta and Yin-Yuan Mo

7 MicroRNA Northern Blotting, Precursor Cloning, and Ago2-Improved RNA Interference 85
Julia Winter and Sven Diederichs

8 miRNA Profiling on High-Throughput OpenArray™ System 101
Elena V. Grigorenko, Elen Ortenberg, James Hurley, Andrew Bond, and Kevin Munnelly

9 Silicon Nanowire Biosensor for Ultrasensitive and Label-Free Direct Detection of miRNAs 111
Guo-Jun Zhang

10 High-Throughput and Reliable Protocols for Animal MicroRNA Library Cloning 123
Caide Xiao

11 MicroRNA Regulation of Growth Factor Receptor Signaling in Human Cancer Cells 147
Keith M. Giles, Andrew Barker, Priscilla M. Zhang, Michael R. Epis, and Peter J. Leedman

12 Epigenetic Regulation of MicroRNA Expression in Cancer 165
Hani Choudhry and James W.F. Catto

13 In Vitro Functional Study of miR-126 in Leukemia 185
Zejuan Li and Jianjun Chen

14 Prediction of the Biological Effect of Polymorphisms Within MicroRNA Binding Sites 197
Debora Landi, Roberto Barale, Federica Gemignani, and Stefano Landi

15 The Guideline of the Design and Validation of MiRNA Mimics 211
Zhiguo Wang

16 Analysis of Targets and Functions Coregulated by MicroRNAs 225
Shu-Jen Chen and Hua-Chien Chen

17 Utilization of SSCprofiler to Predict a New miRNA Gene 243
Anastasis Oulas and Panayiota Poirazi

18 MicroRNA Profiling Using Fluorescence-Labeled Beads: Data Acquisition and Processing 253
Thomas Streichert, Benjamin Otto, and Ulrich Lehmann

Index 269

Contributors

Roberto Barale • *Genetics, Department of Biology, University of Pisa, Pisa, Italy*

Andrew Barker • *Laboratory for Cancer Medicine, Centre for Medical Research, Western Australian Institute for Medical Research, University of Western Australia, Perth, WA, Australia*

Andrew Bond • *BioTrove Inc, Woburn, MA, USA*

James W. F. Catto • *Academic Urology Unit, Division of Clinical Sciences, Royal Hallamshire Hospital, University of Sheffield, Sheffield, UK*

Hua-Chien Chen • *Department of Life Science, Molecular Medicine Research Center, Change Gung University, Taiwan, China*

Jianjun Chen • *Department of Medicine, The University of Chicago, Chicago, IL, USA*

Shu-Jen Chen • *Department of Life Science, Molecular Medicine Research Center, Change Gung University, Taiwan, China*

Hani Choudhry • *Faculty of Science, Academic Urology Unit, Division of Clinical Sciences, Royal Hallamshire Hospital, University of Sheffield, Sheffield, UK; Biochemistry Department, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia*

Sven Diederichs • *Helmholtz-University-Group "Molecular RNA Biology & Cancer," German Cancer Research Center (DKFZ), Institute of Pathology, University of Heidelberg, Heidelberg, Germany*

Michael R. Epis • *Laboratory for Cancer Medicine, Centre for Medical Research, Western Australian Institute for Medical Research, University of Western Australia, Perth, WA, Australia*

Federica Gemignani • *Genetics, Department of Biology, University of Pisa, Pisa, Italy*

Keith M. Giles • *Laboratory for Cancer Medicine, Centre for Medical Research, Western Australian Institute for Medical Research, University of Western Australia, Perth, WA, Australia*

Elena V. Grigorenko • *BioTrove Inc, Woburn, MA, USA*

Ashim Gupta • *Immunology and Cell Biology, Department of Medical Microbiology, School of Medicine, Southern Illinois University, Springfield, IL, USA*

James Hurley • *BioTrove Inc, Woburn, MA, USA*

Nestoras Karathanasis • *Institute for Molecular Biology and Biotechnology (IMBB), Foundation for Research and Technology-Hellas (FORTH), Heraklion, Crete, Greece*

Debora Landi • *Genetics, Department of Biology, University of Pisa, Pisa, Italy*

Stefano Landi • *Genetics, Department of Biology, University of Pisa, Pisa, Italy*

Peter J. Leedman • *Laboratory for Cancer Medicine, Centre for Medical Research, School of Medicine and Pharmacologcxxy, Western Australian Institute for Medical Research, University of Western Australia, Perth, WA, Australia*

Ulrich Lehmann • *Institute of Pathology Medizinische Hochschule Hannover, Hannover, Germany*
Zejuan Li • *Section of Hematology and Oncology, Department of Medicine, The University of Chicago, Chicago, IL, USA*
Yin-Yuan Mo • *Immunology and Cell Biology, Department of Medical Microbiology, School of Medicine, Southern Illinois University, Springfield, IL, USA*
Kevin Munnelly • *BioTrove Inc, Woburn, MA, USA*
Elen Ortenberg • *BioTrove Inc, Woburn, MA, USA*
Benjamin Otto • *Department of Clinical Chemistry/Central Laboratories, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany*
Anastasis Oulas • *Institute for Molecular Biology and Biotechnology (IMBB), Foundation for Research and Technology-Hellas (FORTH), Heraklion, Crete, Greece*
Panayiota Poirazi • *Computational Biology Laboratory, Institute for Molecular Biology and Biotechnology (IMBB), Foundation for Research and Technology-Hellas (FORTH), Heraklion, Crete, Greece*
Thomas Streichert • *Department of Clinical Chemistry/Central Laboratories, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany*
Zhiguo Wang • *Department of Medicine, Montreal Heart Institute, University of Montreal, Montreal, QC, Canada*
Julia Winter • *Helmholtz-University-Group "Molecular RNA Biology & Cancer," German Cancer Research Center (DKFZ) & Institute of Pathology, University of Heidelberg, Heidelberg, Germany*
Wei Wu • *Institute for Biocomplexity and Informatics, Department of Biological Science, University of Calgary, Calgary, AB, Canada*
Caide Xiao • *Institute for Biocomplexity and Informatics, Department of Biological Science, University of Calgary, Calgary, AB, Canada*
Guo-Jun Zhang • *Institute of Microelectronics, A*STAR (Agency for Science, Technology and Research), Singapore, Singapore*
Priscilla M. Zhang • *Laboratory for Cancer Medicine, Centre for Medical Research, Western Australian Institute for Medical Research, University of Western Australia, Perth, WA, Australia*

Chapter 1

MicroRNA Biogenesis and Cancer

Julia Winter and Sven Diederichs

Abstract

MicroRNAs are short non-coding RNA molecules that are involved in diverse physiological and developmental processes by controlling the gene expression of target mRNAs. They play important roles in almost all kinds of cancer where they modulate key processes during tumorigenesis such as metastasis, apoptosis, proliferation, or angiogenesis. Depending on the mRNA targets they regulate, they can act as oncogenes or as tumor suppressor genes. Multiple links between microRNA biogenesis and cancer highlight its significance for tumor diseases. However, mechanisms of their own regulation on the transcriptional and posttranscriptional level in health and disease are only beginning to emerge. Here, we review the microRNA-processing pathway as well as recent insights into posttranscriptional regulation of microRNA expression.

Key words: microRNA, miRNA, Biogenesis, Processing, Cancer, Tumor, Posttranscriptional regulation

1. From an Oddity of Nematodes to the Human MicroRNA World

The first microRNA (miRNA) was discovered in 1993 in the nematode *Caenorhabditis elegans*. This noncoding RNA termed lin-4 modulates the expression of lin-14, a protein-coding gene that is relevant for developmental timing in *C. elegans* (1–3). The authors already proposed a regulatory mechanism through binding of the miRNA to the 3′ untranslated region (3′-UTR) of the target mRNA (1, 3).

Also the next miRNA identified, let-7, was a 22 nucleotide (nt) regulatory RNA important in the larval development of *C. elegans* (4, 5). Shortly after its discovery in *C. elegans*, let-7 had been identified in humans, *Drosophila melanogaster*, and several other bilaterian animals, and many other miRNA genes have been discovered in a wide range of species (6–10).

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_1, © Springer Science+Business Media, LLC 2011

2. Never Change a Winning Team: miRNAs are Highly Conserved Molecules

Mammalian miRNAs are short (20–23 nt), highly conserved molecules (11, 12) that seem to have evolved independently of plant miRNAs as their sequences and biogenesis pathways differ substantially (13, 14).

They are localized throughout the whole genome: in intronic or exonic regions of protein-coding or non-coding genes. Approximately 50% of all human miRNAs are interconnected in genomic clusters and transcribed from a single polycistronic transcription unit (15–17). Many miRNAs have multiple paralogues that could be the result of gene duplications. The sequence as well as the genomic loci of the pri-miRNA varies while identical mature miRNAs are generated and consequently identical mRNAs are targeted by these isoforms. To date, more than 700 miRNAs are registered for the human species alone and the number is still increasing (18).

3. Perfect Couples Not Required: Imperfect Complementarity Increases the Number of Potential Targets

MiRNAs bind to their targets through imperfect base pairing. Perfectly matched sequence complementarity is required only between the "seed" region of the miRNA (=nt 2–7 of the mature sequence) and the target mRNA (19). Such binding leads to either degradation, destabilization, or translational inhibition of the mRNA and consequently silences gene expression (20). As these binding sites are primarily located in the 3′ UTR of the target mRNA, some genes avoid miRNA regulation with the help of very short 3′ UTRs that are specifically depleted of these sites (21). However, miRNA-binding sites can also be located in the 5′ UTR or the coding region of a gene (22, 23). As binding between miRNA and target mRNA does not require perfect complementarity, a single miRNA can affect a broad range of mRNAs and consequently the whole miRNA family possesses the potential to target and regulate thousands of genes (24, 25). The estimation that approximately one third of the protein-coding genes are controlled by miRNAs indicates that almost all cellular pathways are directly or indirectly influenced by miRNAs.

Because imperfect base pairing between the miRNA and its target as well as alternative 3′ UTRs have to be taken into consideration, the identification of miRNA targets is still very challenging (26–29). Recently, a new in vivo method has been published that possesses the potential to substitute the algorithm-based bioinformatic target search. Argonaute high-throughput sequencing of RNAs isolated by crosslinking immunoprecipitation (HITS-CLIP)

directly identifies protein–RNA interactions in living tissues by covalently crosslinking the complexes and subsequently sequencing the Ago-bound RNA molecules (30–32).

4. A Multi-talented Family: miRNAs and Cancer: Tumor Suppressors and Oncogenes from One Tribe

Aberrant miRNA expression and function is frequently observed in many diseases including cancer (33–35). MiRNAs exhibit tumor suppressing (e.g., miR-15a/16-1 (36) and let-7 (37, 38)) or oncogenic (e.g., miR-17-92 (39)) activities depending on their respective target genes (40). As miRNA expression patterns strongly correlate with tumor type and stage (41, 42), miRNAs can be used as clinical markers for cancer diagnosis and prognosis (43, 44). Genome-wide miRNA profiling uncovered a general downregulation of miRNAs in cancerous tissues, but next to this general trend, individual miRNAs can be upregulated (35). At least three different mechanisms have been proposed to be responsible for altered miRNA expression in tumors: localization of miRNAs inside or close to cancer-associated genomic regions (42, 45, 46), epigenetic regulation of miRNA expression (47–50), or abnormalities in miRNA processing (51–53).

In cancer, multiple lines of evidence suggest an important role of posttranscriptional regulation of miRNA expression during the miRNA-processing pathway (54). First, the global downregulation of miRNA expression in tumors points toward processing defects in cancer affecting the general expression level of mature miRNAs (35). In addition, individual miRNAs exhibit defects in processing with normal pre-miRNA levels coinciding with reduced mature miRNA levels (35, 42, 52, 55–57) as shown for miR-143 and miR-145 in colorectal adenocarcinomas (53) or miR-7 in glioblastomas (51). Owing to retained pre-miR-31 in the nucleus, mature miR-31 is absent in several cell-lines, for example, the breast cancer cell-line MCF-7, despite substantial pri- and precursor levels (52).

Hence, in this introductory chapter, we focus on the miRNA biogenesis pathway (Fig. 1), regulatory mechanisms therein, and their aberrations in cancer. In addition, the miRNA pathway is also of greatest importance for future drug developments because a functional miRNA machinery is a compulsory prerequisite for any RNA interference (RNAi)-based therapy approach. The siRNA-based drug only provides the specificity component that targets, for example, a driving oncogene in a tumor cell. However, the miRNA machinery must be functional to mediate the siRNA-induced knockdown. Thus, a detailed understanding of the mechanisms and players in miRNA biogenesis and function in physiological settings as well as malignant diseases is essential for the successful development of RNAi-based drugs.

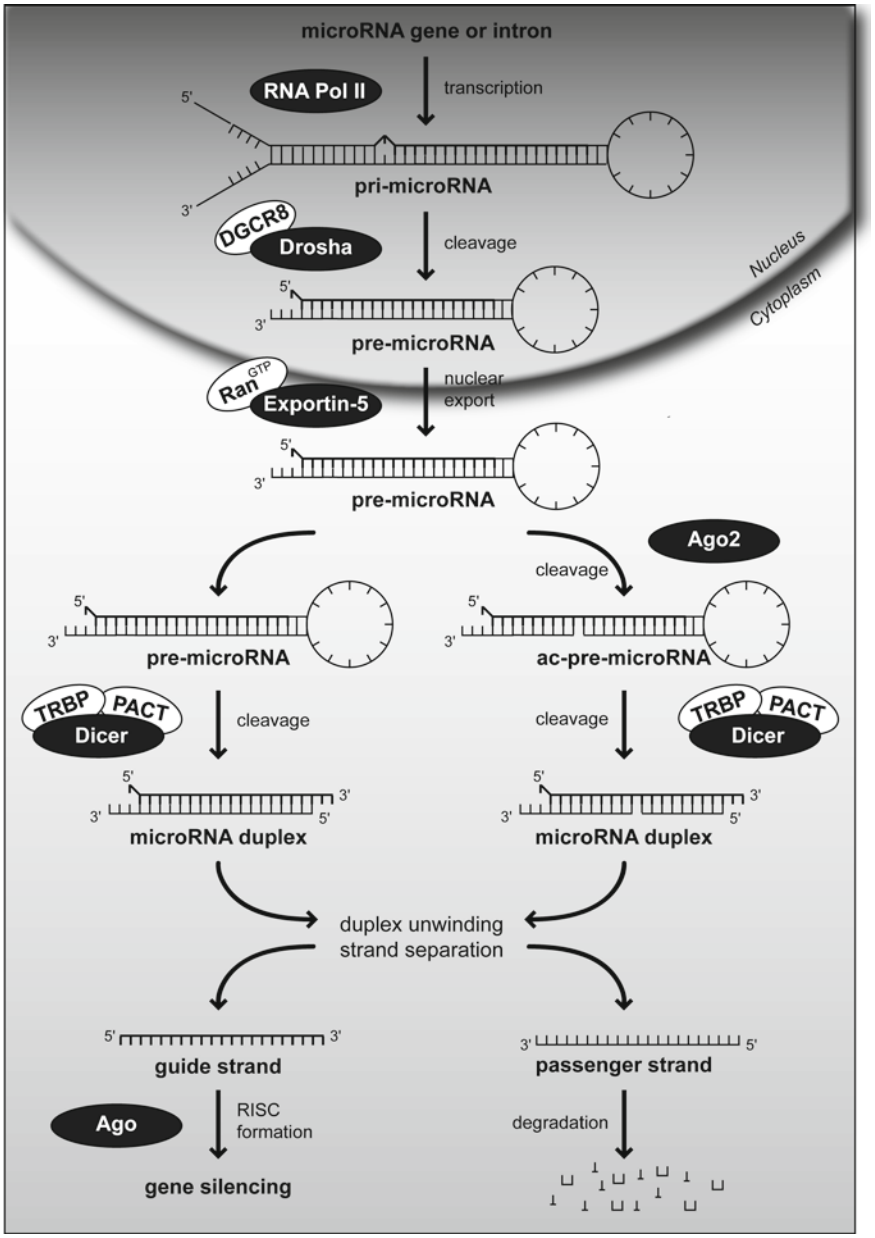


Fig. 1. The canonical miRNA biogenesis pathway. MiRNAs are mainly transcribed by Polymerase II before Drosha cleaves off the ss flanking region to generate the pre-miRNA that is exported to the cytoplasm in a Ran-GTP-dependent manner by Exportin-5. Several miRNAs are cleaved by Ago2, generating the ac-pre-miRNA. Subsequent to Dicer cleavage, the miRNA-duplex is unwound and the passenger strand degraded. The guide strand is then incorporated into the RISC where gene silencing can be accomplished via mRNA target cleavage, translational repression, or mRNA deadenylation.

5. A Stitch in Time… miRNA Regulation Begins at the Transcriptional Level

The majority of miRNAs is polyadenylated and capped, indicators of Polymerase II (Pol II)-mediated transcription (58, 59). In addition, an association with Pol III instead of Pol II was reported for the promoter of the miR-517-cluster (C19MC) (60). But as more recent data proposed a nonprotein-coding Pol II to be responsible for the transcription of this cluster (61), evidence is not clear whether other polymerases than Pol II are involved in miRNA transcription.

MiRNAs are under the control of a wide range of transcription factors, including tumorigenic regulators. Both tumor suppressors and oncogenes can influence miRNA expression: While c-Myc binds to the miR-17-92-locus and activates the expression of this cluster (62), miR-34 is a direct target of p53 (63). However, also the methylation status of promoter sequences significantly contributes to transcriptional miRNA regulation: inhibiting DNA methylation and histon deacetylases upregulates the expression of several miRNAs, especially miR-127 (48, 49). Let-7a-3, a gene that is located in CpG islands and heavily methylated in healthy human tissue, is hypomethylated in some lung adenocarcinomas (64).

6. The First Cut is the Deepest: Drosha Generates the Pre-miRNA

After transcription, pri-miRNAs, which are usually composed of a 33 base pair (bp) hairpin stem, a terminal loop, and flanking single-stranded (ss) DNA regions of several kilobases (kb) length, are processed by Drosha (RNASEN). This cleavage step occurs 11 bp apart from the stem of the hairpin and therefore already defines the 5′ end of the mature miRNA (65). Drosha generates a 2 nt overhang at its 3′ end, a characteristic feature of RNase III endonucleases (66, 67). This can occur co-transcriptionally and therefore prior to splicing (68, 69).

Drosha is part of two different complexes: It is either affiliated with several RNA-associated proteins including RNA helicases and nuclear riboproteins or a component of the so-called microprocessor complex that predominantly facilitates pri-miRNA cleavage (70). The microprocessor complex is composed of Drosha and its directly interacting dsRNA-binding protein (dsRBP) DGCR8 (Di George syndrome Critical Region 8, called Pasha in *D. melanogaster* or *C. elegans*) (70–74) that regulate each other via feedback mechanisms. While DGCR8 stabilizes Drosha through an interaction with its C-terminal domain, the endonuclease cleaves two hairpin structures in the 5′ UTR and coding region of *Dgcr8* mRNA subsequently resulting in its degradation (75).

Even though Drosha contains a dsRNA-binding domain, pri-miRNA association depends on DGCR8. DGCR8 downregulation leads to accumulating pri-miRNA levels and reduced levels of mature miRNAs in several organisms (70–73).

DGCR8 possesses a second feature: It acts as a molecular ruler and determines the Drosha cleavage site that is located ~11 bp apart from the ssRNA–dsRNA junction (SD junction). Therefore, DGCR8 recognizes the SD junction prior to transiently interacting with the pri-miRNA to facilitate binding between the processing center and the cleavage site (72). The resulting pre-miRNA is now composed of a ~22 nt stem and a terminal loop. The structure of the single-stranded flanking regions is critical for efficient Drosha cleavage – high secondary structures and blunt ends disrupt processing while neither the stem loop structure nor the sequence of the miRNA is essential (76, 77).

Notable exceptions are the so-called "mirtrons" – pre-miRNAs that correspond precisely to spliced-out introns and therefore circumvent the first parts of the miRNA biogenesis pathway (78–80). Even though Drosha-mediated processing is dispensable for mirtrons, the biogenesis pathway in the cytoplasm is identical to other miRNAs processed from pri-miRNAs by Drosha.

7. Changing the Settings: Exportin-5 Mediates Nuclear Export into the Cytoplasm

Exportin-5, a Ran-GTP-dependent dsRNA-binding protein, transports various tRNAs from the nucleus into the cytoplasm and also mediates the export of pre-miRNAs (81–83). The specificity of this process depends on structural determinants on the miRNA, including the 3′ overhangs and a defined length of the stem. Consequently, Exportin-5 does not only export pre-miRNAs but is also capable of protecting them against nuclear degradation (83, 84). As this process is saturable and thus carrier-mediated, shRNA overexpression and efficiency of RNA experiments are enhanced by Exportin-5 cotransfection (85).

8. Getting Organized: Dicer, TRBP/PACT, and Ago2 form the RISC-Loading Complex

Prior to Dicer cleavage, the RISC-loading complex (RLC) is generated to provide a platform for RISC (RNA-induced silencing complex) assembly and cytoplasmic processing. Dicer and its dsRBD proteins, TRBP (Tar RNA-binding protein), and PACT (protein activator of PKR) associate to a complex that is

subsequently replenished by the core component Ago2 (Argonaute-2). The exported pre-miRNA only accedes after the formation of this ternary complex (86–90).

Although neither TRBP nor PACT is required for processing activity itself and the in vitro reconstitution of the RLC is accomplished by Dicer, TRBP, and Ago2 alone, both dsRBD proteins seem to enhance RISC formation because depletion of TRBP or PACT reduces the efficiency of posttranscriptional gene silencing and both proteins have been shown to participate in the recruitment of Ago2 (88–90).

9. Generating an Additional Precursor: Ago2 Cleaves the Pre-miRNA

For some highly complementary miRNAs, Ago2 cleaves the 3′ arm in the middle of the hairpin to generate a nicked hairpin that is processed as efficiently as noncleaved miRNAs, the ac-pre-miRNA (91). This additional cleavage step might explain the early association of Ago2 with the RLC, even before the pre-miRNA (86–90). Even though the function and the molecular determinants for this cleavage reaction remain to be elucidated in detail, a role in strand separation can be proposed as shown for siRNAs. Therefore, passenger strand cleavage affects thermodynamic stability and subsequently facilitates strand separation of siRNAs (92–95).

10. The Final Cut: Dicer Cleaves off the Loop

Once the RLC is formed, the endonuclease Dicer cleaves off the loop of the pre-miRNA to generate the miRNA duplex (96). As Drosha already determined one end of the mature miRNA, Dicer cleavage defines the other mature miRNA end with 2 nt protruding overhangs (97–101).

In contrast to Drosha, Dicer does not require a “molecular ruler” protein. Structural determinants of the pre-miRNA hairpin are sufficient to predict the cleavage site (72). The Dicer endonuclease is composed of a PAZ domain located in the middle region of the protein that is connected to two RNase III domains (RIIID) via a long positively charged helix (65, 102). The PAZ domain binds to the protruding ends while the helix spans the stem to direct the catalytic centers to the predicted cleavage sites (103–106). The two RNase III domains act as an intramolecular dimer in which the two catalytic sites are located closely to each other. RIIIDa cleaves the 3′ strand of the duplex, whereas RIIIDb cuts the complementary strand generating the characteristic 2 nt

overhang (65, 102). The distance between the PAZ domain and the catalytic center was calculated as 65 Å which is consistent with the length of the cleaved product (65, 102).

Dicer is an evolutionary highly conserved protein that is indispensable in all organisms tested so far (98–100). Its abrogation leads to decreased or absent levels of mature miRNAs and even to lethality in early developmental stages of mice, accentuating the imperative of miRNA biogenesis (107). Even though the human genome contains only one Dicer gene, the number varies among species. *Drosophila melanogaster* contains two Dicer homologues that associate with different dsRBDs and maintain distinct functions. Dicer-1, associated with Loquacious-1 (Loqs-1), is essential for miRNA biogenesis, whereas Dicer-2 together with R2D2 plays an important role in siRNA production (108–110).

11. Unwinding the Duplex: Separating the Passenger and Guide Strand

Following Dicer cleavage, the ternary complex dissociates and the miRNA duplex is separated into the functional guide strand and the subsequently degraded passenger strand (87).

Little is known about the process of strand separation. Even though an association between several helicases (e.g., p68, p72, Gemin3/4, Mov10) and RISC formation has been reported, a general helicase responsible for duplex unwinding has not been identified so far (110–114). As RISC loading is an ATP-independent process (93), helicases might be dispensable but helpful tools for the unwinding of several miRNAs (115). P68 is sufficient to separate the let-7-duplex as shown via recombinant protein and knockdown experiments but has no effect on siRNA unwinding (110–114). As mentioned above, Ago2-cleavage in the middle of the passenger strand might also facilitate strand separation (91–95).

12. Passenger or Guide: That is the Question

Both strands of the miRNA duplex possess the potential to be either incorporated into the RISC or to be degraded. Consequently, one miRNA duplex is able to give rise to two individual mature miRNAs with varying targets due to differing seed regions (19). In mammals, thermodynamic stability of the two terminal base pairs provides the decisive criterion for strand selection. The strand with the more stable base pair at the 5′ end is typically degraded, whereas the one with the less stable terminal base pair is incorporated into the RISC (116, 117).

Owing to the existence of multiple Ago proteins, the guide strands are arranged into various RISCs. While sorting in *A. thaliana* depends on the 5′ terminal nucleotide (118), *D. melanogaster* takes the number of bulges and mismatches of the duplex as determinants for strand selection (108, 119). Despite distinct biogenesis pathways, miRNAs and siRNAs participate in a common sorting step where siRNA-like duplexes are incorporated into an Ago2-RISC by R2D2 while less complementary duplexes are incorporated into a RISC composed of Ago1 (108, 119). A sorting step in humans has not been identified so far. Even though antibody experiments accentuated the increased association between miRNAs and Ago2 or Ago3 (120), all four Ago proteins are capable of binding miRNAs sequence independently (121).

13. Sssshhh… Silencing the Target Gene

At least three independent mechanisms of miRNA-mediated mRNA silencing have been discovered so far: target cleavage, translational inhibition, and mRNA decay (20, 122). Because exclusively human Ago2 possesses slicing activity, different RISCs might fulfill different gene-silencing strategies (121, 123). In situations of perfect complementarity, the Ago2-RISC cleaves the mRNA (124). The decision whether RISCs containing the mature miRNA in combination with Ago1, Ago3, or Ago4 undergo translational inhibition or decay depends on specific features of the mRNA rather than the Ago protein or the miRNA, which was shown to be able to act via both pathways (125). The number and localization of miRNA-binding sites, the RNA context, and the structure of the miRNA-mRNA-duplex including number, type, and position of mismatches seem to be pivotal (26).

Translational inhibition is achieved via various mechanisms: Ago proteins compete with eIF6 or eIF4E to inhibit the association of the CAP structure with the elongation factor or the joining of the small and large subunits of the ribosome while configuration of the closed loop mRNA is prevented by an ill-defined mechanism that includes deadenylation (126, 127). However, translation-elongation blockage, ribosome drop-off promoting premature dissociation of the mRNA, or potentially cotranslational protein degradation by so far unknown proteases inhibit translation as well (128, 129).

In addition, miRNAs trigger deadenylation and subsequent decapping to initiate degradation of the mRNA target (130–133). The accumulation of miRNAs and Ago proteins in P-bodies, sites of enriched untranslated mRNAs and mRNA turnover enzymes, might be a consequence rather than a cause of target mRNA silencing (125, 134).

14. Recycling of miRNAs

In general, mature miRNAs might be rather stable. More than 48 h after the depletion of miRNA-processing factors, mature miRNAs are still detected (70, 73, 135) and binding to Ago proteins even extends the life-span of miRNAs (136). Because certain miRNAs (e.g., miR-29b) are degraded much more rapidly than others (136) and degradation occurs at the precursor level as well, this might be an additional regulation step of the processing pathway. In contrast to *C. elegans* and *A. thaliana*, nucleases promoting the degradation of human miRNAs remain to be elucidated (137, 138).

15. Regulating the Regulators: Aberrant Expression of Processing Factors in Tumors

Globally decreased miRNA abundance coinciding with constant pre-miRNA levels in various cancerous tissues is partially due to aberrant expression of processing components (42, 52). Inhibiting the biogenesis pathway via shRNAs against Drosha, DGCR8, or Dicer enhances tumorigenesis and transformation properties in vivo and in vitro, including colony formation and growth in soft agar. Injecting these processing-defective cells into nude mice increases the invasiveness of the tumor as well as protein levels of the oncogenes *k-ras* and *c-myc* (139).

Aberrant expression levels of Drosha and Dicer have been reported in a number of cancers (139–143). Dicer activity can also be reduced by frameshift mutations in the *TRBP*-gene, prohibiting TRBP binding to the amino-terminal DExD/H-box-helicase domain (144). This interaction generally leads to conformational changes and subsequent activation of Dicer (145). In addition, *Dicer* mRNA is controlled by let-7 targeting its 3′ UTR creating a negative feedback loop with its product (146).

Increased mature miRNA levels are detected after Ago1–4 overexpression, most probably due to the stabilizing activity of this protein family on miRNAs (91). Three human Ago proteins (Ago1, Ago3, and Ago4) are frequently deleted in Wilms tumors while Ago1 was reported to be overexpressed in renal tumors that lack the Wilms tumor suppressor gene *WT1* (147, 148). Generally, Ago proteins are controlled by posttranslational modifications: Hydroxylation stabilizes and phosphorylation guides Ago2 to P-bodies (149–151).

Ectopic Ago2 expression increases gene silencing of siRNA duplexes or shRNA constructs that possess perfect complementarity to their targets. Depending on the endonuclease activity of Ago2, this observation can be used to improve siRNA experiments (91, 152).

16. Everybody Is Unique: Regulation Is Not Universal to All miRNAs

For some time, miRNA biogenesis was thought to be controlled mainly at the transcriptional level. As mentioned above, different activators or repressors as well as the methylation status of distinct promoters are able to fulfill this regulation (48, 49, 64, 153). However, not only the observation that each miRNA located in the same cluster can be transcribed and regulated independently but also the restriction of individual posttranscriptional regulation steps to single miRNAs led to the conclusion that processing is neither universal to all miRNAs nor restricted to the regulation of processing factors (54).

16.1. The Regulatory Roles of hnRNPA1, SMAD Proteins, and Lin-28

MiRNA-specific posttranscriptional regulation has been described predominantly at the level of the microprocessor complex.

The heterogeneous ribonucleoprotein A1 (hnRNP A1), an RNA-binding protein implicated in various pathways of RNA processing, exclusively binds the terminal loop of the potential oncogene pri-miR-18a but not to any of the other six neighboring miRNAs that are located in the miR-17-92-cluster. This association increases Drosha processing mediated by structural modifications (154). In the light of hnRNPA1's ability to bind to the terminal loops of let-7a-1 and miR-101, as well, and the finding that ~14% of all miRNAs contain highly conserved terminal loops and could thus be subject to similar regulatory mechanisms, it needs to be determined whether this mechanism is restricted to pri-miR-18a (154, 155).

Posttranscriptionally enhanced Drosha processing is also detected for miR-21, a miRNA that is upregulated in almost all types of cancer, mediating invasion and metastasis (156). In vascular smooth muscle cells, TGF-β and BMP signaling recruit SMAD proteins together with the helicase p68 to pri-miR-21 resulting in increased Drosha cleavage and consequently upregulated levels of mature miR-21 (157). P68 has also been shown to facilitate pri- to pre-miRNA processing of several miRNAs after p53 interaction (158).

Lin-28 can affect let-7 processing at both levels, processing by Drosha or Dicer. As this pluripotent RNA-binding protein is exclusively highly expressed in undifferentiated and cancer cells, it was considered as a potential target for cancer treatment (159). Binding to conserved nucleotides in the loop region selectively blocks Drosha cleavage (160–162) while the activation of TUT4 (terminal uridylytransferase 4) induces uridylation of pre-let-7 at its 3′ end. This subsequently leads to pre-miRNA degradation and inhibition of Dicer cleavage (163–165).

16.2. Single Nucleotide Polymorphisms and RNA-Editing

Posttranscriptional regulation also relies on sequence variations in the small regulatory RNAs, for example, RNA-editing events or single nucleotide polymorphisms (SNPs). Despite the dispensability of specific sequences for Drosha cleavage (45, 84), a G-U

SNP in miR-125a blocks processing and consequently reduces the levels of mature miR-125a (166).

Adenosine-to-Inosine editing either takes place in the nucleus or in the cytoplasm. Enzymes that mediate deamination (ADARs) modify miRNA sequences and subsequently influence processing by Drosha (167, 168), Dicer (169, 170), or even change the targets of the miRNA (171). Editing of pri-miR-142 prevents Drosha cleavage resulting in subsequent degradation of the miRNA by the ribonuclease Tudor-SN (167, 168), whereas the editing event in pri-miR-151 inhibits Dicer cleavage (169, 170). However, this posttranscriptional regulation can also increase processing and consequently the number of mature miRNAs (169).

17. Conclusions and Outlook: miRNA Biogenesis and Cancer

In summary, deregulated miRNA processing occurs in tumorous cells and might be one reason for aberrant expression patterns frequently observed in cancer with a notable global downregulation of mature miRNAs (35, 44). Reduced expression of Dicer associates with poor prognosis in lung cancer (142), and the knockdown of *Drosha*, *Dgcr8*, or *Dicer* enhances tumorigenesis and transformation properties in vivo and in vitro (139). MiRNA biogenesis is not universal to all miRNAs but can even be specific for a single miRNA that is posttranscriptionally regulated by editing or miRNA-specific processing factors such as hnRNP A1, SMAD proteins, or lin-28 (54, 172). Therefore, unraveling the mechanisms underlying miRNA regulation is a central challenge in the coming years to gain knowledge of their many roles in health and disease. In addition, the miRNA pathway is also of great importance to future RNAi-based drug developments because these therapy approaches require an efficient and properly working miRNA machinery. Hence, gaining deeper insights in the mechanisms and modulators of miRNA biogenesis and function is also essential for the successful development and application of RNAi-based drugs.

Acknowledgments

We apologize to the many researchers whose important work in the microRNA field could not be cited due to space constraints. We are grateful for the support of our research by the Helmholtz Society (VH-NG-504), the Marie Curie Programme of the European Union (239308), the German Research Foundation

(TRR77 TP B03), the German Cancer Research Center, and the Institute of Pathology at the University of Heidelberg. We would also like to thank all members of our lab for helpful discussions on the manuscript.

References

1. Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14, *Cell 75*, 843–854.
2. Wightman, B., Burglin, T. R., Gatto, J., Arasu, P., and Ruvkun, G. (1991) Negative regulatory sequences in the lin-14 3′-untranslated region are necessary to generate a temporal switch during *Caenorhabditis elegans* development, *Genes Dev 5*, 1813–1824.
3. Wightman, B., Ha, I., and Ruvkun, G. (1993) Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*, *Cell 75*, 855–862.
4. Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*, *Nature 403*, 901–906.
5. Slack, F. J., Basson, M., Liu, Z., Ambros, V., Horvitz, H. R., and Ruvkun, G. (2000) The lin-41 RBCC gene acts in the *C. elegans* heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor, *Mol Cell 5*, 659–669.
6. Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs, *Science 294*, 853–858.
7. Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*, *Science 294*, 858–862.
8. Lee, R. C., and Ambros, V. (2001) An extensive class of small RNAs in *Caenorhabditis elegans*, *Science 294*, 862–864.
9. Lai, E. C., Tomancak, P., Williams, R. W., and Rubin, G. M. (2003) Computational identification of Drosophila microRNA genes, *Genome Biol 4*, R42.
10. Lim, L. P., Glasner, M. E., Yekta, S., Burge, C. B., and Bartel, D. P. (2003) Vertebrate microRNA genes, *Science 299*, 1540.
11. Bartel, D. P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function, *Cell 116*, 281–297.
12. Ibanez-Ventoso, C., Vora, M., and Driscoll, M. (2008) Sequence relationships among *C. elegans*, *D. melanogaster* and human microRNAs highlight the extensive conservation of microRNAs in biology, *PLoS One 3*, e2818.
13. Chapman, E. J., and Carrington, J. C. (2007) Specialization and evolution of endogenous small RNA pathways, *Nat Rev Genet 8*, 884–896.
14. Millar, A. A., and Waterhouse, P. M. (2005) Plant and animal microRNAs: similarities and differences, *Funct Integr Genomics 5*, 129–135.
15. Lee, Y., Jeon, K., Lee, J. T., Kim, S., and Kim, V. N. (2002) MicroRNA maturation: stepwise processing and subcellular localization, *EMBO J 21*, 4663–4670.
16. Song, G., and Wang, L. (2008) MiR-433 and miR-127 arise from independent overlapping primary transcripts encoded by the miR-433-127 locus, *PLoS One 3*, e3574.
17. Song, G., and Wang, L. (2008) Transcriptional mechanism for the paired miR-433 and miR-127 genes by nuclear receptors SHP and ERRgamma, *Nucleic Acids Res 36*, 5727–5735.
18. Griffiths-Jones, S. (2004) The microRNA Registry, *Nucleic Acids Res 32*, D109–D111.
19. Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets, *Cell 120*, 15–20.
20. Filipowicz, W., Bhattacharyya, S. N., and Sonenberg, N. (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet 9*, 102–114.
21. Stark, A., Brennecke, J., Bushati, N., Russell, R. B., and Cohen, S. M. (2005) Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3′ UTR evolution, *Cell 123*, 1133–1146.
22. Lytle, J. R., Yario, T. A., and Steitz, J. A. (2007) Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5′ UTR as in the 3′ UTR, *Proc Natl Acad Sci U S A 104*, 9667–9672.

23. Tay, Y., Zhang, J., Thomson, A. M., Lim, B., and Rigoutsos, I. (2008) MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation, *Nature 455*, 1124–1128.
24. Baek, D., Villen, J., Shin, C., Camargo, F. D., Gygi, S. P., and Bartel, D. P. (2008) The impact of microRNAs on protein output, *Nature 455*, 64–71.
25. Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008) Widespread changes in protein synthesis induced by microRNAs, *Nature 455*, 58–63.
26. Grimson, A., Farh, K. K., Johnston, W. K., Garrett-Engele, P., Lim, L. P., and Bartel, D. P. (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing, *Mol Cell 27*, 91–105.
27. John, B., Enright, A. J., Aravin, A., Tuschl, T., Sander, C., and Marks, D. S. (2004) Human MicroRNA targets, *PLoS Biol 2*, e363.
28. John, B., Sander, C., and Marks, D. S. (2006) Prediction of human microRNA targets, *Methods Mol Biol 342*, 101–113.
29. Krek, A., Grun, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., MacMenamin, P., da Piedade, I., Gunsalus, K. C., Stoffel, M., and Rajewsky, N. (2005) Combinatorial microRNA target predictions, *Nat Genet 37*, 495–500.
30. Beitzinger, M., Peters, L., Zhu, J. Y., Kremmer, E., and Meister, G. (2007) Identification of human microRNA targets from isolated argonaute protein complexes, *RNA Biol 4*, 76–84.
31. Chi, S. W., Zang, J. B., Mele, A., and Darnell, R. B. (2009) Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps, *Nature 460*(7254), 479–486.
32. Landthaler, M., Gaidatzis, D., Rothballer, A., Chen, P. Y., Soll, S. J., Dinic, L., Ojo, T., Hafner, M., Zavolan, M., and Tuschl, T. (2008) Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs, *RNA 14*, 2580–2596.
33. Xiao, C., and Rajewsky, K. (2009) MicroRNA control in the immune system: basic principles, *Cell 136*, 26–36.
34. Ma, L., and Weinberg, R. A. (2008) Micromanagers of malignancy: role of microRNAs in regulating metastasis, *Trends Genet 24*, 448–456.
35. Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R., and Golub, T. R. (2005) MicroRNA expression profiles classify human cancers, *Nature 435*, 834–838.
36. Bonci, D., Coppola, V., Musumeci, M., Addario, A., Giuffrida, R., Memeo, L., D'Urso, L., Pagliuca, A., Biffoni, M., Labbaye, C., Bartucci, M., Muto, G., Peschle, C., and De Maria, R. (2008) The miR-15a-miR-16-1 cluster controls prostate cancer by targeting multiple oncogenic activities, *Nat Med 14*, 1271–1277.
37. Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K. L., Brown, D., and Slack, F. J. (2005) RAS is regulated by the let-7 microRNA family, *Cell 120*, 635–647.
38. Lee, Y. S., and Dutta, A. (2007) The tumor suppressor microRNA let-7 represses the HMGA2 oncogene, *Genes Dev 21*, 1025–1030.
39. He, L., Thomson, J. M., Hemann, M. T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S. W., Hannon, G. J., and Hammond, S. M. (2005) A microRNA polycistron as a potential human oncogene, *Nature 435*, 828–833.
40. Esquela-Kerscher, A., and Slack, F. J. (2006) Oncomirs – microRNAs with a role in cancer, *Nat Rev Cancer 6*, 259–269.
41. Calin, G. A., Ferracin, M., Cimmino, A., Di Leva, G., Shimizu, M., Wojcik, S. E., Iorio, M. V., Visone, R., Sever, N. I., Fabbri, M., Iuliano, R., Palumbo, T., Pichiorri, F., Roldo, C., Garzon, R., Sevignani, C., Rassenti, L., Alder, H., Volinia, S., Liu, C. G., Kipps, T. J., Negrini, M., and Croce, C. M. (2005) A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia, *N Engl J Med 353*, 1793–1801.
42. Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., and Croce, C. M. (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers, *Proc Natl Acad Sci U S A 101*, 2999–3004.
43. Wang, V., and Wu, W. (2009) MicroRNA-based therapeutics for cancer, *BioDrugs 23*, 15–23.
44. Visone, R., and Croce, C. M. (2009) MiRNAs and cancer, *Am J Pathol 174*, 1131–1138.
45. Diederichs, S., and Haber, D. A. (2006) Sequence variations of microRNAs in human

cancer: alterations in predicted secondary structure do not affect processing, *Cancer Res 66*, 6097–6104.

46. Makunin, I. V., Pheasant, M., Simons, C., and Mattick, J. S. (2007) Orthologous microRNA genes are located in cancer-associated genomic regions in human and mouse, *PLoS One 2*, e1133.
47. Chuang, J. C., and Jones, P. A. (2007) Epigenetics and microRNAs, *Pediatr Res 61*, 24R–29R.
48. Lujambio, A., Calin, G. A., Villanueva, A., Ropero, S., Sanchez-Cespedes, M., Blanco, D., Montuenga, L. M., Rossi, S., Nicoloso, M. S., Faller, W. J., Gallagher, W. M., Eccles, S. A., Croce, C. M., and Esteller, M. (2008) A microRNA DNA methylation signature for human cancer metastasis, *Proc Natl Acad Sci USA 105*, 13556–13561.
49. Lujambio, A., Ropero, S., Ballestar, E., Fraga, M. F., Cerrato, C., Setien, F., Casado, S., Suarez-Gauthier, A., Sanchez-Cespedes, M., Git, A., Spiteri, I., Das, P. P., Caldas, C., Miska, E., and Esteller, M. (2007) Genetic unmasking of an epigenetically silenced microRNA in human cancer cells, *Cancer Res 67*, 1424–1429.
50. Yang, N., Coukos, G., and Zhang, L. (2008) MicroRNA epigenetic alterations in human cancer: one step forward in diagnosis and treatment, *Int J Cancer 122*, 963–968.
51. Kefas, B., Godlewski, J., Comeau, L., Li, Y., Abounader, R., Hawkinson, M., Lee, J., Fine, H., Chiocca, E. A., Lawler, S., and Purow, B. (2008) microRNA-7 inhibits the epidermal growth factor receptor and the Akt pathway and is down-regulated in glioblastoma, *Cancer Res 68*, 3566–3572.
52. Lee, E. J., Baek, M., Gusev, Y., Brackett, D. J., Nuovo, G. J., and Schmittgen, T. D. (2008) Systematic evaluation of microRNA processing patterns in tissues, cell lines, and tumors, *RNA 14*, 35–42.
53. Michael, M. Z., Susan, M. O. C., van Holst Pellekaan, N. G., Young, G. P., and James, R. J. (2003) Reduced accumulation of specific microRNAs in colorectal neoplasia, *Mol Cancer Res 1*, 882–891.
54. Winter, J., Jung, S., Keller, S., Gregory, R. I., and Diederichs, S. (2009) Many roads to maturity: microRNA biogenesis pathways and their regulation, *Nat Cell Biol 11*, 228–234.
55. Gaur, A., Jewell, D. A., Liang, Y., Ridzon, D., Moore, J. H., Chen, C., Ambros, V. R., and Israel, M. A. (2007) Characterization of microRNA expression levels and their biological correlates in human cancer cell lines, *Cancer Res 67*, 2456–2468.
56. Porkka, K. P., Pfeiffer, M. J., Waltering, K. K., Vessella, R. L., Tammela, T. L., and Visakorpi, T. (2007) MicroRNA expression profiling in prostate cancer, *Cancer Res 67*, 6130–6135.
57. Visone, R., Pallante, P., Vecchione, A., Cirombella, R., Ferracin, M., Ferraro, A., Volinia, S., Coluzzi, S., Leone, V., Borbone, E., Liu, C. G., Petrocca, F., Troncone, G., Calin, G. A., Scarpa, A., Colato, C., Tallini, G., Santoro, M., Croce, C. M., and Fusco, A. (2007) Specific microRNAs are downregulated in human thyroid anaplastic carcinomas, *Oncogene 26*, 7590–7595.
58. Cai, X., Hagedorn, C. H., and Cullen, B. R. (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs, *RNA 10*, 1957–1966.
59. Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H., and Kim, V. N. (2004) MicroRNA genes are transcribed by RNA polymerase II, *EMBO J 23*, 4051–4060.
60. Borchert, G. M., Lanier, W., and Davidson, B. L. (2006) RNA polymerase III transcribes human microRNAs, *Nat Struct Mol Biol 13*, 1097–1101.
61. Bortolin-Cavaille, M. L., Dance, M., Weber, M., and Cavaille, J. (2009) C19MC microRNAs are processed from introns of large Pol-II, non-protein-coding transcripts, *Nucleic Acids Res 37*, 3464–3473.
62. Dews, M., Homayouni, A., Yu, D., Murphy, D., Sevignani, C., Wentzel, E., Furth, E. E., Lee, W. M., Enders, G. H., Mendell, J. T., and Thomas-Tikhonenko, A. (2006) Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster, *Nat Genet 38*, 1060–1065.
63. He, L., He, X., Lim, L. P., de Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., Ridzon, D., Jackson, A. L., Linsley, P. S., Chen, C., Lowe, S. W., Cleary, M. A., and Hannon, G. J. (2007) A microRNA component of the p53 tumour suppressor network, *Nature 447*, 1130–1134.
64. Brueckner, B., Stresemann, C., Kuner, R., Mund, C., Musch, T., Meister, M., Sultmann, H., and Lyko, F. (2007) The human let-7a-3 locus contains an epigenetically regulated microRNA gene with oncogenic function, *Cancer Res 67*, 1419–1423.
65. Blaszczyk, J., Tropea, J. E., Bubunenko, M., Routzahn, K. M., Waugh, D. S., Court, D. L., and Ji, X. (2001) Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage, *Structure 9*, 1225–1236.

66. Filippov, V., Solovyev, V., Filippova, M., and Gill, S. S. (2000) A novel type of RNase III family proteins in eukaryotes, *Gene 245*, 213–221.
67. Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V. N. (2003) The nuclear RNase III Drosha initiates microRNA processing, *Nature 425*, 415–419.
68. Morlando, M., Ballarino, M., Gromak, N., Pagano, F., Bozzoni, I., and Proudfoot, N. J. (2008) Primary microRNA transcripts are processed co-transcriptionally, *Nat Struct Mol Biol 15*, 902–909.
69. Pawlicki, J. M., and Steitz, J. A. (2008) Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production, *J Cell Biol 182*, 61–76.
70. Gregory, R. I., Yan, K. P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004) The microprocessor complex mediates the genesis of microRNAs, *Nature 432*, 235–240.
71. Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F., and Hannon, G. J. (2004) Processing of primary microRNAs by the microprocessor complex, *Nature 432*, 231–235.
72. Han, J., Lee, Y., Yeom, K. H., Nam, J. W., Heo, I., Rhee, J. K., Sohn, S. Y., Cho, Y., Zhang, B. T., and Kim, V. N. (2006) Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex, *Cell 125*, 887–901.
73. Landthaler, M., Yalcin, A., and Tuschl, T. (2004) The human DiGeorge syndrome critical region gene 8 and Its *D. melanogaster* homolog are required for miRNA biogenesis, *Curr Biol 14*, 2162–2167.
74. Yeom, K. H., Lee, Y., Han, J., Suh, M. R., and Kim, V. N. (2006) Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing, *Nucleic Acids Res 34*, 4622–4629.
75. Han, J., Pedersen, J. S., Kwon, S. C., Belair, C. D., Kim, Y. K., Yeom, K. H., Yang, W. Y., Haussler, D., Blelloch, R., and Kim, V. N. (2009) Posttranscriptional crossregulation between Drosha and DGCR8, *Cell 136*, 75–84.
76. Zeng, Y., and Cullen, B. R. (2003) Sequence requirements for micro RNA processing and function in human cells, *RNA 9*, 112–123.
77. Zeng, Y., and Cullen, B. R. (2005) Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences, *J Biol Chem 280*, 27595–27603.
78. Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M., and Lai, E. C. (2007) The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila, *Cell 130*, 89–100.
79. Ruby, J. G., Jan, C. H., and Bartel, D. P. (2007) Intronic microRNA precursors that bypass Drosha processing, *Nature 448*, 83–86.
80. Berezikov, E., Chung, W. J., Willis, J., Cuppen, E., and Lai, E. C. (2007) Mammalian mirtron genes, *Mol Cell 28*, 328–336.
81. Bohnsack, M. T., Czaplinski, K., and Gorlich, D. (2004) Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs, *RNA 10*, 185–191.
82. Calado, A., Treichel, N., Muller, E. C., Otto, A., and Kutay, U. (2002) Exportin-5-mediated nuclear export of eukaryotic elongation factor 1A and tRNA, *EMBO J 21*, 6216–6224.
83. Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E., and Kutay, U. (2004) Nuclear export of microRNA precursors, *Science 303*, 95–98.
84. Zeng, Y., and Cullen, B. R. (2004) Structural requirements for pre-microRNA binding and nuclear export by Exportin 5, *Nucleic Acids Res 32*, 4776–4785.
85. Yi, R., Doehle, B. P., Qin, Y., Macara, I. G., and Cullen, B. R. (2005) Overexpression of exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs, *RNA 11*, 220–226.
86. Maniataki, E., and Mourelatos, Z. (2005) A human, ATP-independent, RISC assembly machine fueled by pre-miRNA, *Genes Dev 19*, 2979–2990.
87. Gregory, R. I., Chendrimada, T. P., Cooch, N., and Shiekhattar, R. (2005) Human RISC couples microRNA biogenesis and posttranscriptional gene silencing, *Cell 123*, 631–640.
88. Haase, A. D., Jaskiewicz, L., Zhang, H., Laine, S., Sack, R., Gatignol, A., and Filipowicz, W. (2005) TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing, *EMBO Rep 6*, 961–967.
89. Lee, Y., Hur, I., Park, S. Y., Kim, Y. K., Suh, M. R., and Kim, V. N. (2006) The role of PACT in the RNA silencing pathway, *EMBO J 25*, 522–532.
90. MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V., and Doudna, J. A. (2008) In vitro reconstitution of the human RISC-loading complex, *Proc Natl Acad Sci USA 105*, 512–517.
91. Diederichs, S., and Haber, D. A. (2007) Dual role for argonautes in microRNA processing

and posttranscriptional regulation of microRNA expression, *Cell 131*, 1097–1108.

92. Leuschner, P. J., Ameres, S. L., Kueng, S., and Martinez, J. (2006) Cleavage of the siRNA passenger strand during RISC assembly in human cells, *EMBO Rep 7*, 314–320.
93. Matranga, C., Tomari, Y., Shin, C., Bartel, D. P., and Zamore, P. D. (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes, *Cell 123*, 607–620.
94. Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H., and Siomi, M. C. (2005) Slicer function of Drosophila Argonautes and its involvement in RISC formation, *Genes Dev 19*, 2837–2848.
95. Rand, T. A., Petersen, S., Du, F., and Wang, X. (2005) Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation, *Cell 123*, 621–629.
96. Carmell, M. A., and Hannon, G. J. (2004) RNase III enzymes and the initiation of gene silencing, *Nat Struct Mol Biol 11*, 214–218.
97. Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference, *Nature 409*, 363–366.
98. Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D. L., Fire, A., Ruvkun, G., and Mello, C. C. (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing, *Cell 106*, 23–34.
99. Hutvagner, G., McLachlan, J., Pasquinelli, A. E., Balint, E., Tuschl, T., and Zamore, P. D. (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA, *Science 293*, 834–838.
100. Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J., and Plasterk, R. H. (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*, *Genes Dev 15*, 2654–2659.
101. Knight, S. W., and Bass, B. L. (2001) A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*, *Science 293*, 2269–2271.
102. Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004) Single processing center models for human Dicer and bacterial RNase III, *Cell 118*, 57–68.
103. Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004) Nucleic acid 3′-end recognition by the Argonaute2 PAZ domain, *Nat Struct Mol Biol 11*, 576–577.
104. Ma, J. B., Ye, K., and Patel, D. J. (2004) Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain, *Nature 429*, 318–322.
105. Song, J. J., Liu, J., Tolia, N. H., Schneiderman, J., Smith, S. K., Martienssen, R. A., Hannon, G. J., and Joshua-Tor, L. (2003) The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes, *Nat Struct Biol 10*, 1026–1032.
106. Yan, K. S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M. M. (2003) Structure and conserved RNA binding of the PAZ domain, *Nature 426*, 468–474.
107. Bernstein, E., Kim, S. Y., Carmell, M. A., Murchison, E. P., Alcorn, H., Li, M. Z., Mills, A. A., Elledge, S. J., Anderson, K. V., and Hannon, G. J. (2003) Dicer is essential for mouse development, *Nat Genet 35*, 215–217.
108. Forstemann, K., Horwich, M. D., Wee, L., Tomari, Y., and Zamore, P. D. (2007) Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1, *Cell 130*, 287–297.
109. Liu, Q., Rand, T. A., Kalidas, S., Du, F., Kim, H. E., Smith, D. P., and Wang, X. (2003) R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway, *Science 301*, 1921–1925.
110. Tomari, Y., Matranga, C., Haley, B., Martinez, N., and Zamore, P. D. (2004) A protein sensor for siRNA asymmetry, *Science 306*, 1377–1380.
111. Chu, C. Y., and Rana, T. M. (2006) Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54, *PLoS Biol 4*, e210.
112. Meister, G., Landthaler, M., Peters, L., Chen, P. Y., Urlaub, H., Luhrmann, R., and Tuschl, T. (2005) Identification of novel argonaute-associated proteins, *Curr Biol 15*, 2149–2155.
113. Robb, G. B., and Rana, T. M. (2007) RNA helicase A interacts with RISC in human cells and functions in RISC loading, *Mol Cell 26*, 523–537.
114. Salzman, D. W., Shubert-Coleman, J., and Furneaux, H. (2007) P68 RNA helicase unwinds the human let-7 microRNA precursor duplex and is required for let-7-directed silencing of gene expression, *J Biol Chem 282*, 32773–32779.
115. Fukuda, T., Yamagata, K., Fujiyama, S., Matsumoto, T., Koshida, I., Yoshimura, K.,

Mihara, M., Naitou, M., Endoh, H., Nakamura, T., Akimoto, C., Yamamoto, Y., Katagiri, T., Foulds, C., Takezawa, S., Kitagawa, H., Takeyama, K., O'Malley, B. W., and Kato, S. (2007) DEAD-box RNA helicase subunits of the Drosha complex are required for processing of rRNA and a subset of microRNAs, *Nat Cell Biol 9*, 604–611.

116. Khvorova, A., Reynolds, A., and Jayasena, S. D. (2003) Functional siRNAs and miRNAs exhibit strand bias, *Cell 115*, 209–216.
117. Schwarz, D. S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P. D. (2003) Asymmetry in the assembly of the RNAi enzyme complex, *Cell 115*, 199–208.
118. Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., Wu, L., Li, S., Zhou, H., Long, C., Chen, S., Hannon, G. J., and Qi, Y. (2008) Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5′ terminal nucleotide, *Cell 133*, 116–127.
119. Tomari, Y., Du, T., and Zamore, P. D. (2007) Sorting of Drosophila small silencing RNAs, *Cell 130*, 299–308.
120. Azuma-Mukai, A., Oguri, H., Mituyama, T., Qian, Z. R., Asai, K., Siomi, H., and Siomi, M. C. (2008) Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing, *Proc Natl Acad Sci USA 105*, 7964–7969.
121. Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs, *Mol Cell 15*, 185–197.
122. Eulalio, A., Huntzinger, E., and Izaurralde, E. (2008) Getting to the root of miRNA-mediated gene silencing, *Cell 132*, 9–14.
123. Liu, J., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J. J., Hammond, S. M., Joshua-Tor, L., and Hannon, G. J. (2004) Argonaute2 is the catalytic engine of mammalian RNAi, *Science 305*, 1437–1441.
124. Yekta, S., Shih, I. H., and Bartel, D. P. (2004) MicroRNA-directed cleavage of HOXB8 mRNA, *Science 304*, 594–596.
125. Eulalio, A., Rehwinkel, J., Stricker, M., Huntzinger, E., Yang, S. F., Doerks, T., Dorner, S., Bork, P., Boutros, M., and Izaurralde, E. (2007) Target-specific requirements for enhancers of decapping in miRNA-mediated gene silencing, *Genes Dev 21*, 2558–2570.
126. Kiriakidou, M., Tan, G. S., Lamprinaki, S., De Planell-Saguer, M., Nelson, P. T., and Mourelatos, Z. (2007) An mRNA m7G cap binding-like motif within human Ago2 represses translation, *Cell 129*, 1141–1151.
127. Chendrimada, T. P., Finn, K. J., Ji, X., Baillat, D., Gregory, R. I., Liebhaber, S. A., Pasquinelli, A. E., and Shiekhattar, R. (2007) MicroRNA silencing through RISC recruitment of eIF6, *Nature 447*, 823–828.
128. Maroney, P. A., Yu, Y., Fisher, J., and Nilsen, T. W. (2006) Evidence that microRNAs are associated with translating messenger RNAs in human cells, *Nat Struct Mol Biol 13*, 1102–1107.
129. Maroney, P. A., Yu, Y., and Nilsen, T. W. (2006) MicroRNAs, mRNAs, and translation, *Cold Spring Harb Symp Quant Biol 71*, 531–535.
130. Beilharz, T. H., Humphreys, D. T., Clancy, J. L., Thermann, R., Martin, D. I., Hentze, M. W., and Preiss, T. (2009) microRNA-mediated messenger RNA deadenylation contributes to translational repression in mammalian cells, *PLoS One 4*, e6783.
131. Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A. E. (2005) Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation, *Cell 122*, 553–563.
132. Behm-Ansmant, I., Rehwinkel, J., and Izaurralde, E. (2006) MicroRNAs silence gene expression by repressing protein expression and/or by promoting mRNA decay, *Cold Spring Harb Symp Quant Biol 71*, 523–530.
133. Wu, L., and Belasco, J. G. (2005) Micro-RNA regulation of the mammalian lin-28 gene during neuronal differentiation of embryonal carcinoma cells, *Mol Cell Biol 25*, 9198–9208.
134. Pauley, K. M., Eystathioy, T., Jakymiw, A., Hamel, J. C., Fritzler, M. J., and Chan, E. K. (2006) Formation of GW bodies is a consequence of microRNA genesis, *EMBO Rep 7*, 904–910.
135. Han, J., Lee, Y., Yeom, K. H., Kim, Y. K., Jin, H., and Kim, V. N. (2004) The Drosha-DGCR8 complex in primary microRNA processing, *Genes Dev 18*, 3016–3027.
136. Hwang, H. W., Wentzel, E. A., and Mendell, J. T. (2007) A hexanucleotide element directs microRNA nuclear import, *Science 315*, 97–100.
137. Kennedy, S., Wang, D., and Ruvkun, G. (2004) A conserved siRNA-degrading RNase negatively regulates RNA interference in *C. elegans*, *Nature 427*, 645–649.
138. Ramachandran, V., and Chen, X. (2008) Degradation of microRNAs by a family of exoribonucleases in Arabidopsis, *Science 321*, 1490–1492.

139. Kumar, M. S., Lu, J., Mercer, K. L., Golub, T. R., and Jacks, T. (2007) Impaired microRNA processing enhances cellular transformation and tumorigenesis, *Nat Genet 39*, 673–677.

140. Blenkiron, C., Goldstein, L. D., Thorne, N. P., Spiteri, I., Chin, S. F., Dunning, M. J., Barbosa-Morais, N. L., Teschendorff, A. E., Green, A. R., Ellis, I. O., Tavare, S., Caldas, C., and Miska, E. A. (2007) MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype, *Genome Biol 8*, R214.

141. Chiosea, S., Jelezcova, E., Chandran, U., Luo, J., Mantha, G., Sobol, R. W., and Dacic, S. (2007) Overexpression of Dicer in precursor lesions of lung adenocarcinoma, *Cancer Res 67*, 2345–2350.

142. Karube, Y., Tanaka, H., Osada, H., Tomida, S., Tatematsu, Y., Yanagisawa, K., Yatabe, Y., Takamizawa, J., Miyoshi, S., Mitsudomi, T., and Takahashi, T. (2005) Reduced expression of Dicer associated with poor prognosis in lung cancer patients, *Cancer Sci 96*, 111–115.

143. Muralidhar, B., Goldstein, L. D., Ng, G., Winder, D. M., Palmer, R. D., Gooding, E. L., Barbosa-Morais, N. L., Mukherjee, G., Thorne, N. P., Roberts, I., Pett, M. R., and Coleman, N. (2007) Global microRNA profiles in cervical squamous cell carcinoma depend on Drosha expression levels, *J Pathol 212*, 368–377.

144. Melo, S. A., Ropero, S., Moutinho, C., Aaltonen, L. A., Yamamoto, H., Calin, G. A., Rossi, S., Fernandez, A. F., Carneiro, F., Oliveira, C., Ferreira, B., Liu, C. G., Villanueva, A., Capella, G., Schwartz, S., Jr., Shiekhattar, R., and Esteller, M. (2009) A TARBP2 mutation in human cancer impairs microRNA processing and DICER1 function, *Nat Genet 41*, 365–370.

145. Ma, E., MacRae, I. J., Kirsch, J. F., and Doudna, J. A. (2008) Autoinhibition of human dicer by its internal helicase domain, *J Mol Biol 380*, 237–243.

146. Forman, J. J., Legesse-Miller, A., and Coller, H. A. (2008) A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence, *Proc Natl Acad Sci U S A 105*, 14879–14884.

147. Carmell, M. A., Xuan, Z., Zhang, M. Q., and Hannon, G. J. (2002) The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis, *Genes Dev 16*, 2733–2742.

148. Nelson, P., Kiriakidou, M., Sharma, A., Maniataki, E., and Mourelatos, Z. (2003) The microRNA world: small is mighty, *Trends Biochem Sci 28*, 534–540.

149. Qi, H. H., Ongusaha, P. P., Myllyharju, J., Cheng, D., Pakkanen, O., Shi, Y., Lee, S. W., Peng, J., and Shi, Y. (2008) Prolyl 4-hydroxylation regulates Argonaute 2 stability, *Nature 455*, 421–424.

150. Rudel, S., and Meister, G. (2008) Phosphorylation of Argonaute proteins: regulating gene regulators, *Biochem J 413*, e7–e9.

151. Zeng, Y., Sankala, H., Zhang, X., and Graves, P. R. (2008) Phosphorylation of Argonaute 2 at serine-387 facilitates its localization to processing bodies, *Biochem J 413*, 429–436.

152. Diederichs, S., Jung, S., Rothenberg, S. M., Smolen, G. A., Mlody, B. G., and Haber, D. A. (2008) Coexpression of Argonaute-2 enhances RNA interference toward perfect match binding sites, *Proc Natl Acad Sci U S A 105*, 9284-9289.

153. Saito, K., Ishizuka, A., Siomi, H., and Siomi, M. C. (2005) Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells, *PLoS Biol 3*, e235.

154. Guil, S., and Caceres, J. F. (2007) The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a, *Nat Struct Mol Biol 14*, 591–596.

155. Michlewski, G., Guil, S., Semple, C. A., and Caceres, J. F. (2008) Posttranscriptional regulation of miRNAs harboring conserved terminal loops, *Mol Cell 32*, 383–393.

156. Chan, J. A., Krichevsky, A. M., and Kosik, K. S. (2005) MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells, *Cancer Res 65*, 6029–6033.

157. Davis, B. N., Hilyard, A. C., Lagna, G., and Hata, A. (2008) SMAD proteins control DROSHA-mediated microRNA maturation, *Nature 454*, 56–61.

158. Suzuki, H. I., Yamagata, K., Sugimoto, K., Iwamoto, T., Kato, S., and Miyazono, K. (2009) Modulation of microRNA processing by p53, *Nature 460*, 529–533.

159. Lu, L., Katsaros, D., Shaverdashvili, K., Qian, B., Wu, Y., de la Longrais, I. A., Preti, M., Menato, G., and Yu, H. (2009) Pluripotent factor lin-28 and its homologue lin-28b in epithelial ovarian cancer and their associations with disease outcomes and expression of let-7a and IGF-II, *Eur J Cancer 45*(12), 2212–2218.

160. Newman, M. A., Thomson, J. M., and Hammond, S. M. (2008) Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing, *RNA 14*, 1539–1549.

161. Piskounova, E., Viswanathan, S. R., Janas, M., LaPierre, R. J., Daley, G. Q., Sliz, P., and Gregory, R. I. (2008) Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28, *J Biol Chem 283*, 21310–21314.

162. Viswanathan, S. R., Daley, G. Q., and Gregory, R. I. (2008) Selective blockade of microRNA processing by Lin28, *Science 320*, 97–100.

163. Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V. N. (2008) Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA, *Mol Cell 32*, 276–284.

164. Heo, I., Joo, C., Kim, Y. K., Ha, M., Yoon, M. J., Cho, J., Yeom, K. H., Han, J., and Kim, V. N. (2009) TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation., *Cell 138*, 696–708.

165. Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E. E., Nitsch, R., and Wulczyn, F. G. (2008) A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment, *Nat Cell Biol 10*, 987–993.

166. Duan, R., Pak, C., and Jin, P. (2007) Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA, *Hum Mol Genet 16*, 1124–1131.

167. Scadden, A. D. (2005) The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage, *Nat Struct Mol Biol 12*, 489–496.

168. Yang, W., Chendrimada, T. P., Wang, Q., Higuchi, M., Seeburg, P. H., Shiekhattar, R., and Nishikura, K. (2006) Modulation of microRNA processing and expression through RNA editing by ADAR deaminases, *Nat Struct Mol Biol 13*, 13–21.

169. Kawahara, Y., Megraw, M., Kreider, E., Iizasa, H., Valente, L., Hatzigeorgiou, A. G., and Nishikura, K. (2008) Frequency and fate of microRNA editing in human brain, *Nucleic Acids Res 36*, 5270–5280.

170. Kawahara, Y., Zinshteyn, B., Chendrimada, T. P., Shiekhattar, R., and Nishikura, K. (2007) RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex, *EMBO Rep 8*, 763–769.

171. Kawahara, Y., Zinshteyn, B., Sethupathy, P., Iizasa, H., Hatzigeorgiou, A. G., and Nishikura, K. (2007) Redirection of silencing targets by adenosine-to-inosine editing of miRNAs, *Science 315*, 1137–1140.

172. Kim, V. N., Han, J., and Siomi, M. C. (2009) Biogenesis of small RNAs in animals, *Nat Rev Mol Cell Biol 10*, 126–139.

Chapter 2

Computational Identification of miRNAs Involved in Cancer

Anastasis Oulas, Nestoras Karathanasis, and Panayiota Poirazi

Abstract

Changes in the structure and/or the expression of protein-coding genes were thought to be the major cause of cancer for many decades. However, the recent discovery of non-coding RNA (ncRNA) transcripts suggests that the molecular biology of cancer is far more complex. MicroRNAs (miRNAs) are key players of the family of ncRNAs and they have been under extensive investigation because of their involvement in carcinogenesis, often taking up roles of tumor suppressors or oncogenes. Owing to the slow nature of experimental identification of miRNA genes, computational procedures have been applied as a valuable complement to cloning. Numerous computational tools, implemented to recognize the characteristic features of miRNA biogenesis, have resulted in the prediction of multiple novel miRNA genes. Computational approaches provide valuable clues as to which are the dominant features that characterize these regulatory units and furthermore act by narrowing down the search space making experimental verification faster and significantly cheaper. Moreover, in combination with large-scale, high-throughput methods, such as deep sequencing and tilling arrays, computational methods have aided in the discovery of putative molecular signatures of miRNA deregulation in human tumors. This chapter focuses on existing computational methods for identifying miRNA genes, provides an overview of the methodology undertaken by these tools, and underlies their contribution toward unraveling the role of miRNAs in cancer.

Key words: MicroRNAs, Gene prediction, Software tools comparison, Cancer

1. Introduction

Current estimates show that while over 30% of vertebrate genomes are transcribed (1), only 1% represents protein-coding genes; the rest are believed to be various types of non-coding RNA (ncRNA) genes. Currently, only ~700 human microRNA (miRNA) genes exist in the miRNA registry,[1] and it is anticipated that there may be thousands more. The role of these molecules in cancer has

[1] miRBase, release 13.0.

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_2, © Springer Science+Business Media, LLC 2011

lately received a great deal of the scientific community's attention. Specifically, recent findings indicate that alterations in the expression of several miRNAs are often present in human cancers, suggesting potential roles of miRNAs in carcinogenic processes. For example, the expression levels of *let-7* (2), *miR-15a/miR-16-1* cluster (3), and neighboring *miR-143/miR-145* (4) are found to be reduced in some malignancies, suggesting their potential role as tumor suppressors. In contrast, some other miRNAs such as the *miR-17-92* cluster (5–7) and *miR-155/BIC* (8) are overexpressed in various cancers, suggesting a possible oncogenic role. Furthermore, some miRNAs with altered expression levels appear to be associated with certain genetic alterations, such as deletion, amplification, and mutation. Regions that are prone to such genetic alterations are commonly referred to as cancer-associated genomic regions (CAGRs) and fragile sites (FRA) (9). MiRNA genes located within or in close proximity to these regions have been suggested to be associated with chromosomal events leading to carcinogenesis, as graphically illustrated in Fig. 1.

The large amount of unexplored non-coding regions in the human genome combined with the increasing importance of miRNAs in cancer highlights the need for fast, flexible, and reliable miRNA identification methods. Toward this goal, a number of different computational methods have been used to identify

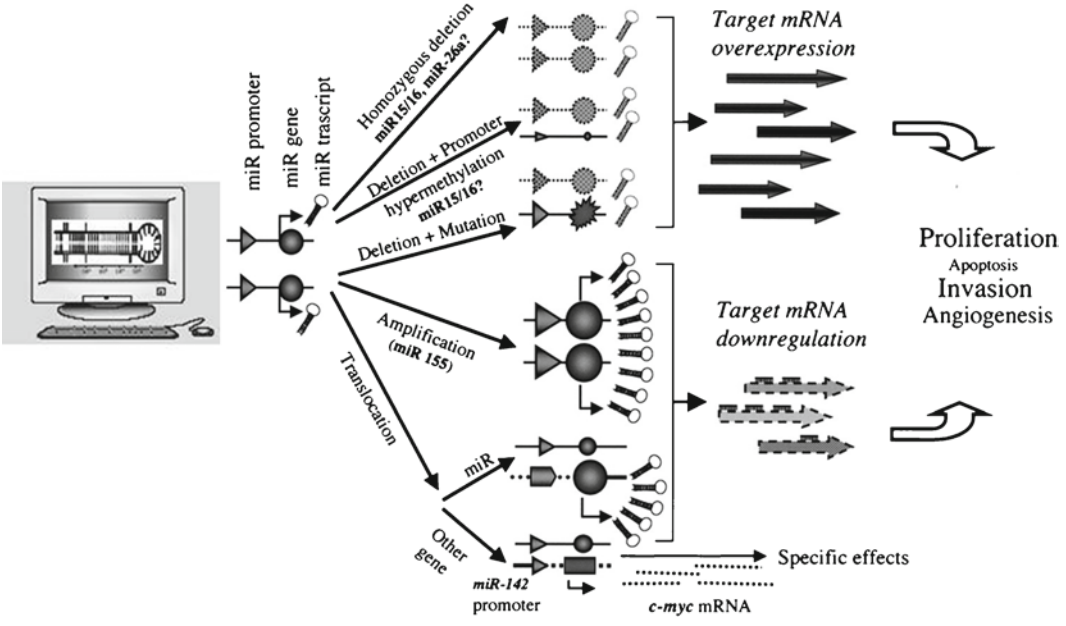


Fig. 1. MiRNAs as cancer players. Computational prediction initiates the search for putative miRNAs that play a role in tumorigenesis. Some of these proposed mechanisms are experimentally proven, like the deletion of *miR-15a/miR-16a* cluster in B-CLL (3, 49), the c-myc overexpression by the reposition near a putative *miR* promoter (9), or *miR143/miR-145* cluster downregulation in colon cancers (4). Figure adopted with permission from Calin et al. (9).

miRNA genes. Early studies focused on scanning for hairpin structures conserved between closely related species such as *Caenorhabditis elegans* and *Caenorhabditis briggsae* (10, 11), or using homology between known miRNAs and other regions in aligned genomes like human and mouse (12). Other approaches relied on conserved regions of synteny – conserved clustering of miRNAs in closely related genomes – to predict novel miRNAs (12). Subsequent computational studies utilized profile-based detection (13) as well as secondary structure alignment (14) of miRNAs using sequence conservation across multiple, highly divergent organisms (i.e., mouse and fugu).

The main drawback of the abovementioned tools is that they undertake a pipeline approach by applying stringent cut-offs and eliminating candidate miRNAs as the pipeline proceeds (10, 11). This results in the loss of numerous true miRNAs along the line. The use of homology by some tools (12–14) to detect novel miRNAs based on their similarity to previously identified miRNAs is another drawback. These methods obviously fall short when scanning distantly related sequences or when novel miRNAs lack detectable homologs.

The next generation of computational tools relied on more sophisticated machine learning algorithms such as support vector machine tools (SVMs) capable of taking into account multiple biological features such as free energy of the hairpin structure, paired bases, loop length, and stem conservation to predict novel miRNAs (15–17). Two very effective computational studies utilized hidden Markov models (HMMs) and a Bayesian classifier (18, 19) to simultaneously consider sequence and structure features at the nucleotide level for predicting miRNA genes. These studies, however, did not integrate conservation information in their algorithms, an important feature of the majority of miRNA genes. More recently, two computational tools miRRim (20) and SSCprofiler (21) also employing HMMs proved to be very effective, achieving high performance on identifying miRNAs in the human genome.

With the advent of large-scale, high-throughput methods such as tiling arrays or deep sequencing, the identification of novel miRNA genes is taking a different turn (22–24). These methods are exceptionally useful as they produce large datasets that offer a relatively accurate expression map for small RNAs in the genome. However, as large-scale expression data are usually limited by the specific tissue and developmental stage of their samples, only the coupling of such data to computational tools (as done in two recent studies (20, 21)) can facilitate rapid and precise detection of novel miRNAs while at the same time giving greater credence to computational predictions.

2. Tool Comparison

In the following paragraphs, we overview the representative examples of miRNA gene prediction tools and highlight their most important characteristics. The tools are organized according to the biological features they implement.

2.1. Sequence-Based Prediction

The initial computational tools for miRNA identification were based on sequence conservation with already cloned miRNAs. For instance, Quintana et al. (25) predicted 34 novel miRNAs using tissue-specific cloning. Almost all of these miRNAs were conserved in the human genome and frequently in non-mammalian vertebrate genomes such as pufferfish. One interesting observation was that certain miRNAs showed increased expression in heart, liver, or brain when compared with the entire miRNA population known at the time, proposing a role of these miRNAs in tissue specification or cell lineage decision.

2.2. Sequence, Structure, and Closely Related Species Conservation

Early computational methods for miRNA gene prediction relied on rules derived from sequence and structural features of miRNA precursors as well as their degree of conservation across species. The use of conservation was usually limited to pairwise comparison of closely related species. MiRscan (11) and MiRseeker (10) are two representative approaches for this category of tools. The MiRscan (11) tool implements a probabilistic method that uses known miRNAs as a training set in order to derive new miRNAs based on their degree of similarity. Specifically, the tool was developed as follows:

1. A total of 36,000 conserved sequences between *C. elegans* and *C. briggsae* (WU-BLAST cut-off $E<1.8$) were scanned using RNAfold in search for hairpin structures.
2. Fifty of them, which have previously been reported as real miRNAs, were used as the training set in order to derive statistical information regarding a set of characteristic features (shown in Fig. 2).
3. Using the trained algorithm, all 36,000 sequences were evaluated with respect to their potential as miRNA precursors. Identification of new miRNA precursors by MiRscan is achieved via the use of a sliding window of 21 nt that is shifted across each hairpin candidate searching for the presence of specific miRNA features. These features include:
    - Complementary base pairing;
    - High degree of 5′ and 3′ conservation;
    - Certain degree of bulge symmetry;
    - Specific distance from the loop;
    - Initial pentamer properties.

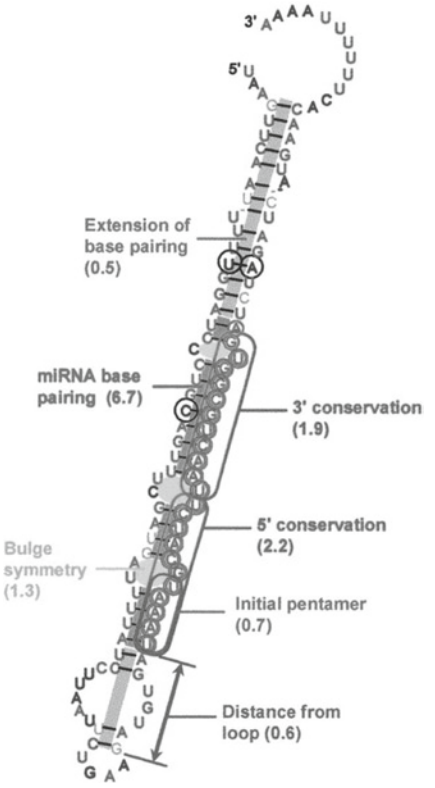


Fig. 2. Criteria used by MiRscan to identify miRNA genes among aligned segments of two genomes. The typical hairpin-like structure of miRNAs and the seven components of the MiRscan score for mir-232 of *Caenorhabditis elegans*/*Caenorhabditis briggsae*. Figure adopted with permission from Lim et al. (11).

MiRseeker (10) is another tool belonging to the same category. It has been used for predicting miRNA genes in two Drosophila species, namely *Drosophila melanogaster* and *Drosophila pseudoobscura*, via the use of the following three-step pipeline approach:

1. The two genomes are aligned to find all conserved regions.
2. MiRseeker, which includes MFOLD, is used to identify miRNA genes. A very important feature of MiRseeker which enhances its prediction power is the consideration of pattern divergence. As per the authors' report, there is less selective pressure and hence less conservation in the loop sequences.
3. The method is evaluated according to its ability to assign high scores to 24 known Drosophila miRNAs.

2.3. Multiple Species Conservation and Conserved Regions of Synteny

Another approach that has been used to predict novel human miRNAs is known as phylogenetic shadowing (26, 27). This approach uses multiple sequence comparisons of closely related species and allows the identification of conserved regions.

The method is based on observing conserved islands of a specific length, which are considered likely to be miRNA genes, in an otherwise unconserved region (26). An example of phylogenetic shadowing application to the identification of miRNA genes is found in (27), using the following steps:

1. Sequences from more than 100 miRNA regions in ten different primates are compared in order to infer a characteristic profile:
   - Variation in the loop sequences;
   - Conservation in stem of hairpins;
   - Significant decrease in conservation of sequence flanking the hairpins.
2. This profile is used to identify new miRNAs in pairwise alignments of more divergent species such as human and mouse or human and rat.
3. Additional filtering is performed according to the folding energy of candidate sequences.

A total of 976 potential human miRNAs have been identified using this method. This set contains over 80% of all known human miRNAs in version 3.1 of the miRNA registry (http://microrna.sanger.ac.uk/sequences/). Northern blot analyses combined with database searches reach a conservative estimate of 200–300 verified novel human miRNAs, a twofold increase over previous studies (28). Strong conservation over all species is only evident for two well-known miRNA genes.

Another approach is the use of full genome sequence alignment (12, 29, 30). The motivation being that human and mouse miRNAs should reside in conserved regions of synteny. A representative study using full genome alignments is found in (12), whereby:

1. The BLAT comparison tool (31) is used to compare the entire set of human and mouse precursor and mature miRNAs in the miRNA registry, version 2.2 (http://microrna.sanger.ac.uk/sequences/).
2. The results are further filtered using secondary structure prediction tools, like MFOLD and other criteria (such as G:U base pairing).
3. Characteristic features of some miRNAs are used to identify miRNA gene clusters and display conservation in the location of the clusters in comparison with other neighboring genes.

The findings of this work included the prediction of 80 new putative miRNAs genes.

The computational approaches described so far utilize information regarding closely related species and homology searches

using sequences of already cloned miRNAs. These methods proved to be successful, leading to the prediction of multiple new miRNA genes. However, as they were bound by species similarity they eventually reached a prediction limit. To overcome this obstacle, researchers turned their attention to the prediction of miRNAs that do not show conservation to other known miRNAs and are not highly conserved across closely related species.

2.4. A More General Model for miRNA Gene Prediction

miRAlign (14) is one of the first tools that detects new miRNAs based on both sequence and structure alignments without implementing conservation features. The main characteristics that differentiate this tool from the existing homolog search methods are:

1. By applying a relatively loose conservation of the mature miRNA sequence, it has the ability to find distant homologs.
2. It considers more structural properties by introducing a structure alignment strategy that can use each single miRNA as a query for genomic searches.

The tool has been shown to perform better in comparison with other tools such as BLAST or ERPIN (32), and its main advantage is the prediction of more distant miRNA homologs or orthologs.

2.5. The Next Generation of Tools

2.5.1. HMM Tools

A *Hidden Markov Model* (*HMM*) is a statistical model in which the system being modeled is assumed to be a Markov process with unobserved state.

2.5.1.1. ProMir – Nam et al. (18)

1. A highly specific probabilistic co-learning method was handcrafted, based on the paired HMM.
2. This method combines both sequence and structural characteristics of miRNA genes in a probabilistic framework and simultaneously decides if a miRNA gene and the mature miRNA are present by detecting the signals for the site cleaved by Drosha.
3. miRNA gene candidates are finally filtered using conservation across multiple divergent species.

Most recently, two freely available prediction tools (SSCprofiler (21) and miRRim (20)) have been shown to predict miRNA genes with high accuracy.

2.5.1.2. SSCprofiler Oulas et al. (21)

1. SSCprofiler utilizes profile HMM trained to recognize key biological features of miRNAs such as sequence, structure, and conservation in order to identify novel miRNA precursors as shown in Fig. 3.
2. SSCprofiler is trained to learn the characteristic features of human miRNA precursors with high accuracy, and the trained model is applied on CAGRs in search of novel miRNA genes.

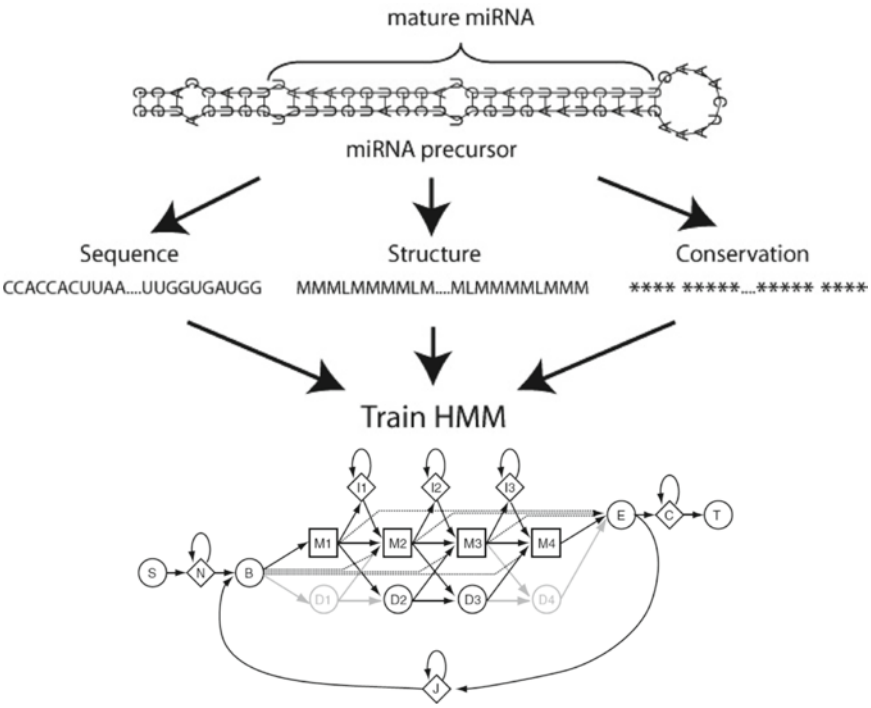


Fig. 3. The supervised procedure of training hidden Markov models (HMMs) for miRNA precursor identification. Biological features of miRNA biogenesis and conservation across other organisms are used as input for training. Initially, secondary structure prediction is performed using programs such as RNAfold. Every nucleotide position is henceforth represented by an "M" for match and an "L" for loop. This information is aligned integrated with conservation and sequence information for every nucleotide position and used to train the HMM. Once trained, the HMM can be utilized to analyze sequences of desired length and assign a likelihood score. The higher the score, the greater the chances of a candidate sequence being a true miRNA precursor. Figure adopted with permission from Oulas et al. (21).

3. Predictions are ranked according to expression information from a recently published full genome tiling array study (23), and the top four scoring candidates have been verified experimentally using Northern blot.

2.5.1.3. MiRRim – Terai et al. (20)

MiRRim is similar to SSCprofiler as it also uses an HMM algorithm that considers structure and conservation features for predicting novel miRNA genes. However, sequence information is not taken into consideration by the algorithm. The main conceptual differences between the two tools are provided in Table 1.

2.5.2. Bayes Classifiers

A *naive Bayes classifier* is a simple probabilistic classifier based on applying Bayes' theorem (from Bayesian statistics) with strong (naive) independence assumptions. In simple terms, a naive Bayes classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature. Depending on the precise nature of the probability model, naive Bayes classifiers can be trained very efficiently in a

Table 1
Comparison of SSCprofiler and miRRim tools

	SSCprofiler	miRRim
Biological features of miRNAs	Sequence, structure, and conservation	Structure and conservation
Negative data	Selected from 3′-UTRs, and filtered according to conservation and a minimum energy score to ensure that they resemble true miRNAs in both structure and conservation	Randomly selected 200 nt-long genomic regions with different degrees of conservation. No requirements for resemblance with true miRNAs
Sensitivity/specificity (validation set)	HMM score: 3 Sens: 88.95% Spec: 84.16%	Sens: ~70% Spec: ~90%
Generalization (blind test set)	Identification of 219 previously unseen miRNAs with an accuracy of 72.15%	No evaluation of performance on a blind test set
Scanning procedure	104 nt sliding window, shifted 1 nt at a time for positive as well as negative data	Size of sliding window and shift step unclear. Positive data ranged between 160 and 236 nt, negative data were 200 nt, implying a larger window and thus a smaller search space
Total number of hits	For a coverage of 96.0% (HMM threshold 11), ~5,800 miRNA hits for 350 MB (CAGRs) of the human genome	For a coverage of 91.0%, ~4,000 miRNA hits for the whole human genome
Expression information using high-throughput methods	Tiling array data from HeLa and HepG2 cells	No expression information is provided
Experimental verification	Successful verification of four top scoring miRNA candidates via Northern blot	No experimental verification is provided

supervised learning setting. An advantage of the naive Bayes classifier is that it requires a small amount of training data to estimate the parameters (means and variances of the variables) necessary for classification.

2.5.2.1. NaiveBayes – Yousef et al. (19)

1. This method generates the model automatically and identifies rules based on the miRNA gene sequence and structure; thus allowing the prediction of nonconserved miRNAs.
2. In addition, the method uses a comparative analysis over multiple species to reduce the false positive (FP) rate.

2.5.3. Other Probabilistic Tools

2.5.3.1. miRDeep – Marc R Friedländer et al. (22)

1. The miRDeep algorithm uses a probabilistic model of miRNA biogenesis to score compatibility of the position and frequency of sequenced RNA (Solexa sequencing as well as 454 deep sequencing) with the secondary structure of the miRNA precursor.
2. The authors demonstrate the accuracy and robustness of the tool using published *C. elegans* data and data generated by deep sequencing human and dog RNAs.
3. miRDeep reports altogether ~230 previously unannotated miRNAs, of which 4 novel *C. elegans* miRNAs have been validated by Northern blot analysis.

2.5.3.2. SVM Tools – (15–17, 33, 34)

SVMs are a set of related supervised learning methods used for classification and regression. Viewing input data as two sets of vectors in an *n*-dimensional space, an SVM will construct a separating hyperplane in that space, one which maximizes the *margin* between the two data sets.

Multiple studies have utilized SVMs to predict novel miRNA genes, some of which are listed below:

1. Xue et al. (17) proposed a set of novel features of local contiguous structure–sequence information for distinguishing the hairpins of real pre-miRNAs and 1,000 pseudo pre-miRNAs. Remarkably, the SVM classifier built on human data can also correctly identify up to 90% of the pre-miRNAs from other species, including plants and virus, without utilizing any comparative genomics information.
2. Sewer et al. (34) focused on genomic regions around already known miRNAs in order to incorporate the observation that miRNAs are occasionally found in clusters. Starting with the known human, mouse, and rat miRNAs, the authors scanned 20 kb of flanking genomic regions for the presence of putative precursor miRNAs. Each genome was analyzed separately, allowing the evaluation of the species-specific identity and genome organization of miRNA loci. Only cross-species comparisons were used to make conservative estimates of the number of novel miRNAs. This ab initio method predicted between 50 and 100 novel pre-miRNAs for each of the considered species. Around 30% of these miRNAs have already been experimentally verified in a large set of cloned mammalian small RNAs (24).

2.6. Methods Designed to Complement Other Tools

In addition to miRNA gene prediction tools, a number of methods have been developed to complement such tools and help to maximize their prediction accuracy.

2.6.1. RNAmicro (16)

RNAmicro is designed specifically to work as a "sub-screen" for large-scale ncRNA surveys with RNAz. RNAz is an SVM-based

classification tool which combines comparative sequence analysis and structure prediction (35). RNAmicro tries to provide an annotation of the RNAz survey data, in order to provide a more balanced trade-off between sensitivity and specificity. Thus, the goal of RNAmicro is a bit different from that of specific surveys for miRNAs in genomic sequences; in the latter case, one is interested in very high specificity so that the candidates selected for experimental verification contain as few FPs as possible.

RNAmicro is an SVM-based method which, in order to classify the miRNA precursors, evaluates the information contained in multiple sequence alignment.

RNAmicro consists of the following components:

1. A pre-processor that performs the following three actions:
   a. Identifies conserved hairpin-structure regions in a multiple sequence alignment. Extracted windows of length L are used in one-nucleotide steps from the input alignment.
   b. RNAalifold algorithm (Vienna RNA Package) is used to compute consensus sequence and structure (36).
   c. The consensus secondary structure which is obtained in "dot-parenthesis" notation is further analyzed.
2. A module that computes a vector of numerical descriptors from each hairpin structure.
3. A SVM that classifies the candidate based on its vector of descriptors.

2.6.2. Microprocessor SVM (15)

Another study describes a SVM classifier that can separate between true and false Drosha processing sites. The biological relevance of this tool is based on the assumption that mature miRNAs are processed from long hairpin transcripts by Drosha, and that this processing defines the mature product and is characteristic for all miRNA genes. This classifier can predict the exact location of 5′ microprocessor processing sites in human 5′-miRNAs with 50% accuracy. It is also important to mention that if the predicted site is wrong then the actual site is within two nucleotides of the predicted site, in about 90% of the cases. Even though Microprocessor SVM is not effective as a standalone tool, it can be useful as:

1. A post-processor for existing tools that only predict whether hairpins are likely miRNAs.
2. A complementary miRNA gene classifier that performs better than currently available methods for predicting unconserved miRNAs. As a consequence, other prediction tools can be improved upon post-processing using this method.

3. Examples of Computationally Predicted miRNAs Involved in Cancer

In order to stress the importance of miRNA gene prediction tools in identifying miRNA genes implicated with various types of cancers, it is necessary to provide representative examples.

1. These include miRNAs predicted by sequence homology to already cloned miRNAs, such as miR-143 (25) involved in colorectal cancer (4), miR-125b (lin-4) and miR-145 which are implicated in breast cancer (37), miR-106a which is believed to play a regulatory role in colon, pancreas, and prostate cancer (38), and miR-155 which is associated with HL, BCL, pediatric BL, breast, and lung cancer as well as poor survival (8, 39–42).
2. Another large study utilized conservation with mouse and Fugu rubripes sequences and the score given by the program MiRscan (28) in order to predict novel miRNAs associated with cancer. These included mir221/222 which are involved in Papillary thyroid carcinoma (43) and glioblastomas (44), hsa-mir-192 which is shown to have reduced expression in colorectal neoplasia (4), hsa-mir-196a-1 which was cloned from human osteoblast sarcoma cells (45), and hsa-mir-210 which is implicated in Kaposi's sarcoma-associated herpes virus infections (46). The majority of these miRNA genes were identified computationally and their implication in cancer confirmed by experimental methods.
3. In a recent large-scale bioinformatics study, Calin et al. (9) made use of the miRNA registry as well as bibliography for CAGRs, to show that over 80 known miRNAs reside within CAGR and fragile sites. This was one of the first large-scale computational studies to indicate a direct connection between genomic location of miRNAs and regions prone to genetic alteration in cancer. This study has provided the raw material for undergoing experiments aiming to verify these connections.
4. Recently, a sophisticated computational tool (SSCprofiler (21)) scanned over 70 CAGRs and resulted in the prediction of four novel candidate miRNAs, all of which were verified by tilling array data as well as Northern blot analysis.

4. Tool Criteria for Success

This section underlies the capabilities and limitations of the numerous available computational methods for miRNA gene prediction.

1. The general rule is that, the more biological information integrated in the computational tool the more successful it will be.
2. As indicated in Table 2, early tools making use of sequence and the closely related species conservation alone did not achieve very high prediction accuracy.
3. Subsequent methodologies taking into account additional biological information, such as structure and multiple species conservation for filtering their predictions, showed significant improvements.
4. Another important criterion for success is the simultaneous consideration of biological information. Considering all features at once, preferably at the nucleotide level, is more informative and more efficient than undertaking a pipeline approach, which utilizes the different features sequentially to predict novel miRNAs.
5. Sophisticated machine learning algorithms process all the biological features in parallel in order to build predictive models. This is a significant improvement to linear approaches adopted by initial brute-force methods (see Table 2).
6. Successful training of learning algorithms requires caution when selecting positive and negative training examples. Online databases may contain false-positives, and the definition of a negative miRNA is still uncertain.
7. The use of 3′ UTR regions to draw negative miRNA genes has been the norm in most studies mostly because there was no documented miRNA gene within these regions. However, recently a small percentage of miRNA genes have been shown to exist within 3′ UTRs.
8. Evaluating the performance of prediction tools by measuring sensitivity as well as specificity. However, these measures are greatly affected by the quality as well as the number of negative and positive samples (19). Hence, the prediction accuracy of one tool may change if the dataset from another study is used and vice versa. In general, computational miRNA gene prediction lacks a benchmark dataset.
9. There is still ample space for improvement in the field of computational prediction of miRNA genes, besides the great advances in the last years.
10. As more biological information regarding miRNA biogenesis and regulation is made available, computational tools incorporating this information will become much more effective.
11. Perhaps novel biological information will shed some light into one of the bottlenecks of in silico prediction of miRNAs, namely the identification of the mature miRNA sequence on the miRNA precursor.

Table 2
Comparison of miRNA identification methods

Features	Cloning	MiRscan	MiR seeker	Phylogenetic shadowing	Blatting	MiR Align	ProMir	Bayes-classifier	Xue	Sewer	RNA micro	SSC profiler	MiR Deep	MiR Rim	Microprocessor SVM
Sequence															
Directly[a]	x					x	x	x	x	x	x	x			x
Indirectly		x	x	x	x								x	x	
Structure															
Base pairing		x				x	x	x	x	x	x	x	x	x	x
Hairpin		x				x	x	x	x	x	x	x	x	x	x
Bulges/loops		x				x	x	x	x	x	x	x	x	x	x
Mature location		x					x					x			
Thermodynamic temp			x	x		x	x	x	x	x	x	x	x	x	x
Conservation															
Pairwise	x	x	x												
Conserved synteny					x										
Conserved clustering										x					x
Multiple species				x			x	x				x		x	

Methodology														
Brute-force	x	x												
Homology based				x										
SVM								x	x	x				x
Probablistics					x	x	x				x		x	
Complement other tools										x				x
Performance														
Sensitivity (%)	74	75				73	97	93.30	64	90	88.95		~70	90
Specificity (%)						96	91	88.10	64		84.16		~90	78
Species	Nematode +human	Droso-phila				Human	Mouse	Multi-species	Human, mouse, and rat	Human	Human	*Caenor-habditis elegans*, human, and dog	Human	Human
Use of large-scale, high-throughput data											x	x		
Experimental verification	x		x			x					x	x		
Reference														

[a]Directly – in the sense that the nucleotide distribution in the sequence is taken into consideration (i.e., GC content). Indirectly refers to the use of sequence to derive structure

12. It is also important to mention that miRNA gene prediction is currently addressed as a 2D problem. Secondary structure prediction algorithms only follow 2D rules and do not portray a complete tertiary picture. Tools capable of predicting tertiary structure of miRNAs (such as pseudoknots (47)) will transform a 2D problem into a 3D one, reflecting the conditions found in the cell more accurately.
13. As a final note, it is important to stress that the development of computational tools is tightly linked to biological research. The successful evolution of these tools demands that developers keep up with novel biological findings that may change the way information should be used. A characteristic example of this is the use of Drosha processing sites in the Microprocessor (15) study mentioned above. It was recently shown that intronic microRNA precursors may bypass Drosha processing (48).

Acknowledgments

This work was supported by the action 8.3.1 (Reinforcement Program of Human Research Manpower – "PENED 2003" (03ED842)) of the operational program "competitiveness" of the Greek General Secretariat for Research and Technology, a Marie Curie Fellowship of the European Commission (PIOF-GA-2008-219622), and the National Science Foundation (NSF 0515357).

References

1. Fantom, C. (2005) The transcriptional landscape of the mammalian genome, *Science 309*, 1559–1563.
2. Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., Harano, T., Yatabe, Y., Nagino, M., Nimura, Y., Mitsudomi, T., and Takahashi, T. (2004) Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival, *Cancer Res 64*, 3753–3756.
3. Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F., and Croce, C. M. (2002) Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia, *Proc Natl Acad Sci U S A 99*, 15524–15529.
4. Michael, M. Z., Susan M. O. C., van Holst Pellekaan, N. G., Young, G. P., and James, R. J. (2003) Reduced accumulation of specific microRNAs in colorectal neoplasia, *Mol Cancer Res 1*, 882–891.
5. Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., Yatabe, Y., Kawahara, K., Sekido, Y., and Takahashi, T. (2005) A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation, *Cancer Res 65*, 9628–9632.
6. He, L., Thomson, J. M., Hemann, M. T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S. W., Hannon, G. J., and Hammond, S. M. (2005) A microRNA polycistron as a potential human oncogene, *Nature 435*, 828–833.

7. Tagawa, H., and Seto, M. (2005) A microRNA cluster as a target of genomic amplification in malignant lymphoma, *Leukemia 19*, 2013–2016.
8. Metzler M. W. M., Busch K., Viehmann S., and Borkhardt A. (2003) High Expression of precursor microRNA-155/BIC RNA in children with Burkitt lymphoma, *Genes Chromosomes Cancer 2*, 167–169.
9. Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., and Croce, C. M. (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers, *PNAS 101*, 2999–3004.
10. Lai, E. C., Tomancak, P., Williams, R. W., and Rubin, G. M. (2003) Computational identification of Drosophila microRNA genes, *Genome Biol 4*, R42–R61.
11. Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., and Yekta, S. (2003) The microRNAs of *Caenorhabditis elegans*, *Genes Dev 16*, 991–1008.
12. Weber, M. J. (2005) New human and mouse microRNA genes found by homology search, *FEBS J 272*, 59–73.
13. Legendre, M., Lambert, A., and Gautheret, D. (2004) Profile-based detection of microRNA precursors in animal genomes, *Bioinformatics 21*, 841–845.
14. Wang, X., Zhang, J., Li, F., Gu, J., He, T., Zhang, X., and Li, Y. (2005) MicroRNA identification based on sequence and structure alignment, *Bioinformatics 21*, 3610–3614.
15. Helvik, S. A., Snove, O., Jr., and Saetrom, P. (2006) Reliable prediction of Drosha processing sites improves microRNA gene prediction, *Bioinformatics 23*, 142–149.
16. Hertel, J., and Stadler, P. F. (2006) Hairpins in a Haystack: recognizing microRNA precursors in comparative genomics data, *Bioinformatics 22*, e197–202.
17. Xue, C., Li, F., He, T., Liu, G. P., Li, Y., and Zhang, X. (2005) Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine, *BMC Bioinformatics 6*, 310–316.
18. Nam, J. W., Shin, K. R., Han, J., Lee, Y., Kim, V. N., and Zhang, B. T. (2005) Human microRNA prediction through a probabilistic co-learning model of sequence and structure, *Nucleic Acid Res 33*, 3570–3581.
19. Yousef, M., Nebozhyn, M., Shatkay, H., Kanterakis, S., Showe, L. C., and Showe, M. K. (2006) Combining multi-species genomic data for microRNA identification using a Naive Bayes classifier, *Bioinformatics 22*, 1325–1334.
20. Terai, G., Komori, T., Asai, K., and Kin, T. (2007) miRRim: a novel system to find conserved miRNAs with high sensitivity and specificity, *RNA 13*, 2081–2090.
21. Oulas, A., Boutla, A., Gkirtzou, K., Reczko, M., Kalantidis, K., and Poirazi, P. (2009) Prediction of novel microRNA genes in cancer-associated genomic regions – a combined computational and experimental approach, *Nucleic Acids Res 37*, 3276–3287.
22. Friedlander, M. R., Chen, W., Adamidi, C., Maaskola, J., Einspanier, R., Knespel, S., and Rajewsky, N. (2008) Discovering microRNAs from deep sequencing data using miRDeep, *Nat Biotechnol 26*, 407–415.
23. Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., Stadler, P. F., Hertel, J., Hackermuller, J., Hofacker, I. L., Bell, I., Cheung, E., Drenkow, J., Dumais, E., Patel, S., Helt, G., Ganesh, M., Ghosh, S., Piccolboni, A., Sementchenko, V., Tammana, H., and Gingeras, T. R. (2007) RNA maps reveal new RNA classes and a possible function for pervasive transcription, *Science 316*, 1484–1488.
24. Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A. O., Landthaler, M., Lin, C., Socci, N. D., Hermida, L., Fulci, V., Chiaretti, S., Foa, R., Schliwka, J., Fuchs, U., Novosel, A., Muller, R. U., Schermer, B., Bissels, U., Inman, J., Phan, Q., Chien, M., Weir, D. B., Choksi, R., De Vita, G., Frezzetti, D., Trompeter, H. I., Hornung, V., Teng, G., Hartmann, G., Palkovits, M., Di Lauro, R., Wernet, P., Macino, G., Rogler, C. E., Nagle, J. W., Ju, J., Papavasiliou, F. N., Benzing, T., Lichter, P., Tam, W., Brownstein, M. J., Bosio, A., Borkhardt, A., Russo, J. J., Sander, C., Zavolan, M., and Tuschl, T. (2007) A mammalian microRNA expression atlas based on small RNA library sequencing, *Cell 129*, 1401–1414.
25. Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., and Tuschl, T. (2002) Identification of tissue-specific microRNAs from mouse, *Curr Biol 12*, 735–739.
26. Boffelli, D., McAuliffe, J., Ovcharenko, D., Lewis, K. D., Ovcharenko, I., Pachter, L., and Rubin, E. M. (2003) Phylogenetic shadowing of primate sequences to find functional regions of the human genome, *Science (New York, NY) 299*, 1391–1394.

27. Berezikov, E., Guryev, V., van de Belt, J., Wienholds, E., Plasterk, R. H., and Cuppen, E. (2005) Phylogenetic shadowing and computational identification of human microRNA genes, *Cell 120*, 21–24.
28. Lim, L. P., Glasner, M. E., Yekta, S., Burge, C. B., and Bartel, D. P. (2003) Vertebrate microRNA genes, *Science 299*, 1540.
29. Artzi, S., Kiezun, A., and Shomron, N. (2008) miRNAminer: a tool for homologous microRNA gene search, *BMC Bioinformatics 9*, 39.
30. Sunkar, R., and Jagadeeswaran, G. (2008) In silico identification of conserved microRNAs in large number of diverse plant species, *BMC Plant Biol 8*, 37.
31. Kent, W. J. (2002) BLAT – the BLAST-like alignment tool, *Genome Res 12*, 656–664.
32. Lambert, A., Fontaine, J. F., Legendre, M., Leclerc, F., Permal, E., Major, F., Putzer, H., Delfour, O., Michot, B., and Gautheret, D. (2004) The ERPIN server: an interface to profile-based RNA motif identification, *Nucleic Acids Res 32*, W160–W165.
33. Buck, A. H., Santoyo-Lopez, J., Robertson, K. A., Kumar, D. S., Reczko, M., and Ghazal, P. (2007) Discrete clusters of virus-encoded microRNAs are associated with complementary strands of the genome and the 7.2-kilobase stable intron in murine cytomegalovirus, *J Virol 81*, 13761–13770.
34. Sewer, A., Paul, N., Landgraf, P., Aravin, A., Pfeffer, S., Brownstein, M. J., Tuschl, T., van Nimwegen, E., and Zavolan, M. (2005) Identification of clustered microRNAs using an ab initio prediction method, *BMC Bioinformatics 6*, 267–281.
35. Washietl, S., Hofacker, I. L., and Stadler, P. F. (2005) Fast and reliable prediction of noncoding RNAs, *Proc Natl Acad Sci U S A 102*, 2454–2459.
36. Hofacker, I. L. (2003) Vienna RNA secondary structure server, *Nucleic Acids Res 31*, 3429–3431.
37. Iorio, M. V., Ferracin, M., Liu, C. G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M., Menard, S., Palazzo, J. P., Rosenberg, A., Musiani, P., Volinia, S., Nenci, I., Calin, G. A., Querzoli, P., Negrini, M., and Croce, C. M. (2005) MicroRNA gene expression deregulation in human breast cancer, *Cancer Res 65*, 7065–7070.
38. Volinia, S., Calin, G. A., Liu, C. G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., Prueitt, R. L., Yanaihara, N., Lanza, G., Scarpa, A., Vecchione, A., Negrini, M., Harris, C. C., and Croce, C. M. (2006) A microRNA expression signature of human solid tumors defines cancer gene targets, *Proc Natl Acad Sci USA 103*, 2257–2261.
39. Eis, P. S., Tam, W., Sun, L., Chadburn, A., Li, Z., Gomez, M. F., Lund, E., and Dahlberg, J. E. (2005) Accumulation of miR-155 and BIC RNA in human B cell lymphomas, *Proc Natl Acad Sci U S A 102*, 3627–3632.
40. Kluiver, J., Haralambieva, E., de Jong, D., Blokzijl, T., Jacobs, S., Kroesen, B. J., Poppema, S., and van den Berg, A. (2006) Lack of BIC and microRNA miR-155 expression in primary cases of Burkitt lymphoma, *Genes Chromosomes Cancer 45*, 147–153.
41. Kluiver, J., Poppema, S., de Jong, D., Blokzijl, T., Harms, G., Jacobs, S., Kroesen, B. J., and van den Berg, A. (2005) BIC and miR-155 are highly expressed in Hodgkin, primary mediastinal and diffuse large B cell lymphomas, *J Pathol 207*, 243–249.
42. Yanaihara, N., Caplen, N., Bowman, E., Seike, M., Kumamoto, K., Yi, M., Stephens, R. M., Okamoto, A., Yokota, J., Tanaka, T., Calin, G. A., Liu, C. G., Croce, C. M., and Harris, C. C. (2006) Unique microRNA molecular profiles in lung cancer diagnosis and prognosis, *Cancer Cell 9*, 189–198.
43. He, H., Jazdzewski, K., Li, W., Liyanarachchi, S., Nagy, R., Volinia, S., Calin, G. A., Liu, C. G., Franssila, K., Suster, S., Kloos, R. T., Croce, C. M., and de la Chapelle, A. (2005) The role of microRNA genes in papillary thyroid carcinoma, *Proc Natl Acad Sci USA 102*, 19075–19080.
44. Ciafre, S. A., Galardi, S., Mangiola, A., Ferracin, M., Liu, C. G., Sabatino, G., Negrini, M., Maira, G., Croce, C. M., and Farace, M. G. (2005) Extensive modulation of a set of microRNAs in primary glioblastoma, *Biochem Biophys Res Commun 334*, 1351–1358.
45. Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A., and Tuschl, T. (2003) New microRNAs from mouse and human, *RNA 9*, 175–179.
46. Cai, X., Lu, S., Zhang, Z., Gonzalez, C. M., Damania, B., and Cullen, B. R. (2005) Kaposi's sarcoma-associated herpesvirus expresses an array of viral microRNAs in latently infected cells, *Proc Natl Acad Sci USA 102*, 5570–5575.
47. Reeder, J., and Giegerich, R. (2004) Design, implementation and evaluation of a practical pseudoknot folding algorithm based on thermodynamics, *BMC Bioinformatics 5*, 104–115.

48. Ruby, J. G., Jan, C. H., and Bartel, D. P. (2007) Intronic microRNA precursors that bypass Drosha processing, *Nature 448*, 83–86.
49. Cimmino, A., Calin, G. A., Fabbri, M., Iorio, M. V., Ferracin, M., Shimizu, M., Wojcik, S. E., Aqeilan, R. I., Zupo, S., Dono, M., Rassenti, L., Alder, H., Volinia, S., Liu, C. G., Kipps, T. J., Negrini, M., and Croce, C. M. (2005) miR-15 and miR-16 induce apoptosis by targeting BCL2, *Proc Natl Acad Sci USA 102*, 13944–13949.

Chapter 3

The Principles of MiRNA-Masking Antisense Oligonucleotides Technology

Zhiguo Wang

Abstract

MiRNA-masking antisense oligonucleotides technology (miR-mask) is an anti-*microRNA* antisense oligodeoxyribonucleotide (AMO) approach of a different sort. A standard miR-mask is single-stranded 2′-*O*-methyl-modified oligoribonucleotide (or other chemically modified) that is a 22-nt antisense to a protein-coding mRNA as a target for an endogenous miRNA of interest. Instead of binding to the target miRNA like an AMO, an miR-mask does not directly interact with its target miRNA but binds to the binding site of that miRNA in the 3′ UTR of the target mRNA by fully complementary mechanism. In this way, the miR-mask covers up the access of its target miRNA to the binding site so as to derepress its target gene (mRNA) via blocking the action of its target miRNA. The anti-miRNA action of an miR-mask is gene-specific because it is designed to be fully complementary to the target mRNA sequence of an miRNA. The anti-miRNA action of an miR-mask is miRNA-specific as well because it is designed to target the binding site of that particular miRNA. The miR-mask approach is a valuable supplement to the AMO technique; while AMO is indispensable for studying the overall function of an miRNA, the miR-mask might be more appropriate for studying the specific outcome of regulation of the target gene by the miRNA. This technology was first established by my research group in 2007 (Xiao et al., J Cell Physiol 212:285–292; Wang et al., J Mol Med 86:772–783, 2008) and a similar approach with the same concept was subsequently reported by Schier's laboratory (Choi et al., Science 318:271–274, 2007).

Key words: MiRNAs, Gene expression, MiRNA-masking antisense oligonucleotides, Anti-miRNA antisense

1. Introduction

Each single miRNA may regulate as many as 1,000 protein-coding genes and each gene may be regulated by multiple miRNAs (1, 2). This implies that the binding action of miRNAs is sequence-specific but not gene-specific; any genes carrying the binding motifs for a particular miRNA may be the targets for this miRNA. Similarly, the action of anti-miRNA antisense (AMO) (3–7) is not gene-specific

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_3, © Springer Science+Business Media, LLC 2011

either, but miRNA-specific. These properties of miRNAs and AMOs may present the obstacles in miRNA research and development as therapeutic agents since they may elicit unwanted side effects and toxicity through their nongene-specific functional profiles. For example, knockdown of *miR-21* in cultured glioblastoma cells resulted in a significant drop in cell number. This reduction was accompanied by increases in caspase-3 and -7 enzymatic activities and TUNEL staining (8, 9). Similarly, in MCF-7 human breast cancer cells, *miR-21* also elicits antiapoptotic effects (10, 11). In HeLa cells, however, *miR-21* does the opposite; the inhibition of *miR-21* increased the number of surviving cells (12). Moreover, the inhibition of *miR-21* in A549 human lung cancer cells fails to alter cell death or growth (12). Evidently, a single miRNA can have three different actions – antiapoptotic, proapoptotic, or neutral in different cell types. Ji et al. (13) found that *miR-21* inhibition upregulates and *miR-21* overexpression downregulates the expression of PTEN protein in vascular smooth muscle cells (VSMCs). They further revealed that the effects of *miR-21* on its downstream Akt that mediates survival signal in a cell, with *miR-21* inhibition downregulating and overexpression upregulating Akt protein level, are consistent with the expression changes of PTEN. In contrast to PTEN, *miR-21* knockdown decreases and overexpression increases the expression of antiapoptotic Bcl-2 protein. The authors suggested that Bcl-2 might be an indirect target of *miR-21* in VSMCs by suppressing expression of a gene that negatively regulates Bcl-2 expression or that *miR-21* might be able to directly affect Bcl-2 expression via binding to the sequence outside the 3′ UTR (13). Similar upregulation of Bcl-2 by *miR-21* was also observed in MCF-7 breast cancer cells (10). Regulation of PTEN/Akt and Bcl-2 by *miR-21* is in agreement with its cytoprotective ability against apoptosis. Two most recent studies revealed that the tumor suppressor protein programmed cell death 4 (PDCD4) is a functionally important target for *miR-21* in breast cancer cells (14) and in colorectal cancer cells (15). Based on these targeting mechanisms, miR-21 is expected to affect tumorigenesis. Taking the concept of "miRNA as a regulator of a cellular function", one can just focus on what miR-21 does on cancer progression. However, if one wants to understand the mechanisms using miR-21 loss-of-function strategy, the AMO technology will fall short due to their lack of gene specificity. By knocking down miR-21, one will potentially alter the expression of all miR-21 target genes mentioned above and it would be hard to predict whether miR-21's action would be antiapoptotic, proapoptotic, or neutral.

To emasculate the problem, we have developed the miRNA-masking antisense oligonucleotides technology (miR-mask) that provides a gene-specific strategy of miRNA loss-of-function for studying miRNA function and mechanisms. Using the miR-mask approach, one is now able to dissect the role of each of the target

genes of an miRNA, say the survival and apoptotic genes for miR-21. For instance, one can use an miR-mask on PTEN specifically to explore the role of miR-21 regulation of cell death. Soon after our publication of this technology, Choi et al. (16) also published a study using essentially the same strategy and named the technology "target protector". For convenience and clarity, I suggest using miR-mask as a unified name.

The miR-mask technology is an alternative to the AMO approach. But unlike AMO, which acts in a nongene-specific manner, the miR-mask finds its particular value in targeting miRNA in a gene-specific fashion. It is particularly useful when inhibiting miRNA action on a specific protein-coding gene without affecting the level of this miRNA and its silencing effects on other genes are required.

We have validated the miR-mask technology by testing its application to the cardiac pacemaker channel-encoding genes *HCN2* and *HCN4* (17). We created miR-masks that are able to bind to HCN2 and HCN4 and prevent the repressive actions of miR-1 and miR-133. These miR-masks resulted in enhanced protein expression of the pacemaker channels and increased pacemaker activities revealed by whole-cell patch-clamp recordings. Functionally, the miR-masks for HCN channels cause acceleration of heart rate in rats, simulating "biological pacemakers" (17).

This technology has also been validated by a recent study in which the authors investigated the role of zebrafish miR-430 in regulating expression of TGF-β nodal agonist squint and antagonist lefty, the key regulators of mesendoderm induction and left–right axis formation (16). They designed miR-masks, which they called target protector morpholinos, complementary to miRNA binding sites in target mRNAs in order to disrupt the interaction of specific miRNA–mRNA pairs. Protection of squint or lefty mRNAs from miR-430 resulted in enhanced or reduced nodal signaling.

2. The Advantage of the miR-Mask Technology

The miR-mask strategy was developed to interfere with function of the endogenous miRNAs in a gene-specific and miRNA-specific manner. The idea behind that is to use an miR-mask to regulate the protein expression of the target gene (mRNA) by interfering with the action of a particular miRNA on this gene. Two prerequisites must be fulfilled: (1) the presence of a recognition motif for an miRNA within the 3′ UTR of the target gene and (2) the presence of a fragment with unique sequences containing the miRNA-binding motif which is sufficiently long (~22 nts) for miR-mask binding. The first prerequisite ensures the miRNA specificity of miR-mask action and the second the gene specificity of miR-mask action.

An miRNA-mask is designed to fully base pair with the binding motif of an endogenous miRNA in the 3′ UTR of the target mRNA. Upon delivery into the cell, the miR-mask is expected to bind itself to the region so as to block the access of that endogenous miRNA to the site of action. In this way, the miR-mask disrupts miRNA:mRNA interaction to relieve the repressive action of the miRNA on the target gene to promote the protein expression of that gene.

The miR-mask approach differs from the AMO approach in several aspects, despite that they both can result in enhancement of gene expression by removing the repressive effects of a particular miRNA on protein translation of the target mRNA.

1. An AMO is designed to entirely base pair with the sequence of the target miRNA, whereas an miR-mask is designed based on the sequence of the target site for an miRNA in the 3′ UTR of the target mRNA. In other words, miR-mask acts like a protector of the gene from being inhibited by miRNA.
2. An AMO interacts with (binds to) its target miRNA and may well cause degradation of that miRNA such that all functions of that miRNA are deemed to be eliminated, whereas an miR-mask interacts with (binds to) its target mRNA and does not induce miRNA degradation such that the function of that miRNA on other genes is intact. In this sense, an miR-mask is not only a target protector but also an miRNA protector.
3. The action of an AMO is miRNA-specific but not gene-specific and it may well induce enhancement of expression of multiple genes regulated by the same target miRNA, whereas an miR-mask is expected to be gene-specific because it is fully complementary to the target mRNA sequence and is miRNA-specific as well because it is designed to target the binding site of that miRNA. Hence, an miR-mask is expected to derepress only the target gene.
4. An AMO acts to disrupt miRNA:mRNA interaction by creating an AMO:miRNA interaction, whereas an miR-mask disrupts miRNA:mRNA interaction by creating an ASO:mRNA interaction (ASO: anti-mRNA antisense oligomer).

The miR-mask approach also differs from the conventional antisense technique in the following two aspects, despite that they both are entirely complementary to the target sequences.

1. A conventional antisense oligodeoxynucleotide (ASO) in theory can be designed to target any part of the protein-coding region of a gene (though the sequences from the translation start codon are frequently used), whereas an miR-mask is limited to the target site of an miRNA in the 3′ UTR of a protein-coding gene.

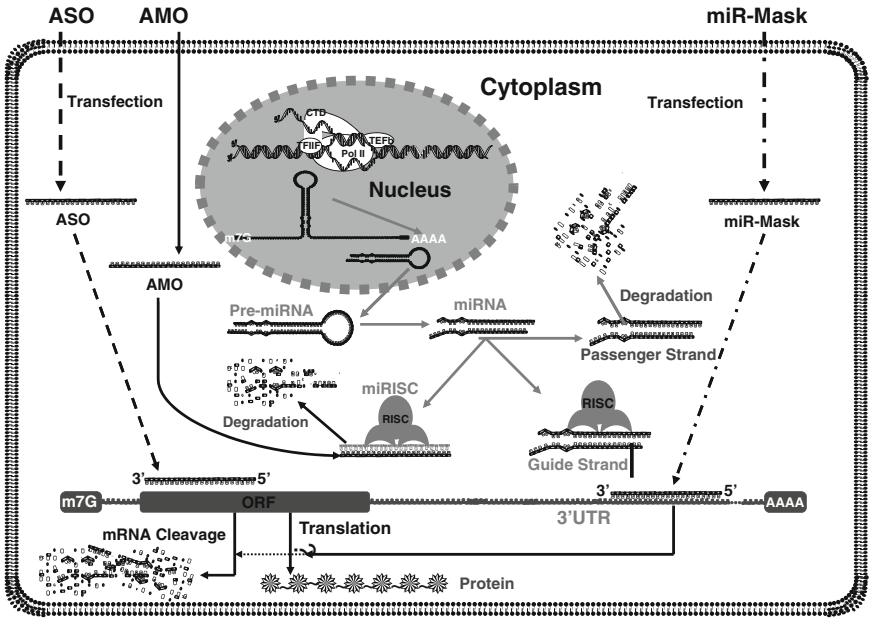


Fig. 1. Schematic presentation of actions of miRNA-masking antisense oligonucleotide (miR-mask) compared with the conventional antisense oligodeoxynucleotide (ASO) and anti-miRNA antisense inhibitor oligonucleotide (AMO) technologies. Synthetic nucleic acids are introduced into the cells. ASO binds to the coding region of the target mRNA and hinders the translation process; AMOs bind to the target miRNA, resulting in miRNA cleavage; miR-masks bind to the binding site of miRNAs in 3′ UTR of the target mRNA and prevent miRNAs from binding to the target mRNA, leading to a relief of translational repression without affecting miRNAs.

2. The efficacy of an miR-mask on gene expression depends on the basal activity of the endogenous miRNA on the target mRNA, while that of a conventional ASO depends on the interaction between the ASO and the target gene.
3. A conventional antisense ODN binds to its target site in the coding region of a gene and hinders the protein translation process. Conversely, an miR-mask binds to the 3′ UTR and masks the target site of an miRNA so as to block the action of the endogenous miRNA and enhance protein translation. Thus, the two techniques produce exactly opposite outcomes: one inhibits but the other enhances gene expression.

A comparison of the miR-mask, the AMO, and the conventional antisense ODN techniques is summarized in Fig. 1.

3. Designing and Validating miR-Masks

1. Analyze sequences to identify the binding site(s) for an endogenous miRNA of interest, in the 3′ UTR of the target mRNA. For example, we have shown that the muscle-specific

miRNA miR-133 represses the protein expression of HERG K^+ channel gene *KCNH2*, contributing to the increased risk of pathologic long QT syndrome in diabetic cardiomyopathy (17). We did a study to see if we could relieve the repression to reduce the arrhythmogenic potential in diabetic hearts. For this end, the first step is to analyze the 3′ UTR of KCNH2 around the region containing the binding sequence for miR-133.

2. Design an oligonucleotide fragment of around 22 nts or longer fully antisense to the region covering the binding sequence of the miRNA of interest.
3. Blast search to verify the uniqueness of the fragment to ensure gene specificity. Once confirmed, the fragment is considered an miR-mask.
4. Chemically synthesize the miR-mask using services provided by commercial companies such as IDT Technologies or Ambion. Remember to chemically modify the oligonucleotide.
5. Transfect the miR-mask into cells to study the enhancing effects on protein expression of the target gene. In the case of an miR-mask for KCNH2 and miR-133, we expect to see an increase in the HERG protein level.

Like miR-Mimics (*see* Chapter 15), the functional activities of miR-masks must be verified with several approaches including luciferase reporter gene assay, western blot analysis, qRT–PCR quantification, and functional assays.

4. Summary

The miR-mask approach is a valuable supplement to the AMO technique; while AMO is indispensable for studying the overall function of an miRNA, the miR-mask might be more appropriate for studying the specific outcome of regulation of the target gene by the miRNA. (1) The major advantage of this technology is that it offers a gene-specific miRNA-interfering strategy, which in many situations is highly desirable. (2) This expression-enhancing action of miR-mask is unique and could have many applications. (3) The characteristic dual specificities (miRNA specificity and gene specificity) of miR-mask may be particularly useful for the miRNA:mRNA interactions consequent to polymorphisms in the protein-coding genes that create new binding sites for miRNAs.

Acknowledgments

This work was supported in part by the Canadian Institute of Health Research, Heart and Stroke Foundation of Quebec, and Fonds de la Recherche de l'Institut de Cardiologie de Montreal. Dr. Z. Wang is a Changjiang Scholar Professor of the Ministry of Education of China and a Longjiang Scholar Professor of Heilongjiang, China. The authors thank XiaoFan Yang for her excellent technical support.

References

1. Stenvang J, Kauppinen S (2008) MicroRNAs as targets for antisense-based therapeutics. Expert Opin Biol Ther 8:59–81.
2. Watanabe Y, Tomita M, Kanai A (2007) Computational methods for microRNA target prediction. Meth Enzymol 427:65–86.
3. Hutvágner G, Simard MJ, Mello CC, Zam PD (2004) Sequence-specific inhibition of small RNA function. PLoS Biol 2:465–475.
4. Meister G, Landthaler M, Dorsett Y, Tuschl T (2004) Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. RNA 10:544–550.
5. Krutzfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, Manoharan M, Stoffel M (2005) Silencing of microRNAs in vivo with "antagomirs". Nature 438:685–689.
6. Lu Y, Xiao J, Lin H, Bai Y, Luo X, Wang Z, Yang B (2009) Complex antisense inhibitors offer a superior approach for microRNA research and therapy. Nucleic Acids Res 37:e24–e33.
7. Ebert MS, Neilson JR, Sharp PA (2007) MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. Nat Methods 4:721–726.
8. Chan JA, Krichevsky AM, Kosik KS (2005) MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. Cancer Res 65:6029–6033.
9. Corsten MF, Miranda R, Kasmieh R, Krichevsky AM, Weissleder R, Shah K (2007) MicroRNA-21 knockdown disrupts glioma growth in vivo and displays synergistic cytotoxicity with neural precursor cell delivered S-TRAIL in human gliomas. Cancer Res 67:8994–9000.
10. Si ML, Zhu S, Wu H, Lu Z, Wu F, Mo YY (2007) miR-21-mediated tumor growth. Oncogene 26:2799–1803.
11. Zhu S, Si ML, Wu H, Mo YY (2007) MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1). J Biol Chem 282:14328–14336.
12. Cheng AM, Byrom MW, Shelton J, Ford LP (2005) Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. Nucleic Acids Res 33:1290–1297.
13. Ji R, Cheng Y, Yue J, Yang J, Liu X, Chen H, Dean DB, Zhang C (2007) MicroRNA expression signature and antisense-mediated depletion reveal an essential role of microRNA in vascular neointimal lesion formation. Circ Res 100:1579–1588.
14. Asangani IA, Rasheed SA, Nikolova DA, Leupold JH, Colburn NH, Post S, Allgayer H (2008) MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. Oncogene 27:2128–2136.
15. Frankel LB, Christoffersen NR, Jacobsen A, Lindow M, Krogh A, Lund AH (2008) Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells. J Biol Chem 283:1026–1033.
16. Choi WY, Giraldez AJ, Schier AF (2007) Target protectors reveal dampening and balancing of nodal agonist and antagonist by miR-430. Science 318:271–274.
17. Xiao J, Yang B, Lin H, Lu Y, Luo X, Wang Z (2007) Novel approaches for gene-specific interference via manipulating actions of microRNAs: examination on the pacemaker channel genes *HCN2* and *HCN4*. J Cell Physiol 212:285–292.
18. Wang Z, Luo X, Lu Y, Yang B (2008) MiRNAs at the heart of the matter. J Mol Med 86:772–783.

Chapter 4

The Concept of Multiple-Target Anti-miRNA Antisense Oligonucleotide Technology

Zhiguo Wang

Abstract

The multiple-target AMO technology or MT-AMO technology is an innovative strategy, which confers on a single AMO fragment the capability of targeting multiple miRNAs. This modified AMO is single-stranded 2′-*O*-methyl-modified oligoribonucleotides carrying multiple AMO units, which are engineered into a single unit and are able to simultaneously silence multiple-target miRNAs or multiple miRNA seed families. Studies suggest that the MT-AMO is an improved approach for miRNA target finding and miRNA function validation; it not only enhances the effectiveness of targeting miRNAs but also confers diversity of actions. It has been successfully used to identify target genes and cellular function of several oncogenic miRNAs and of the muscle-specific miRNAs (Lu et al., Nucleic Acids Res 37:e24–e33, 2009). This novel strategy may find its broad application as a useful tool in miRNA research for exploring biological processes involving multiple miRNAs and multiple genes, and the potential as an miRNA therapy for human disease such as cancer and cardiac disorders. This technology was developed by my research laboratory in collaboration with Yang's group (Lu et al., Nucleic Acids Res 37:e24–e33, 2009), and it is similar but distinct from the miRNA Sponge technology developed by Sharp's laboratory in 2007 (Ebert et al., Nat Methods 4:721–726, 2007) and modified by Gentner et al. (Nat Methods 6:63–66, 2009).

Key words: miRNAs, Multiple-target AMO technology, AMO, One-drug, multiple-targets

1. Introduction

One of the indispensable approaches in miRNA research as well as in miRNA therapy is to inhibit miRNAs to achieve miRNA-loss-of-function outcomes. The base pair interaction between miRNAs and mRNAs is essential for the function of miRNAs; therefore, the most logical approach of silencing miRNAs is to use a nucleic acid that is antisense to the miRNA (1–5). These anti-miRNA oligonucleotides (AMOs) specifically and stoichiometrically bind, and efficiently and irreversibly silence,

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_4, © Springer Science+Business Media, LLC 2011

their target miRNAs. The AMO approach has been used in numerous studies to identify many cellular functions of miRNAs and has been considered a plausible strategy for miRNA therapy of human disease (1–5).

However, it has become clear that a particular condition may be associated with multiple miRNAs, and a given gene may be regulated by multiple miRNAs. For example, a study directed to the human heart identified 67 significantly upregulated miRNAs and 43 significantly downregulated miRNAs in failing left ventricles vs. normal hearts (6). No less than five different miRNAs have been shown to critically involve in cardiac hypertrophy (7–11). Similarly, Volinia et al. conducted a large-scale miRNA analysis on 540 solid tumor samples, including lung, breast, stomach, prostate, colon, and pancreatic tumors (12). Their survey revealed 15 miRNAs upregulated and 12 downregulated in human breast cancer tissues. Similar changes of multiple miRNAs were also found in other five solid tumor types. Many of these miRNAs have been reported to regulate cell proliferation or apoptosis, and some of them have been considered oncogenic miRNAs or tumor-suppressor miRNAs. These observations have led many in the field to speculate that multiple miRNAs work in a combinatorial fashion and act in concert to target a single transcript (13). Moreover, many miRNAs are members of families that share a seed sequence, but may have one or more nucleotide changes in the remaining sequence. These related miRNAs are expected to regulate similar target mRNAs, and if coexpressed in a tissue it may be necessary to inhibit all of them at once to observe phenotypic effects.

These properties of miRNA regulation may well create some uncertainties of outcomes by applying the AMO technology to silence miRNAs, since knocking down a single miRNA may not be sufficient, and certainly not optimal, to achieve the expected interference of cellular process and gene expression, which are regulated by multiple miRNAs. These facts also prompted us to raise several pertinent questions. Whether targeting a single miRNA is adequate for tackling a pertinent pathological condition? Whether simultaneously targeting multiple miRNAs relevant to a particular condition offers an improved approach than targeting a single miRNA using the regular AMO techniques? How can we concomitantly silence multiple miRNAs to achieve an interference of a cellular function?

Tools that allow multi-miRNA knockdown will be essential for the identification and validation of miRNA targets. For such applications, cotransfection of multiple AMOs targeting various isoforms is possible (14). In this respect, genetic approaches are superior for studying individual miRNA family members, whereas miRNA sponges (15) or multiple AMOs are appropriate for studying miRNA families whose members only contain a common seed sequence; single AMO studies may simplify the study of nearly identical miRNA paralogs. Combinations of

AMOs targeting unrelated miRNAs have also been used to disrupt more than one miRNA in the same transfected cells, obviating the need to make and combine multiple genetic knockouts. Because combinatorial control of targets by miRNAs may be common (16), this approach may prove particularly important for uncovering networks of miRNAs that act together. Vermeulen et al. (17) showed that cotransfection of a six AMO mixture can effectively derepress reporters for each individual miRNA. Functional studies also suggest that cotransfection of AMOs is effective. For instance, Pedersen et al. (18) tested that the efficacy of five interferon-β-induced miRNAs with seed matches to Hepatitis C genes in antiviral activity; indeed, simultaneous cotransfection of all five AMOs, but not controls, significantly enhanced the Hepatitis C RNA production.

Coapplication of multiple AMOs, while effective in some cases, may be problematic in that control of equal transfection efficiency is difficult if not impossible. To tackle the problem, an innovative strategy, the multiple-target AMO technology or MT-AMO technology, which confers a single AMO fragment the capability of targeting multiple miRNAs has been developed in my laboratory (19), which follows our original "single-agent, multiple-targets" theory (20). This modified AMO carries multiple antisense units, which are engineered into a single unit that is able to simultaneously silence multiple-target miRNAs. Studies suggest the MT-AMO is an improved approach for miRNA target gene finding and for studying functions of miRNAs. This novel strategy may find its broad application as a useful tool in miRNA research for exploring biological processes involving multiple miRNAs and multiple genes, and the potential as an miRNA therapy for human disease such as cancer and cardiac disorders.

We have validated the technique with two separate MT-AMOs: anti-miR-21/anti-miR-155/anti-miR-17-5p (MT-$AMO_{21/155/17}$) and anti-miR-1/anti-miR-133 (MT-$AMO_{1/133}$). miR-21, miR-155, and miR-17-5p are proven oncogenic miRNAs overexpressing in several solid cancers (12, 21), and miR-1 and miR-133 are muscle-specific miRNAs crucial for myogenesis. We first tested the ability of MT-$AMO_{21/155/17}$ and MT-$AMO_{1/133}$ to inactivate their respective target miRNAs by luciferase reporter assays. We then evaluated the effects of the two MT-AMOs on the protein levels of predicted target genes tumor suppressor genes TGFBI, APC, and BCL2L11 for miR-21, miR-155, and miR-17-5p, respectively, and HCN2 (a subunit of pacemaker channel) and Cav1.2 (α-subunit of L-type Ca^{2+} channel encoded by CACNA1C) for miR-1 and miR-133, respectively. We further demonstrated that MT-$AMO_{21/155/17}$ induced MCF-7 human breast cancer cell death with a greater efficacy and potency than the regular singular AMOs.

Most human diseases are multifactorial and multistep processes. Targeting a single factor (molecule) may not be adequate and certainly not optimal in cancer therapy, because single agent is

limited by incomplete efficacy and dose-limiting adverse effects. If related factors are concomitantly attacked, better outcomes are expected, and the current combination pharmacotherapy was developed for this reason: a combination of two or more drugs or therapeutic agents given as a single treatment that successfully saves lives. The "drug-cocktail" therapy of AIDS is one example of such a strategy (22), and similar approaches have been used for a variety of other diseases, including cancers (23–27). However, the current drug-cocktail therapy is costly and may involve complicated treatment regimen, undesired drug–drug interactions, and increased side effects (23). There is a need to develop a strategy to avoid these problems, and our "one-drug, multiple-target" strategy is highly desirable (20). However, it is nearly impossible to confer the ability of single compound to act on multiple-target molecules on the traditional pharmaceutical approaches or the currently known antigene strategies.

The MT-AMO technology was developed to enable an AMO to target multiple miRNAs in order to effectively interfere with expression of a protein-coding gene that is regulated by these multiple miRNAs so as to effectively alter a relevant cellular function and physiological process. In theory, it mimics the well-known drug-cocktail therapy. But it is devoid of the weaknesses of the drug-cocktail therapy, involving complicated treatment regimens, undesirable drug–drug interactions, and increased side effects. The MT-AMO technology offers resourceful combinations of varying AMOs for concomitantly targeting multiple miRNAs for studying or treating biological and pathophysiological processes involving multiple factors.

2. Designing MT-AMOs

(a) Select a particular gene, a particular cellular function, or a particular disease for your study. Then determine, based on published studies, the miRNAs that can target the gene under test or that are known to be associated with or implicated in this gene or cellular function or disease (Fig. 1);

(b) Design anti-miRNA oligonucleotides fragments (AMOs) exactly antisense to the selected miRNAs for your study. Then link these AMOs in the orientation of 5′-end to 3′-end together to form a long MT-AMO. Note that an AMO can be a RNA or a DNA (ODN, oligodeoxynucleotide); for commercial synthesis, an ODN is less costly than a RNA fragment. For example, we have tested an MT-AMO designed to integrate the AMOs against *miR-21*, *miR-155*, and *miR-17-5p* into one AMO ($MT\text{-}AMO_{21/155/17}$);

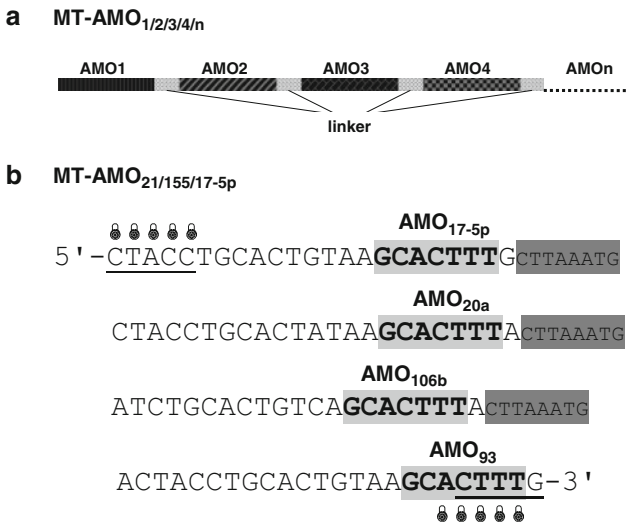


Fig. 1. Design of multiple-target anti-miRNA antisense oligomers (MT-AMO). (**a**) Schematic illustration of an MT-AMO. An MT-AMO incorporates multiple AMO units targeting different miRNAs. The number of AMO units in an MT-AMO is in theory unlimited; but longer sequence of an MT-AMO may create secondary structure and may also increase the difficulty of being up-taken into a cell. (**b**) Example of an MT-AMO designed to carry three AMOs targeting miR-21, miR-155, and miR-17-5p used in our previous study (19). The sequences corresponding to the seed sites of the miRNAs are highlighted in *light gray*. To enhance the stability and affinity, MT-AMO is chemically modified to have 5 nts at both ends locked with methylene bridges (LNA). An 8-nts linker (highlighted in *dark gray*) is inserted to connect the two adjacent AMO units.

(c) A linker (CTTAAATG) may be inserted into between every two AMOs to connect the two adjacent antisense units. This design gives you an MT-AMO that contains all the AMOs needed to knockdown the miRNAs relevant to the gene or cellular function or disease of your interest.

(d) Synthesize the designed MT-AMO fragment with chemical modifications

(e) Store the MT-AMO construct at −80°C for future use;

(f) Construct a negative control MT-AMO (NC MT-AMO) for verifying the effects and specificity of the effects of the MT-AMO. This NC MT-AMO should be designed based on the sequence of the MT-AMO; simply modifying the MT-AMO to contain **~5 nts mismatches at the 5′-end "seed site" to make it expectedly able to destruct the binding to the target miRNAs.

3. Summary

The MT-AMO technology is a simple and efficient approach to study a complicated, multifactorial cellular process that is regulated by multiple miRNAs. When containing AMO units towards

miRNA seed families, MT-AMOs are capable of knocking down members of multiple miRNA seed families.

However, many important issues remain unresolved in validating the MT-AMO technology as a gene therapy strategy. The optimal combination of targets for an MT-AMO remains unknown. In the example study described above, we tested "three-in-one" MT-AMOs. In theory, "*N*-in-one" MT-AMOs (*N* could be any number of AMO units) can be designed to include more relevant target TFs; however, larger MT-AMOs may hinder their penetration into the cells and compromise the effectiveness. More rigorous studies are warranted to define the optimal combination of length and accessibility of MT-AMOs to optimize desired effectiveness. Our work did not allow us to draw any conclusions as to what the optimal organization is for multiple AMO units to be placed in a single MT-AMO molecule. Moreover, efficient delivery of MT-AMOs into a cell is another challenge using MT-AMOs as therapeutic agents, as in other nucleotide-based technologies such as siRNA, antisense, ribozyme, aptamers, etc. Still another difficulty is to maintain an effective concentration of MT-AMO within a cell for a sufficient period of time. At present, investigation on modifications of MT-AMOs to enhance efficiency of transfection and to strengthen the stability within a cell so as to prolong the duration of actions is an active field of research. Constructing MT-AMO into virus vectors, such as adenovirus, lentivirus, etc., might be a reasonable approach to offset the weakness of the nucleotide technologies.

Acknowledgments

This work was supported in part by the Canadian Institute of Health Research, Heart and Stroke Foundation of Quebec and Fonds de la Recherche de l'Institut de Cardiologie de Montreal. Dr. Z. Wang is a Changjiang Scholar Professor of the Ministry of Education of China and a Longjiang Scholar Professor of Heilongjiang, China. The authors thank XiaoFan Yang for her excellent technical supports.

References

1. Krutzfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, Manoharan M, Stoffel M (2005) Silencing of microRNAs in vivo with 'antagomirs'. Nature 438:685–689.
2. Hammond SM (2006) MicroRNA therapeutics: a new niche for antisense nucleic acids. TiMM 12:99–101.
3. Cheng AM, Byrom MW, Shelton J, Ford LP, Cheng AM, Byrom MW, Shelton J, Ford LP (2005) Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. Nucleic Acids Res 33:1290–1297.

4. Stenvang J, Kauppinen S (2008) MicroRNAs as targets for antisense-based therapeutics. Expert Opin Biol Ther 8:59–81.
5. Eckstein F (2007) The versatility of oligonucleotides as potential therapeutics. Expert Opin Biol Ther 7:1021–1034.
6. Thum T, Galuppo P, Wolf C, Fiedler J, Kneitz S, van Laake LW, Doevendans PA, Mummery CL, Borlak J, Haverich A, Gross C, Engelhardt S, Ertl G, Bauersachs J (2007) MicroRNAs in the human heart: a clue to fetal gene reprogramming in heart failure. Circulation 116:258–267.
7. van Rooij E, Sutherland LB, Liu N, Williams AH, McAnally J, Gerard RD, Richardson JA, Olson EN (2006) A signature pattern of stress-responsive microRNAs that can evoke cardiac hypertrophy and heart failure. Proc Natl Acad Sci U S A 103: 18255–18260.
8. Sayed D, Hong C, Chen IY, Lypowy J, Abdellatif M (2007) MicroRNAs play an essential role in the development of cardiac hypertrophy. Circ Res 100:416–424.
9. Carè A, Catalucci D, Felicetti F, Bonci D, Addario A, Gallo P, Bang ML, Segnalini P, Gu Y, Dalton ND, Elia L, Latronico MV, Høydal M, Autore C, Russo MA, Dorn GW, Ellingsen O, Ruiz-Lozano P, Peterson KL, Croce CM, Peschle C, Condorelli G (2007) MicroRNA-133 controls cardiac hypertrophy. Nat Med 13:613–618.
10. Cheng Y, Ji R, Yue J, Yang J, Liu X, Chen H, Dean DB, Zhang C (2007) MicroRNAs are aberrantly expressed in hypertrophic heart. Do they play a role in cardiac hypertrophy? Am J Pathol 170:1831–1840.
11. Tatsuguchi M, Seok HY, Callis TE, Thomson JM, Chen JF, Newman M, Rojas M, Hammond SM, Wang DZ (2007) Expression of microRNAs is dynamically regulated during cardiomyocyte hypertrophy. J Mol Cell Cardiol 42:1137–1141.
12. Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M, Prueitt RL, Yanaihara N, Lanza G, Scarpa A, Vecchione A, Negrini M, Harris CC, Croce CM (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A 103:2257–2261.
13. Krek A, Grun D, Poy M, Wolf R, Rosenberg L, Epstein E, MacMenamin P, da Piedade I, Gunsalus K, Stoffel M, Rajewsky N (2005) Combinatorial microRNA target predictions. Nat Genet 37:495–500.
14. Bommer GT, Gerin I, Feng Y, Kaczorowski AJ, Kuick R, Love RE, Zhai Y, Giordano TJ, Qin ZS, Moore BB, MacDougald OA, Cho KR, Fearon ER (2007) p53-mediated activation of miRNA34 candidate tumor-suppressor genes. Curr Biol 17:1298–1307.
15. Ebert MS, Neilson JR, Sharp PA (2007) MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. Nat Methods 4:721–726.
16. Bartel DP, Chen CZ (2004) Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. Nat Rev Genet 5:396–400.
17. Vermeulen A, Robertson B, Dalby AB, Marshall WS, Karpilow J, Leake D, Khvorova A, Baskerville S (2007) Double-stranded regions are essential design components of potent inhibitors of RISC function. RNA 13:723–730.
18. Pedersen IM, Cheng G, Wieland S, Volinia S, Croce CM, Chisari FV, David M (2007) Interferon modulation of cellular microRNAs as an antiviral mechanism. Nature 449:919–922.
19. Lu Y, Xiao J, Lin H, Bai Y, Luo X, Wang Z, Yang B (2009) Complex antisense inhibitors offer a superior approach for microRNA research and therapy. Nucleic Acids Res 37:e24–e33.
20. Gao H, Xiao J, Sun Q, Lin H, Bai Y, Yang L, Yang B, Wang H, Wang Z (2006) A single decoy oligodeoxynucleotides targeting multiple oncoproteins produces strong anticancer effects. Mol Pharmacol 70:1621–1629.
21. Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, Magri E, Pedriali M, Fabbri M, Campiglio M, Menard S, Palazzo JP, Rosenberg A, Musiani P, Volinia S, Nenci I, Calin GA, Querzoli P, Negrini M, Croce CM (2005) MicroRNA gene expression deregulation in human breast cancer. Cancer Res 65: 7065–7070.
22. Henkel J (1999) Attacking AIDS with a 'cocktail' therapy? FDA Consum 33:12–17.
23. Konlee M (1998) An evaluation of drug cocktail combinations for their immunological value in preventing/remitting opportunistic infections. Posit Health News 16:2–4.
24. Charpentier G (2002) Oral combination therapy for type 2 diabetes. Diabetes Metab Res Rev 18(Suppl 3):S70–S76.
25. Ogihara T (2003) The combination therapy of hypertension to prevent cardiovascular events (COPE) trial: rationale and design. Hypertens Res 28:331–338.
26. Kumar P (2005) Combination treatment significantly enhances the efficacy of antitumor therapy by preferentially targeting angiogenesis. Lab Investig 85:756–767.
27. Nabholtz JM, Gligorov J (2005) Docetaxel/trastuzumab combination therapy for the treatment of breast cancer. Expert Opin Pharmacother 6:1555–1564.

Chapter 5

Modulation of MicroRNAs for Potential Cancer Therapeutics

Wei Wu

Abstract

MicroRNA (miRNA) is a nonprotein coding small RNA molecule that negatively regulates gene expression by degradation of mRNA or suppression of mRNA translation. MiRNA plays important roles in physiological processes such as cellular development, differentiation, proliferation, apoptosis, and stem cell self-renewal. Studies show that the deregulation of miRNA expression is closely associated with tumorigenicity, invasion, and metastasis. The functionality of miRNAs may act as oncogenes or tumor suppressors during tumor initiation and progression. miRNomes in almost all types of cancers started to develop the regulatory network of miRNA::mRNA interaction in the view of systems biology. Experimental evidence demonstrates that the modulation of specific miRNA alterations in cancer cells using miRNA replacement or anti-miRNA technologies can restore miRNA activities and repair the gene regulatory network and signaling pathways, in turn, reverse the phenotype of cancerous cells. Numerous animal studies for miRNA-based therapy offer the hope of targeting miRNAs as alternative cancer treatment. Developing the small molecules to interfere with miRNAs could be of great pharmaceutical interest in the future. Interestingly, specific miRNA is capable of reprogramming the cancer cells into a pluriopotent embryonic stem cell-like state (mirPS), which could be induced into tissue-specific mature cell types. This chapter will present the various strategies of modulation of miRNAs in vitro and in vivo.

Key words: MicroRNA, MiRNA interference, MicroRNA mimics, AMO, Antagomir, Cancer reprogramming, mirPS

1. Introduction

MicroRNAs (miRNAs) are a class of endogenous noncoding small RNA molecules (~22 nucleotides, nt), that play important roles in development, differentiation, proliferation, and apoptosis (1). Since the miRNAs, lin-4 and let-7, were discovered in the *Caenorhabditis elegans* (2, 3), thousands of miRNAs have been identified in different species including worms, flies, plants, and humans (4–6). The most recent release of the miRBase sequence

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_5, © Springer Science+Business Media, LLC 2011

(http://www.mirbase.org/, 14.0 released on September 2009) lists 10,886 different miRNAs identified in all species, among them are 721 human miRNAs. It is predicted that the human genome encodes about 1,000 miRNAs (7). MiRNAs may be transcribed by separate genes or along with the coding genes (located in introns or extrons) (8). The biogenesis of miRNAs from the primary miRNAs (several hundred nucleotides) to pre-miRNAs and the mature miRNAs is a complicated biochemical process (7), (For details on miRNAs biogenesis, see Chapter 1). When the mature miRNAs forming a complex with RISC (RNA-induced silencing complex) bind to the 3′-untranslational region (3′UTR), a cleavage of mRNA transcripts (in the case of perfect miRNA::mRNA complementarity) or translational repression (in the case of imperfect complementarity) occurs (9). The nature of miRNA acting on mRNAs adds an additional layer of fine tuning the gene expression in human genome, offering a complexity of gene regulatory network. Deregulation of miRNA::mRNA interaction contributes to a variety of human diseases including cancer (10).

Mounting evidence shows that miRNAs are aberrantly expressed during cancer development (11, 12), invasion and metastasis (13), angiogenesis (14) and play crucial roles in cancer stem cell regulation (15). MiRNAs may function as oncogenes (OncomiRs) or tumor suppressors (TSmiRs) in tumors. With the intensive study of miRNomes in almost all types of cancers, the miRNA::mRNA interactive network is beginning to be pictured for a deep understanding of cancer biology. Strikingly, the compelling research on targeting miRNAs as experimental therapy in vitro or in vivo are increasing. Diverse technologies of replacing or inhibiting miRNAs to restore the miRNAs functions have been developed, and small-molecule modifiers of miRNAs are proposed to regulate the miRNAs activities. Delivery of miRNAs into the mice and nonhuman primates provides the pharmacokinetic base for possible human trial in the nearest future. Notably, some miRNAs could reprogram the human cancer cells into a pluriopotent ES cell-like state (mirPS), which could be further induced into differentiated cell types. All these processes will be covered in this chapter.

2. miRNAs Involved in Cancer Development

Cancer is both a genetic disorder and a developmental disease. From the pathohistological to molecular genetic studies, we have gained a deep understanding of cancer biology over a century. Within the past decade, we learned that noncoding miRNAs play roles in cancer initiation, progression, invasion, and metastasis (10).

Here, the association of microNRAs and cancers will be briefly described.

Genetic studies show that more than 50% of miRNA genes can be found in cancer-associated genomic regions or in fragile sites, suggesting that miRNAs play an important role in the pathogenesis of neoplasmas (16). Emerging evidence demonstrate that a unique set of miRNA expression exists in a specific type of cancers, and unique expression profiling of miRNAs presents in a variety of cancers (10). Loss of certain miRNAs' function may promote tumorigenicity, these miRNAs act as tumor suppressors (TSmiRs). For example, the miR-15b and miR-16 clusters at 13q14 were identified as the first TSmiRs because they are deleted and/or downregulated in about 68% of B-cell chronic lymphocytic leukemia (CLL) samples (16, 17). MiR-15b and miR-16 have been shown to control the expression of VEGF, a key proangiogenic factor involving in tumor angiogenesis (14). In addition, miR-15b and miR-16 also induce apoptosis of leukemic cells by affecting the antiapoptotic protein Bcl-2 (18).

Reduced expression of another group of tumor suppressive miRNAs, the let-7 family members, is a common genetic event in nonsmall cell lung cancer and is associated with a poor prognosis (19). The let-7 family is able to negatively regulate let-60/RAS, providing a possible mechanism for let-7 deficiency in cancers (19). In addition, let-7 regulates late embryonic development by suppressing the expression of a number of genes, such as c-Myc and RAS, and the embryonic gene, high-mobility group A2 (HMGA2) (20). Let-7 family induces cancer stem cell differentiation in breast cancer (15). Taken together, reduction of TSmiRs expression in cancers activates the oncogenic genes and promotes the tumor initiation and progression.

In contrast, oncogenic miRNAs that promote tumoregenesis when overexpressed in cancer cells have also been identified. The discovery of miR-17-92 cluster provided the first functional evidence of such an "oncomir" (11). The mir-17-92 cluster, which encodes miR-17-5p, miR-17-3p, miR-18a, miR-19a, miR-20a, miR-19b-1, and miR-92-1, was found to be located within a region on chromosome 13 that is commonly amplified in human B-cell lymphomas (21). In vivo analysis confirmed that mir-17-92 accelerated c-Myc-induced lymphomagenesis (11). Furthermore, mir-17-92 expression is elevated in lung cancer (22) and anaplastic thyroid cancer cells (23), although the opposite effect that mir-17-92 may function as TSmiRs in breast cancer cells was reported (24). The oncomir, miR-21, is overexpressed in nine types of solid tumors (lung, breast, stomach, prostate, colon, brain, head and neck, esophagus, and pancreas) as well as in diffuse large B-cell lymphoma, CLL, and malignant hepatocytes, supporting its oncogenic role in cancer pathogenesis (25). Several studies show that miR-21 inhibits numerous tumor suppressor

genes including phosphatase and tensin homolog deleted on chromosome 10 (PTEN) (26), tropomyosin 1 (27), programmed cell death 4 (PDCD 4) (25, 28, 29), RECK and TIMP3 (30), and sprouty2 (31). Therefore, gain of specific miRNA function in cells may inactivate conventional tumor suppressor genes, and in turn, facilitate carcinogenesis.

Because aberrant miRNA expression is now recognized as a common feature of cancers, specific miRNAs are being considered as potential new biomarkers for cancer diagnosis and prognosis (32). Taken together, loss of function or gain of function of miRNAs contributes to the cancer initiation and progression, correction of these miRNAs and their regulated gene network could change the behavior of cancer cells, thereby making them a promising target for new drug development, as it will be discussed below.

3. Modulating miRNAs for Potential Cancer Intervention

MiRNAs provide a particular layer for gene regulation. About 3% of the genes in the genome encode the miRNAs, and 30% of the coding genes could be regulated by miRNAs (33). One miRNA could have multiple mRNA targets, on the other hand, one mRNA may be inhibited by more than one miRNAs. It is speculated that the change of one miRNA expression in the miRNA:: mRNA network could trigger a chain reaction; when the changes reach the trigger points (threshold), the cells may change the biological behavior. This concept of modulating miRNAs in different ways thus possesses great potential both as a novel class of therapeutic targets and as a powerful intervention tool (34).

3.1. miRNA Replacement Therapy

Decreased expression of a subgroup of miRNA is present in certain cancers, where loss of a miRNA inhibitory effect contributes to the oncogene activation. This straightforward strategy of adding the miRNAs missing in the cancer cells to restore their "normal" functions is called miRNA replacement therapy. Numerous studies observed the efficacy of the miRNA replacement therapy in the cultured cells or animal models. This approach has been used to increase Let-7 expression, which is significantly reduced in lung cancer cell lines and lung cancer tissues. Takamizawa et al. (35) designed expression constructs to synthesize mature miRNAs of two predominant *let-7* isoforms, *let-7a* and *let-7f*, under the control of the RNA polymerase III H1-RNA gene promoter. When the let-7 expression vector was introduced into A549 adenocarcinoma cell line, a 78.6% reduction in the number of colonies was observed. In two separate experiments, an overexpression vector bearing let-7 in lung cancer cell lines enhanced lung cancer cell radiosensitivity (36) and altered cell cycle progression and

reduced cell division (37). Moreover, intranasal let-7 administration reduces tumor formation in vivo in the lungs of animals expressing a G12D activating mutation for the K-ras oncogene (38, 39). Intratumoral delivery of let-7 reduces tumor size in a lung cancer xenograft model (39). Similarly, enforced let-7 expression in primary breast tumor initiating cells inhibits tumorigenesis and metastasis (15). These findings indicate that this miRNA may be useful as a novel therapeutic agent in cancers.

3.2. miRNA Mimic Therapy

miRNA mimics are small, chemically modified, double-stranded RNA molecules that mimic endogenous mature miRNA molecules. Xiao et al. developed this method which has been translated into experimental therapy (40) (also see Chapter 15). The advantages of this approach are: (1) miRNA targeting in gene-specific fashion; (2) less unwanted off target effects; (3) relative feasibility of delivery using lipid- or polymer-based nanoparticles for systemic delivery in vivo. The disadvantages of this approach are that miRNA mimic oligos have only a transient effect, are unstable, and may require repeated supplementation.

Recently, a new RNA polymerase II expression vector including endogenous murine miR-155 flanking sequence (41) has been widely employed in in vitro and in vivo studies (42–44). However, these studies only examined miRNA-expressed cells in culture and in xenograft models. How to directly deliver miRNA-mimic oligos or vector-based miRNA expression to specific tumors or organs in animal models remains to be addressed.

3.3. Anti-miR Oligonucleotides

Overexpression of a subgroup of miRNAs contributes to malignant growth in cancer cells. For these miRNAs, a rational goal would be the reduction of expression. Inhibition of specific endogenous miRNAs has been achieved through the administration of synthetic antisense oligonucleotides that are complementary to the endogenous mature miRNAs. Currently, the most widely used anti-miRNA oligonucleotides (AMO) are: (a) 2′-*O*-methyl AMOs, (b) 2′-*O*-methoxyethyl AMOs, and (c) locked nucleic acid (LNA) AMOs (45). Treatment with these modified RNA oligos results in the stability of serum and cellular uptake and a delayed clearance following systemic administration (46). Onco-miR-21 knockdown studies in vitro provide good examples of an anti-miR application. Applying 2′-*O*-methyl- and/or DNA/LNA-mixed oligonucleotides to specifically knockdown miR-21 in cultured glioblastoma and breast cancer cells suppresses cell growth in vitro in association with increased caspase activation-mediated apoptosis (47, 48). Suppression of mir-21 also significantly reduced invasion and lung metastasis in MDA-MB-231 metastatic breast cancer cells (13), RKO human colon cancer cells (28), glioblastoma cells (30), and pancreatic cancer cells (49). In addition, the inhibition of miR-21 and miR-200b with AMO increased the sensitivity to gemcitabine

treatment in malignant cholangiocytes that highly overexpressed miR-21, miR-141, and miR-200b (50). Finally, administration of LNA–anti-miR-21 oligonucleotides disrupted glioma growth in vivo and displayed synergistic cytotoxicity with neural precursor cell-delivered S-TRAIL in human gliomas (51). Anti-miR21, which is now being developed, appears to be a promising anticancer drug, but more stringent pharmacological study and clinical trial remain to be conducted.

3.4. Modified AMO

Several modified AMO technologies have been developed to improve the specificity and efficiency to inhibit the oncomirs. Ebert et al. (52) developed a new form of miRNA inhibitors termed "miRNA sponges" that can be transiently expressed in cultured mammalian cells. MiRNA sponges are transcripts under the control of strong promoters (RNA polymerase II), containing multiple tandem binding sites to a miRNA of interest, which are able to inhibit miRNA targets as strongly as AMO. Xiao et al. designed alternative strategy called "miRNA masking" which refers a sequence with perfect complementarity to the binding site of an endogenous miRNA in the target gene. This "miRNA masking" can form duplex with the target mRNA with higher affinity, therefore blocking the access of endogenous miRNA to its binding site without the potential side effects of mRNA degradation by AMOs (40). Lu et al. (53) recently developed multiple target of AMOs; a single AMO fragment with multiple miRNA sequences has the capability of inhibiting multiple miRNAs (such as miR-21, miR-155, and miR-17-5p). These technologies are currently being tested and are expected to have wide clinical implications.

3.5. Small-Molecule Modifier

Small molecules have been reported to modulate embryonic stem cell fate and somatic cell reprogramming (54). It is speculated that small molecules could induce or inhibit miRNAs functions. Indeed, Gumireddy et al. developed a luciferase reporter method to identify a compound, diazobenzene, which specifically inhibits the transcription of the miR-21 gene into pri-miR-21 (55). One study reported that an established chemotherapeutic agent, 5-fluorouracil, induces several miRNAs including miR-21 (56). Estradiol may upregulate miR-21 (57) or suppress miR-21 expression (58) depending on the cell content. The small-molecule compounds could interact with miRNAs and prevent miRNA process (59). The discovered small-molecule modifiers of the miRNA represent unique tools to manipulate miRNA function. The advantages of the small molecules that are more easily delivered in animals or humans, more stable intracellularly, and less expensive to manufacture draw a great attention for promising developments. However, the in vivo efficacy of the small-molecule modifiers needs to be explored further.

4. Delivery of miRNA In Vivo

To develop a pharmacological approach for silencing miRNAs in vivo, Krutzfeldt et al. (60) designed chemically modified, cholesterol-conjugated, single-stranded RNA analogs (termed "antagomirs"). The antagomirs efficiently inhibited the activity of miR-16, miR-122, miR-192, and miR-194 in mice after three daily intravenous injections resulting in a marked reduction of corresponding miRNA levels in multiple organs (liver, lung, kidney, heart, intestine, fat, skin, bone marrow, muscle, ovaries, and adrenals). The cholesterol-conjugated, single-strand RNA analog and locked nucleotide acid anti-miR-122 effectively knocked down miR-122 expression in the liver thereby normalizing plasma cholesterol and mRNA expression level in a mouse model. The conjugation of RNA oligos with other lipophilic molecules (e.g., high-density lipoprotein) has also been used to successfully deliver anti-miRs to specific organs (61). The study of LNA-mediated targeting of miR-122 in mice and nonhuman primates (African green monkeys) observed: (1) the dose-dependent and sustained decrease in the plasma cholesterol level with systemic delivery; (2) local accumulation of LNA–anti-miR complex in the liver; (3) no acute or subacute toxicity in LNA–anti-miR-treated subjects (62, 63). A miRNA replacement therapy with miR-26a for liver cancer in mice model has been investigated with scAAV8 vector-based in vivo delivery. The enriched miR-26a in the mouse liver induces cancer cell death without toxicity (64). These in vivo animal studies provide valuable pharmacokinetic data regarding miRNA distribution and metabolism in an animal model, paving the way for future preclinical trials using AMO or replacement for anticancer therapy.

5. miRNA Reprogramming Cancer Stem Cells

Reprogramming the differentiated somatic cells with cocktail transcription factors (Oct4-Sox2-c-Myc-klf-4 or Oc4-Sox2-Nanog-Lin28) (65,66) is the breakthrough in stem cell biology. Subsequently, the embryonic stem cell miRNAs are defined to enhance the production of mouse-induced pluripotent stem (iPS) cells. The miRNAs (miR-291-3p, miR-294, and miR-295) increase the efficiency of reprogramming by three transcription factors (Oct4, Sox2, and Klf4), but not by adding c-Myc to these factors. C-Myc binds the promoter of the miRNAs, suggesting that they are downstream effectors of c-Myc during reprogramming (67). The miR-302 family (miR-302s) is expressed most abundantly in slow-growing

human embryonic stem cells and quickly decreased after cell differentiation and proliferation. Therefore, miR-302s are the key factors essential for the maintenance of ES cell renewal and pluripotency. It appears that certain miRNAs can dramatically influence the cancer stem cell fate as well. The miR-302s-transfected human cancer cells (prostate cancer and skin cancer) exhibit many key ES cell markers and maintain a pluripotent ES cell-like state, namely, miRNA-induced pluripotent stem cells (mirPS). The mirPS can be induced into a differentiated state and subsequently develop a diversity of tissue-specific cell types (68). On the other hand, miRNA let-7 family is markedly reduced in breast tumor initiating cells (or breast cancer stem cells) and increased with differentiation (15). Let-7 is considered as a key "keeper" of the differentiated state. On these two scenarios, if we reduce the "stemness miRNAs" expression or increase the "differentiation miRNAs" by exogenous expression or treatment with small molecules, we might be able to switch cancer cells into relative normal cell types leading to renewed differentiation therapy.

6. Conclusion and Outlook

Following the discovery of the first endogenous small noncoding silencing RNA in 1993, the study of miRNAs has revolutionized our understanding of gene expression regulation during oncogenesis. MiRNA replacement or anti-miRNAs with synthesized oligonucleotides or by small molecules open up an avenue for an integrated cancer treatment. MiRNA is believed to be relatively safe and more effective in cancer treatment in early preclinical studies comparing to siRNA-based therapy (34). Nevertheless, there are still several obstacles to overcome before testing miRNA drug treatment clinically, such as delivery of miRNA to a specific tissue or disease site, avoiding off-target effects, optimizing the dose, and minimizing likelihood of immune activation. Moving forward, design of miRNA drugs for use in these trials should consider the following: (1) conjugation with a specific antibody such as trastuzumab, an anti-HER2 monoclonal antibody, to target HER2 positive breast cancer, (2) conjugation with specific ligand binding to the particular receptors, for example, the EGFR, which guide the miRNA to cancer cells, (3) simultaneous treatment with miRNA and a chemodrug to produce synergistic antineoplastic effects (34). With further understanding of miRNA biogenesis, miRNA expression profiling in primary tumors, and the gene regulatory network between miRNAs and mRNAs, translational implications of miRNA research for cancer patients treatment are foreseeable in the nearest future.

References

1. Kato, M., and Slack, F. J. (2008) MicroRNAs: small molecules with big roles – *C. elegans* to human cancer, *Biol Cell 100*, 71–81.
2. Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14, *Cell 75*, 843–854.
3. Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*, *Nature 403*, 901–906.
4. Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs, *Science 294*, 853–858.
5. Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*, *Science 294*, 858–862.
6. Lee, R. C., and Ambros, V. (2001) An extensive class of small RNAs in *Caenorhabditis elegans*, *Science 294*, 862–864.
7. Bartel, D. P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function, *Cell 116*, 281–297.
8. Kim, V. N., and Nam, J. W. (2006) Genomics of microRNA, *Trends Genet 22*, 165–173.
9. Hutvagner, G., and Zamore, P. D. (2002) A microRNA in a multiple-turnover RNAi enzyme complex, *Science 297*, 2056–2060.
10. Wu, W., Sun, M., Zou, G. M., and Chen, J. (2006) MicroRNA and cancer: Current status and prospective, *Int J Cancer 120*, 953–960.
11. He, L., Thomson, J. M., Hemann, M. T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S. W., Hannon, G. J., and Hammond, S. M. (2005) A microRNA polycistron as a potential human oncogene, *Nature 435*, 828–833.
12. O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V., and Mendell, J. T. (2005) c-Myc-regulated microRNAs modulate E2F1 expression, *Nature 435*, 839–843.
13. Zhu, S., Wu, H., Wu, F., Nie, D., Sheng, S., and Mo, Y. Y. (2008) MicroRNA-21 targets tumor suppressor genes in invasion and metastasis, *Cell Res 18*, 350–359.
14. Hua, Z., Lv, Q., Ye, W., Wong, C. K., Cai, G., Gu, D., Ji, Y., Zhao, C., Wang, J., Yang, B. B., and Zhang, Y. (2006) MiRNA-directed regulation of VEGF and other angiogenic factors under hypoxia, *PLoS ONE 1*, e116.
15. Yu, F., Yao, H., Zhu, P., Zhang, X., Pan, Q., Gong, C., Huang, Y., Hu, X., Su, F., Lieberman, J., and Song, E. (2007) let-7 regulates self renewal and tumorigenicity of breast cancer cells, *Cell 131*, 1109–1123.
16. Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., and Croce, C. M. (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers, *Proc Natl Acad Sci U S A 101*, 2999–3004.
17. Calin, G. A., Liu, C. G., Sevignani, C., Ferracin, M., Felli, N., Dumitru, C. D., Shimizu, M., Cimmino, A., Zupo, S., Dono, M., Dell'Aquila, M. L., Alder, H., Rassenti, L., Kipps, T. J., Bullrich, F., Negrini, M., and Croce, C. M. (2004) MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias, *Proc Natl Acad Sci U S A 101*, 11755–11760.
18. Cimmino, A., Calin, G. A., Fabbri, M., Iorio, M. V., Ferracin, M., Shimizu, M., Wojcik, S. E., Aqeilan, R. I., Zupo, S., Dono, M., Rassenti, L., Alder, H., Volinia, S., Liu, C. G., Kipps, T. J., Negrini, M., and Croce, C. M. (2005) miR-15 and miR-16 induce apoptosis by targeting BCL2, *Proc Natl Acad Sci USA 102*, 13944–13949.
19. Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K. L., Brown, D., and Slack, F. J. (2005) RAS is regulated by the let-7 microRNA family, *Cell 120*, 635–647.
20. Park, S. M., Shell, S., Radjabi, A. R., Schickel, R., Feig, C., Boyerinas, B., Dinulescu, D. M., Lengyel, E., and Peter, M. E. (2007) Let-7 prevents early cancer progression by suppressing expression of the embryonic gene HMGA2, *Cell Cycle 6*, 2585–2590.
21. Ota, A., Tagawa, H., Karnan, S., Tsuzuki, S., Karpas, A., Kira, S., Yoshida, Y., and Seto, M. (2004) Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma, *Cancer Res 64*, 3087–3095.
22. Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., Yatabe, Y., Kawahara, K., Sekido, Y., and Takahashi, T. (2005) A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation, *Cancer Res 65*, 9628–9632.
23. Takakura, S., Mitsutake, N., Nakashima, M., Namba, H., Saenko, V. A., Rogounovitch, T. I., Nakazawa, Y., Hayashi, T., Ohtsuru, A.,

and Yamashita, S. (2008) Oncogenic role of miR-17-92 cluster in anaplastic thyroid cancer cells, *Cancer Sci 99*, 1147–1154.

24. Hossain, A., Kuo, M. T., and Saunders, G. F. (2006) Mir-17-5p regulates breast cancer cell proliferation by inhibiting translation of AIB1 mRNA, *Mol Cell Biol 26*, 8191–8201.
25. Lu, Z., Liu, M., Stribinskis, V., Klinge, C. M., Ramos, K. S., Colburn, N. H., and Li, Y. (2008) MicroRNA-21 promotes cell transformation by targeting the programmed cell death 4 gene, *Oncogene 27*, 4373–4379.
26. Meng, F., Henson, R., Wehbe-Janek, H., Ghoshal, K., Jacob, S. T., and Patel, T. (2007) MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer, *Gastroenterology 133*, 647–658.
27. Zhu, S., Si, M. L., Wu, H., and Mo, Y. Y. (2007) MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1), *J Biol Chem 282*, 14328–14336.
28. Asangani, I. A., Rasheed, S. A., Nikolova, D. A., Leupold, J. H., Colburn, N. H., Post, S., and Allgayer, H. (2008) MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer, *Oncogene 27*, 2128–2136.
29. Frankel, L. B., Christoffersen, N. R., Jacobsen, A., Lindow, M., Krogh, A., and Lund, A. H. (2008) Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells, *J Biol Chem 283*, 1026–1033.
30. Gabriely, G., Wurdinger, T., Kesari, S., Esau, C. C., Burchard, J., Linsley, P. S., and Krichevsky, A. M. (2008) MiR-21 promotes glioma invasion by targeting MMP regulators, *Mol Cell Biol 28(17)*, 5369–5380.
31. Sayed, D., Rane, S., Lypowy, J., He, M., Chen, I. Y., Vashistha, H., Yan, L., Malhotra, A., Vatner, D., and Abdellatif, M. (2008) MicroRNA-21 targets Sprouty2 and promotes cellular outgrowths, *Mol Biol Cell 19(8)*, 3272–3282.
32. Yu, S. L., Chen, H. Y., Chang, G. C., Chen, C. Y., Chen, H. W., Singh, S., Cheng, C. L., Yu, C. J., Lee, Y. C., Chen, H. S., Su, T. J., Chiang, C. C., Li, H. N., Hong, Q. S., Su, H. Y., Chen, C. C., Chen, W. J., Liu, C. C., Chan, W. K., Chen, W. J., Li, K. C., Chen, J. J., and Yang, P. C. (2008) MicroRNA signature predicts survival and relapse in lung cancer, *Cancer Cell 13*, 48–57.
33. Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003) Prediction of mammalian microRNA targets, *Cell 115*, 787–798.
34. Wang, V., and Wu, W. (2009) MicroRNA-based therapeutics for cancer, *BioDrugs 23*, 15–23.
35. Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., Harano, T., Yatabe, Y., Nagino, M., Nimura, Y., Mitsudomi, T., and Takahashi, T. (2004) Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival, *Cancer Res 64*, 3753–3756.
36. Weidhaas, J. B., Babar, I., Nallur, S. M., Trang, P., Roush, S., Boehm, M., Gillespie, E., and Slack, F. J. (2007) MicroRNAs as potential agents to alter resistance to cytotoxic anticancer therapy, *Cancer Res 67*, 11111–11116.
37. Johnson, C. D., Esquela-Kerscher, A., Stefani, G., Byrom, M., Kelnar, K., Ovcharenko, D., Wilson, M., Wang, X., Shelton, J., Shingara, J., Chin, L., Brown, D., and Slack, F. J. (2007) The let-7 microRNA represses cell proliferation pathways in human cells, *Cancer Res 67*, 7713–7722.
38. Esquela-Kerscher, A., Trang, P., Wiggins, J. F., Patrawala, L., Cheng, A., Ford, L., Weidhaas, J. B., Brown, D., Bader, A. G., and Slack, F. J. (2008) The let-7 microRNA reduces tumor growth in mouse models of lung cancer, *Cell Cycle 7*, 759–764.
39. Trang, P., Medina, P. P., Wiggins, J. F., Ruffino, L., Kelnar, K., Omotola, M., Homer, R., Brown, D., Bader, A. G., Weidhaas, J. B., and Slack, F. J. (2010) Regression of murine lung tumors by the let-7 microRNA, *Oncogene 29(11)*, 1580–1587.
40. Xiao, J., Yang, B., Lin, H., Lu, Y., Luo, X., and Wang, Z. (2007) Novel approaches for gene-specific interference via manipulating actions of microRNAs: examination on the pacemaker channel genes HCN2 and HCN4, *J Cell Physiol 212*, 285–292.
41. Chung, K. H., Hart, C. C., Al-Bassam, S., Avery, A., Taylor, J., Patel, P. D., Vojtek, A. B., and Turner, D. L. (2006) Polycistronic RNA polymerase II expression vectors for RNA interference based on BIC/miR-155, *Nucleic Acids Res 34*, e53.
42. McLaughlin, J., Cheng, D., Singer, O., Lukacs, R. U., Radu, C. G., Verma, I. M., and Witte, O. N. (2007) Sustained suppression of Bcr-Abl-driven lymphoid leukemia by microRNA mimics, *Proc Natl Acad Sci USA 104*, 20501–20506.
43. Liang, Z., Wu, H., Reddy, S., Zhu, A., Wang, S., Blevins, D., Yoon, Y., Zhang, Y., and Shim, H. (2007) Blockade of invasion and metastasis of breast cancer cells via targeting CXCR4

with an artificial microRNA, *Biochem Biophys Res Commun 363*, 542–546.

44. Li, Z., Zhan, W., Wang, Z., Zhu, B., He, Y., Peng, J., Cai, S., and Ma, J. (2006) Inhibition of PRL-3 gene expression in gastric cancer cell line SGC7901 via microRNA suppressed reduces peritoneal metastasis, *Biochem Biophys Res Commun 348*, 229–237.
45. Weiler, J., Hunziker, J., and Hall, J. (2006) Anti-miRNA oligonucleotides (AMOs): ammunition to target miRNAs implicated in human disease? *Gene Ther 13(6)*, 496–502.
46. Wahlestedt, C., Salmi, P., Good, L., Kela, J., Johnsson, T., Hokfelt, T., Broberger, C., Porreca, F., Lai, J., Ren, K., Ossipov, M., Koshkin, A., Jakobsen, N., Skouv, J., Oerum, H., Jacobsen, M. H., and Wengel, J. (2000) Potent and nontoxic antisense oligonucleotides containing locked nucleic acids, *Proc Natl Acad Sci U S A 97*, 5633–5638.
47. Chan, J. A., Krichevsky, A. M., and Kosik, K. S. (2005) MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells, *Cancer Res 65*, 6029–6033.
48. Si, M. L., Zhu, S., Wu, H., Lu, Z., Wu, F., and Mo, Y. Y. (2006) miR-21-mediated tumor growth, *Oncogene 26(19)*, 2799–2803.
49. Moriyama, T., Ohuchida, K., Mizumoto, K., Yu, J., Sato, N., Nabae, T., Takahata, S., Toma, H., Nagai, E., and Tanaka, M. (2009) MicroRNA-21 modulates biological functions of pancreatic cancer cells including their proliferation, invasion, and chemoresistance, *Mol Cancer Ther 8(5)*, 1067–1074.
50. Meng, F., Henson, R., Lang, M., Wehbe, H., Maheshwari, S., Mendell, J. T., Jiang, J., Schmittgen, T. D., and Patel, T. (2006) Involvement of human micro-RNA in growth and response to chemotherapy in human cholangiocarcinoma cell lines, *Gastroenterology 130*, 2113–2129.
51. Corsten, M. F., Miranda, R., Kasmieh, R., Krichevsky, A. M., Weissleder, R., and Shah, K. (2007) MicroRNA-21 knockdown disrupts glioma growth in vivo and displays synergistic cytotoxicity with neural precursor cell delivered S-TRAIL in human gliomas, *Cancer Res 67*, 8994–9000.
52. Ebert, M. S., Neilson, J. R., and Sharp, P. A. (2007) MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells, *Nat Methods 4*, 721–726.
53. Lu, Y., Xiao, J., Lin, H., Bai, Y., Luo, X., Wang, Z., and Yang, B. (2009) A single anti-microRNA antisense oligodeoxyribonucleotide (AMO) targeting multiple microRNAs offers an improved approach for microRNA interference, *Nucleic Acids Res 37*, e24.
54. Abujarour, R., and Ding, S. (2009) Induced pluripotent stem cells free of exogenous reprogramming factors, *Genome Biol 10*, 220.
55. Gumireddy, K., Young, D. D., Xiong, X., Hogenesch, J. B., Huang, Q., and Deiters, A. (2008) Small-molecule inhibitors of microrna miR-21 function, *Angew Chem Int Ed Engl 47*, 7482–7484.
56. Rossi, L., Bonmassar, E., and Faraoni, I. (2007) Modification of miR gene expression pattern in human colon cancer cells following exposure to 5-fluorouracil in vitro, *Pharmacol Res 56*, 248–253.
57. Bhat-Nakshatri, P., Wang, G., Collins, N. R., Thomson, M. J., Geistlinger, T. R., Carroll, J. S., Brown, M., Hammond, S., Srour, E. F., Liu, Y., and Nakshatri, H. (2009) Estradiol-regulated microRNAs control estradiol response in breast cancer cells, *Nucleic Acids Res 37*, 4850–4861.
58. Wickramasinghe, N. S., Manavalan, T. T., Dougherty, S. M., Riggs, K. A., Li, Y., and Klinge, C. M. (2009) Estradiol downregulates miR-21 expression and increases miR-21 target gene expression in MCF-7 breast cancer cells, *Nucleic Acids Res 37*, 2584–2595.
59. Deiters, A. (2010) Small molecule modifiers of the microRNA and RNA interference pathway, *AAPS J 12(1)*, 51–60.
60. Krutzfeldt, J., Rajewsky, N., Braich, R., Rajeev, K. G., Tuschl, T., Manoharan, M., and Stoffel, M. (2005) Silencing of microRNAs in vivo with 'antagomirs', *Nature 438*, 685–689.
61. Wolfrum, C., Shi, S., Jayaprakash, K. N., Jayaraman, M., Wang, G., Pandey, R. K., Rajeev, K. G., Nakayama, T., Charrise, K., Ndungo, E. M., Zimmermann, T., Koteliansky, V., Manoharan, M., and Stoffel, M. (2007) Mechanisms and optimization of in vivo delivery of lipophilic siRNAs, *Nat Biotechnol 25*, 1149–1157.
62. Elmen, J., Lindow, M., Schutz, S., Lawrence, M., Petri, A., Obad, S., Lindholm, M., Hedtjarn, M., Hansen, H. F., Berger, U., Gullans, S., Kearney, P., Sarnow, P., Straarup, E. M., and Kauppinen, S. (2008) LNA-mediated microRNA silencing in non-human primates, *Nature 452*, 896–899.
63. Elmen, J., Lindow, M., Silahtaroglu, A., Bak, M., Christensen, M., Lind-Thomsen, A., Hedtjarn, M., Hansen, J. B., Hansen, H. F., Straarup, E. M., McCullagh, K., Kearney, P., and Kauppinen, S. (2008) Antagonism of microRNA-122 in mice by systemically administered LNA-antimiR leads

to up-regulation of a large set of predicted target mRNAs in the liver, *Nucleic Acids Res 36*, 1153–1162.

64. Kota, J., Chivukula, R. R., O'Donnell, K. A., Wentzel, E. A., Montgomery, C. L., Hwang, H. W., Chang, T. C., Vivekanandan, P., Torbenson, M., Clark, K. R., Mendell, J. R., and Mendell, J. T. (2009) Therapeutic microRNA delivery suppresses tumorigenesis in a murine liver cancer model, *Cell 137*, 1005–1017.
65. Takahashi, K., and Yamanaka, S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors, *Cell 126*, 663–676.
66. Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., Slukvin, II, and Thomson, J. A. (2007) Induced pluripotent stem cell lines derived from human somatic cells, *Science 318*, 1917–1920.
67. Judson, R. L., Babiarz, J. E., Venere, M., and Blelloch, R. (2009) Embryonic stem cell-specific microRNAs promote induced pluripotency, *Nat Biotechnol 27*, 459–461.
68. Lin, S. L., Chang, D. C., Chang-Lin, S., Lin, C. H., Wu, D. T., Chen, D. T., and Ying, S. Y. (2008) Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state, *RNA 14*, 2115–2124.

Chapter 6

Detection of MicroRNAs in Cultured Cells and Paraffin-Embedded Tissue Specimens by In Situ Hybridization

Ashim Gupta and Yin-Yuan Mo

Abstract

Determination of gene expression is essential for understanding the role of a given gene in normal cell growth or disease processes. Recently, newly described microRNAs have been shown to play a key role in the regulation of gene expression; in particular, deregulation of microRNAs is often associated with a variety of human disorders including cancer. Although microRNAs are small RNA molecules with about 20–23 nucleotides in length and detection of their expression is believed to be challenging, with the introduction of modified nucleotides such as locked nucleic acid, the specificity and sensitivity of detection have been greatly improved. There are many methods developed for microRNA detection, but our focus in this chapter is on in situ hybridization (ISH) detection of microRNAs. We have successfully used ISH to detect several microRNAs in paraffin-embedded tumor specimens or cells-cultured in vitro.

Key words: MicroRNA, In situ hybridization, Expression, Locked nucleic acid

1. Introduction

MicroRNAs are a class of naturally occurring small noncoding RNAs that control gene expression by translational repression or degradation of mRNAs (1–4). Since the discovery of *lin-4* and *let-7* in *Caenorhabditis elegans* (5–7), thousands of miRNAs have been identified to date in a variety of organisms. In humans, over 700 microRNAs have been identified (8). A unique feature of gene silencing by microRNAs is that a single microRNA can have multiple targets, and thus, microRNAs as a group can regulate a large number of protein-coding genes. Deregulation of microRNAs can lead to a variety of human disorders including cancer. In this

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_6, © Springer Science+Business Media, LLC 2011

regard, microRNAs may function as oncogenes to suppress corresponding tumor suppressors or as tumor suppressors by targeting corresponding oncogenes. MicroRNA microarray and real-time PCR analyses have indicated that tumor suppressive microRNAs are frequently downregulated, while oncogenic microRNAs are frequently upregulated in tumors compared to the matched normal tissue. However, much less work has been done on microRNA expression in situ which can be detected by in situ hybridization (ISH). A unique advantage of ISH is its capability to detect microRNA expression at the cellular level so that we are able to pinpoint which type of cells express a given microRNA. In addition, it also provides semiquantitative analysis of microRNA expression. Although it is challenging to perform ISH with microRNAs because mature microRNAs are only about 20–23 nucleotides in length, introduction of locked nucleic acid (LNA) has greatly improved the performance of this technique. There are several papers describing this technique for different microRNAs or tissue/cells such as those listed here [9–11] and they may serve as good references. In this chapter, we will describe the ISH procedures used to detect microRNA expression in paraffin-embedded tissue or cells cultured in vitro with examples of miR-145 and miR-380.

2. Materials

2.1. For Cultured Cells

2.1.1. Cell Culture

1. Dulbecco's Modified Eagle's Medium (DMEM) or RPMI-1640 supplemented with 10% fetal bovine serum, 1% l-glutamine, and 1% penicillin/streptomycin.
2. Twelve-well plates.
3. Microscope cover slips (tissue culture coated).
4. Stably transfected MCF-7 cells with ectopic expression of miR-145 or miR-380.

2.1.2. Processing and Fixation

1. Paraformaldehyde: 1 and 4% solution in phosphate-buffered saline (PBS). Store the solutions at 4°C.
2. 1× PBS solution. Store at room temperature.
3. Triton X-100: 0.2% solution in milli-Q water. Store at 4°C (see Note 1).
4. Diethyl pyrocarbonate (DEPC).
5. Hydrogen peroxide (H_2O_2, 30% in water; Fisher Scientific): 3% solution in PBS and store at 4°C (see Note 2).
6. PBS in DEPC–H_2O: Autoclave and store at room temperature.

2.1.3. Hybridization

1. Formamide (molecular biology grade).
2. SSC (20× solution): 0.3 M sodium citrate, and 3.0 M NaCl, in milli-Q water, pH 7.0. Prepare 5×, 2×, and 0.2× SSC solutions by dilution. Store at room temperature.
3. Tween-20 (enzyme grade): 10% solution. Store at room temperature in amber colored bottle.
4. Citric acid (99.5+%, A.C.S reagent): 0.1 M citric acid. Store at room temperature.
5. Heparin (Heparin sodium salt, Grade 1-A, from porcine intestinal mucosa).
6. Yeast RNA (Ambion): Store at –20°C.
7. Probe (DIG oligonucleotide 3′-end labeling kit, second generation, Roche Diagnostics, Indianapolis, IN, USA): 100-pmol oligonucleotide and sterile double-distilled water to final volume of 10 µl. Add 1× reaction buffer (4 µl), 5 mM $CaCl_2$ solution (4 µl), 0.05 mM DIG-ddUTP solution (1 µl), and 20 U/µl terminal transferase (1 µl) on ice. Mix and incubate at 37°C for 15 min and place on ice. Stop the reaction by adding 2 µl 0.2 M EDTA (pH 8.0). Store at –20°C (see Note 3).
8. Hybridization buffer: 50% formamide, 5× SSC, 0.1% Tween-20, 9.2 mM citric acid, 50 µg/ml heparin, pH 6.0, 50 µl/ml yeast RNA (add fresh). Store at room temperature.

2.1.4. TSA Signal Amplification System

1. TNB: 0.1 M Tris–HCl, 0.15 M NaCl, pH 7.5, 0.5% blocking reagent (Perkin Elmer LAS, Boston, MA, USA). Store at –20°C (see Note 4).
2. Anti-DIG-HRP (monoclonal mouse anti-DIG; Jackson Immuno Research Laboratories, USA). Store at –20°C.
3. Secondary anti-mouse-HRP (goat anti-mouse IgG; Invitrogen, Carlsbad, CA, USA): Store at 4°C.
4. DNP amplification reagent (Perkin Elmer LAS, Boston, MA, USA): Store at 4°C (see Note 2).
5. Anti-DNP–HRP conjugate (Perkin Elmer LAS, Boston, MA, USA): Store at 4°C.
6. AEC chromogen (Thermo Scientific) or other HRP-based chromogen: Store at 4°C.
7. TNT buffer: 0.1 M Tris–HCl, 0.15 M NaCl, pH 7.5, 0.05% Tween-20. Store at room temperature (see Note 5).
8. Counterstain – Shandon Harris Hematoxylin – acidified (Thermo Electron Corporation, Anatomocal Pathology, USA): Store at room temperature.
9. HistoMount™ (Invitrogen). Store at room temperature.

2.2. For Paraffin-Embedded Tissue

2.2.1. Processing and Fixation

1. Xylene (histological grade; Fisher Scientific). Store at room temperature.
2. Ethanol (Decon Labs, USA):100, 95, 80, and 50% ethanol solution. Store at room temperature.
3. DEPC–H_2O: 0.1% DEPC in milli-Q water. Autoclave the solution. Store at room temperature.
4. 1× PBS. Store at room temperature.
5. PBS in DEPC–H_2O: Autoclave and store at room temperature.
6. Hydrogen peroxide (H_2O_2, 30% in water; Fisher Scientific): 3% solution in PBS and store at 4°C (see Note 2).
7. Proteinase K (Invitrogen): Store at –20°C (see Note 6).
8. Glycine (USB Corporation, Cleveland, OH, USA): Prepare 0.2% solution. Store at room temperature.
9. Paraformaldehyde (Sigma-Aldrich, MO, USA): Prepare 4% solution in PBS. Store at 4°C.

2.2.2. Hybridization

1. Formamide (molecular biology grade; Fisher Biotech).
2. SSC (20× solution): 0.3 M sodium citrate (Fisher) and 3.0 M NaCl in milli-Q water, pH 7.0. Prepare 5×, 2×, and 0.2× SSC solutions by dilution. Store at room temperature.
3. Tween-20 (enzyme grade; Fisher): 10% solution. Store at room temperature in amber colored bottle.
4. Citric acid (99.5+%, A.C.S reagent; Sigma-Aldrich): 0.1 M citric acid. Store at room temperature.
5. Heparin (Heparin sodium salt, Grade 1-A, from porcine intestinal mucosa; Sigma-Aldrich).
6. Yeast RNA (Ambion): Store at –20°C.
7. Probe (DIG oligonucleotide 3′-end labeling kit, second generation; Roche Diagnostics, Indianapolis, IN, USA): 100 pmol oligonucleotide and sterile double-distilled water to final volume of 10 µl. Add 1× reaction buffer (4 µl), 5 mM $CaCl_2$ solution (4 µl), 0.05 mM DIG-ddUTP solution (1 µl), and 20 U/µl terminal transferase (1 µl) on ice. Mix and incubate at 37°C for 15 min and place on ice. Stop the reaction by adding 2 µl 0.2 M EDTA (pH 8.0). Store at –20°C (see Note 3).
8. Hybridization buffer: 50% formamide, 5× SSC, 0.1% Tween-20, 9.2 mM citric acid, 50 µg/ml heparin, pH 6.0, 50 µl/ml yeast RNA (add fresh).

2.2.3. TSA Signal Amplification System

1. TNB: 0.1 M Tris–HCl, 0.15 M NaCl, pH 7.5, 0.5% blocking reagent (Perkin Elmer LAS, Boston, MA, USA). Store at –20°C (see Note 4).

2. Anti-DIG–HRP (monoclonal mouse anti-DIG; Jackson Immuno Research Laboratories, USA). Store at –20°C.
3. Secondary anti-mouse-HRP (goat anti-mouse IgG; Invitrogen, Carlsbad, CA, USA): Store at 4°C.
4. DNP amplification reagent (Perkin Elmer LAS, Boston, MA, USA). Store at 4°C.
5. Anti-DNP–HRP conjugate (Perkin Elmer LAS, Boston, MA, USA). Store at 4°C.
6. AEC chromogen (Thermo Scientific) or other HRP-based chromogen. Store at 4°C.
7. TNT buffer: 0.1 M Tris–HCl, 0.15 M NaCl, pH 7.5, 0.05% Tween-20. Store at room temperature (see Note 5).
8. Counterstain – Shandon Harris Hematoxylin – acidified (Thermo Electron Corporation, Anatomocal Pathology, USA). Store at room temperature.
9. HistoMount™ (Invitrogen). Store at room temperature.

3. Methods

The ISH method allows us to visualize microRNA expression at a higher spatial resolution. However, a challenge of detection of microRNAs by ISH is their small size. LNAs are a class of bicyclic high-affinity RNA analogs in which the furanose ring in the sugar-phosphate backbone is chemically locked in a conformation mimicking the North-type (C3′-endo) conformation of RNA [12, 13]. This results in an unprecedented hybridization affinity of LNA toward complementary single-stranded RNA molecules. Recently, LNA-modified DNA probes have been used to determine the spatiotemporal expression patterns of microRNAs in zebrafish and mouse embryos by whole-mount ISH [10, 14] as well as in archived formalin fixed, paraffin-embedded (FFPE) sections of brain tissue [15]. Here, we used 3′ DIG-labeled, LNA-modified DNA oligonucleotide probes complementary to the entire mature sequence of a subset of selected microRNAs. Following hybridization of the LNA probes to the FFPE sections or cultured cells, a tyramide signal amplification reaction was carried out [16] using horseradish peroxidase conjugated to anti-DIG antibodies.

3.1. For Cultured Cells

3.1.1. Preparation of Plate

1. Place sterile cover slips in a 12-well plate.
2. Seed the cells (stable clones or transiently transfected cells or normal cells) in the plate.

3.1.2. Processing and Fixation

1. Cells are at 40–80% confluent before fixation.
2. Aspirate the medium and wash the cells with 1× PBS (500 µl for each).
3. Fix the cells with 1% paraformaldehyde solution (500 µl for each, see Note 7) for 10 min at room temperature.
4. Wash the cells with 1× PBS (500 µl for each) for 5 min on ice (twice).
5. Aspirate the PBS and add 0.2% Triton X-100 (500 µl for each) for 20 min on ice.
6. Wash the cells with 1× PBS (500 µl for each) for 5 min on ice (twice).
7. Incubate with 3% H_2O_2 (500 µl for each) in PBS for 15 min at room temperature.
8. Wash the cells with PBS in DEPC–H_2O (500 µl for each) for 5 min at room temperature (twice).
9. Fix the cells in 4% paraformaldehyde solution (500 µl for each) for 10 min at room temperature.
10. Wash the cells with 1× PBS (500 µl for each) for 5 min at room temperature (twice).

3.1.3. Hybridization

1. Prehybridization: Add hybridization buffer (100 µl for each) without probe and incubate it in a hybridization oven at 37°C for 2–3 h.
2. Add 6-pmol probe in 100 µl of hybridization buffer (for each) and place in a moist chamber and incubate it at 37°C overnight (see Note 8).
3. Soak the cells in 5× SSC (500 µl for each) at room temperature and place it on a shaker for 10 min.
4. Wash the cells with 2× SSC (500 µl for each) at 37°C for 30 min.
5. Wash the cells with 0.2× SSC (500 µl for each) at 37°C for 30 min.
6. Wash the cells with 1× PBS (500 µl for each) for 5 min at room temperature (twice).

3.1.4. TSA Signal Amplification System

1. Block the cells on the cover slip for 30 min in TNB buffer (100 µl for each) at room temperature.
2. Incubate with anti-DIG–HRP (1:50 in TNB buffer) for 2 h in a moist chamber at room temperature (see Notes 9 and 10).
3. Wash with TNT buffer (500 µl for each) for 5 min at room temperature (twice).
4. Incubate with secondary anti-mouse-HRP (1:100 in TNB buffer) for 30–40 min in a moist chamber at room temperature.

5. Wash with TNT buffer (500 µl for each) for 5 min at room temperature (twice).
6. Incubate with DNP amplification reagent working solution (1:50 in amplification diluent) for 10 min in a moist chamber at room temperature.
7. Wash with TNT buffer (500 µl for each) for 5 min at room temperature (twice).
8. Incubate with anti-DNP–HRP (1:100 in TNB buffer) for 40–45 min in a moist chamber at room temperature.
9. Wash with TNT buffer (500 µl for each) for 5 min at room temperature (twice).
10. Add AEC chromogen (1–2 drops) or other HRP-based chromogen for signal development for 30–40 min at room temperature.
11. Wash with TNT buffer (500 µl for each) for 5 min at room temperature (twice).
12. Counterstain and mount.
13. Examples of miR-145 and miR-380 expression in MCF-7 cells are shown in Fig. 1.

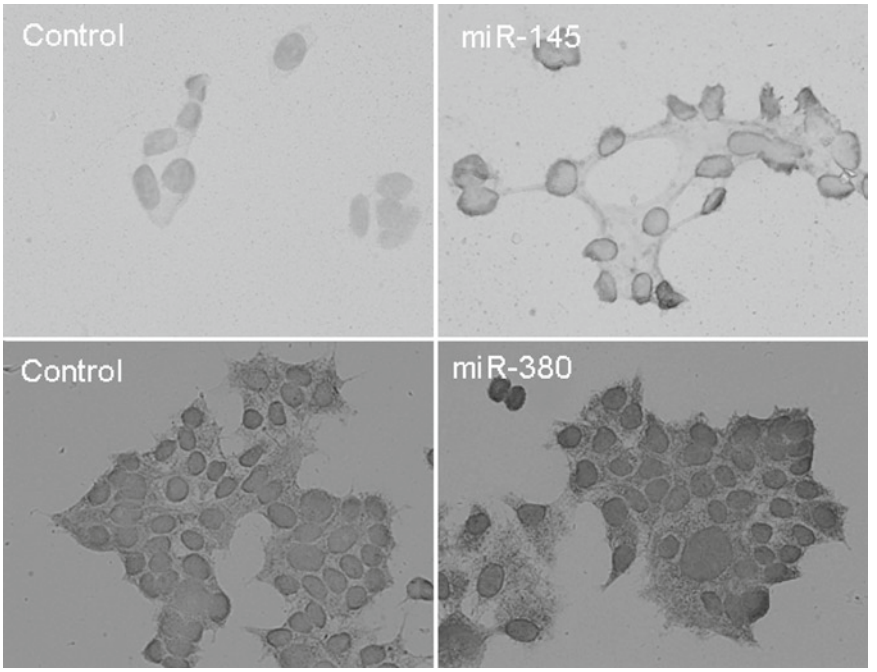


Fig. 1. Detection of microRNA expression in cultured MCF-7 cells. Cells were infected with vector control or miR-145 or miR-380 and cultured overnight in a coated cover slip . In situ hybridization described in the protocol was used to detect their expression. It is apparent that most microRNA signals were localized in the cytoplasm and relative signal intensity is visibly obvious between vector control and microRNA, which is consistent with the real-time PCR results showing 3–4 Ct values difference between vector and miR-145 or miR-380.

3.2. For Paraffin-Embedded Tissue

3.2.1. Processing and Fixation

1. Heat the slides to 60–65°C for 30 min in a hybridization oven (2 h preheated before).
2. Leave the slides in a xylene solution for 12 min at room temperature (three times).
3. Leave the slides in 100% ethanol for 5–10 min at room temperature.
4. Leave the slides in 95% ethanol for 5–10 min at room temperature.
5. Leave the slides in 80% ethanol for 5–10 min at room temperature.
6. Leave the slides in 50% ethanol for 5–10 min at room temperature.
7. Wash the slides with DEPC–H_2O for 5 min at room temperature.
8. Wash the slides with PBS in DEPC–H_2O for 5 min at room temperature (twice). (Prepare the Proteinase K at the same time.)
9. Incubate the slides with 3% H_2O_2 in PBS for 15 min at room temperature.
10. Wash the slides with PBS in DEPC–H_2O for 5 min at room temperature (twice).
11. Proteinase K treatment: Make fresh proteinase K solution in PBS in DEPC–H_2O (10–20 μg/ml), prewarmed at 37°C, and then incubate the slides for 8–10 min at room temperature.
12. Add 0.2% glycine solution for 30 s at room temperature.
13. Wash the slides with PBS for 2 min at room temperature (twice).
14. Fix the slides with 4% paraformaldehyde solution (100 μl for each) for 10 min at room temperature.
15. Wash the slides with PBS in DEPC–H_2O for 5 min at room temperature (twice).

3.2.2. Hybridization

1. Prehybridization: Add hybridization buffer (100 μl for each) without probe and incubate it in a hybridization oven at 37°C for 2–3 h.
2. Add 6-pmol probe in 100 μl of hybridization buffer (for each) and place in a moist chamber and incubate it at 37°C overnight (see Note 8).
3. Soak the slides in 5× SSC at room temperature and leave it on a shaker for 10 min.
4. Wash the slides with 2× SSC at 37°C for 30 min.
5. Wash the slides with 0.2× SSC at 37°C for 30 min.
6. Wash the slides with 1× PBS for 5 min at room temperature (twice).

3.2.3. TSA Signal Amplification System

1. Block the slides for 30 min in TNB buffer (100 µl for each) at room temperature.
2. Incubate with anti-DIG–HRP (1:50 in TNB buffer) for 2 h in a moist chamber at room temperature (see Notes 9 and 10).
3. Wash with TNT buffer for 5 min at room temperature (twice).
4. Incubate with secondary anti-mouse-HRP (1:100 in TNB buffer) for 30–40 min in a moist chamber at room temperature.
5. Wash with TNT buffer for 5 min at room temperature (twice).
6. Incubate with DNP amplification reagent working solution (1:50 in amplification diluent) for 10 min in a moist chamber at room temperature.
7. Wash with TNT buffer for 5 min at room temperature (twice).
8. Incubate with anti-DNP–HRP (1:100 in TNB buffer) for 40–45 min in a moist chamber at room temperature.
9. Wash with TNT buffer for 5 min at room temperature (twice).
10. Add AEC chromogen (1–2 drops) or other HRP-based chromogen for signal development for 30–40 min at room temperature.
11. Wash with TNT buffer for 5 min at room temperature (twice).
12. Counterstain and mount.
13. An example of miR-380 expression in breast tumor and the matched normal breast tissue is shown in Fig. 2.

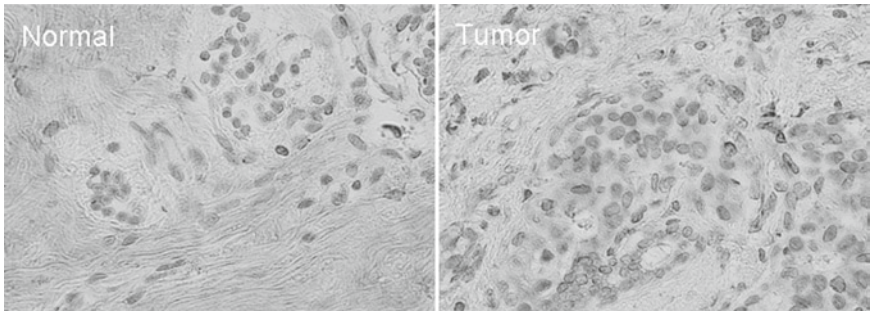


Fig. 2. Expression of miR-380 in matched breast tumor tissue specimens. The level of miR-380 is higher in tumor than in matched normal tissue. Similar to the signals in cultured cells, miR-380 is mainly expressed in the cytoplasm.

4. Notes

1. Triton X-100 is used for cell permeabilization. Several other agents like saponins can also be used.
2. DNP amplification reagent is catalyzed by HRP. To minimize the background, endogenous peroxidase activity if present, must be quenched. This can be done by using 0.3–3% H_2O_2; methanol or PBS as diluents for H_2O_2.
3. It is *not* recommended to increase the amount of oligonucleotide in labeling reaction. Larger amounts of nucleotides can be labeled by increasing the reaction volume and all components proportionally and increasing the incubation time to 1 h.
4. Add the blocking reagent slowly in small increments to buffer while stirring. Heat gradually to 60°C with continuous stirring to completely dissolve the blocking reagent. Aliquot and store at –20°C for long-term use and discard any blocking buffer that has been stored for more than 24 h at room temperature.
5. Other wash buffers such as PBS or substitution of 0.3% Triton X-100 for 0.05% Tween-20 can also be used.
6. Proteinase K is used as a protein digestant to increase reagent penetration prior to probe hybridization. Other protein digestants like 0.005–0.1% pepsin in 0.01 M HCl may be used.
7. The reagent volumes used should be sufficient to completely cover cells or tissue sections on the slides. If larger volumes are required to cover the samples changes should be made in the protocol.
8. Run control slides with each experiment. It includes an *unamplified* control slide or an amplified slide *without* probe.
9. Do not let slides dry out between the steps and all incubations should be done in a humidified/moist chamber (i.e., a damp paper towel in a covered box).
10. Drain off as much of the incubation solution as possible, before the addition of the next solution, to prevent reagent dilution and uneven staining.

Acknowledgment

This work was supported by grant CA102630 from NCI.

References

1. Pillai, R. S. (2005) MicroRNA function: multiple mechanisms for a tiny RNA? *RNA 11*, 175361.
2. Zamore, P. D. and Haley, B. (2005) Ribognome: the big world of small RNAs. *Science 309*, 1519–24.
3. Bartel, D. P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell 116*, 281–97.
4. Ambros, V. (2004) The functions of animal microRNAs. *Nature 431*, 350–5.
5. Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell 75*, 843–54.
6. Wightman, B., Ha, I., and Ruvkun, G. (1993) Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell 75*, 855–62.
7. Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature 403*, 901–6.
8. Griffiths-Jones, S. (2004) The microRNA Registry. *Nucleic Acids Res 32*, D109–11.
9. Yamamichi, N., Shimomura, R., Inada, K., Sakurai, K., Haraguchi, T., Ozaki, Y., Fujita, S., Mizutani, T., Furukawa, C., Fujishiro, M., Ichinose, M., Shiogama, K., Tsutsumi, Y., Omata, M., and Iba, H. (2009) Locked nucleic acid in situ hybridization analysis of miR-21 expression during colorectal cancer development. *Clin Cancer Res 15*, 4009–16.
10. Kloosterman, W. P., Wienholds, E., de Bruijn, E., Kauppinen, S., and Plasterk, R. H. (2006) In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. *Nat Methods 3*, 27–9.
11. Pena, J. T., Sohn-Lee, C., Rouhanifard, S. H., Ludwig, J., Hafner, M., Mihailovic, A., Lim, C., Holoch, D., Berninger, P., Zavolan, M., and Tuschl, T. (2009) miRNA in situ hybridization in formaldehyde and EDC-fixed tissues. *Nat Methods 6*, 139–41.
12. Kauppinen, S., Vester, B., and Wengel, J. (2006) Locked nucleic acid: high-affinity targeting of complementary RNA for RNomics. *Handb Exp Pharmacol 173*, 405–22.
13. Koshkin, A. A. and Wengel, J. (1998) Synthesis of novel 2′,3′-linked bicyclic thymine ribonucleosides. *J Org Chem 63*, 2778–81.
14. Wienholds, E., Kloosterman, W. P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., Horvitz, H. R., Kauppinen, S., and Plasterk, R. H. (2005) MicroRNA expression in zebrafish embryonic development. *Science 309*, 310–1.
15. Nelson, P. T., Baldwin, D. A., Kloosterman, W. P., Kauppinen, S., Plasterk, R. H., and Mourelatos, Z. (2006) RAKE and LNA-ISH reveal microRNA expression and localization in archival human brain. *RNA 12*, 187–91.
16. van Gijlswijk, R. P., Wiegant, J., Vervenne, R., Lasan, R., Tanke, H. J., and Raap, A. K. (1996) Horseradish peroxidase-labeled oligonucleotides and fluorescent tyramides for rapid detection of chromosome-specific repeat sequences. *Cytogenet Cell Genet 75*, 258–62.

Chapter 7

MicroRNA Northern Blotting, Precursor Cloning, and Ago2-Improved RNA Interference

Julia Winter and Sven Diederichs

Abstract

Aberrant expression of microRNAs (miRNAs) and processing defects in their biogenesis pathway are a widespread phenomenon in tumors, conveying great importance to the analysis of miRNA expression, regulation, and biogenesis to gain knowledge about their role in cancer. Besides Drosha and Dicer, Argonaute proteins are key players in miRNA processing. In addition to their role as components of the RNA-induced silencing complex (RISC) executing target silencing, Argonautes mediate posttranscriptional regulation of miRNA maturation by creating an additional intermediate processing step, the Ago2-cleaved precursor miRNA (ac-pre-miRNA), and enhancing the production or stability of mature miRNAs. Here, we describe the detection of miRNA levels by Northern blotting and the identification of the 3′ end of miRNAs by precursor cloning to accentuate two of the many roles of Argonaute proteins. In addition, we describe a method to optimize RNAi experiments by increasing the efficacy and specificity of target silencing via Ago2 cotransfection.

Key words: MicroRNA, miRNA, Biogenesis, Processing, MicroRNA precursor cloning, Northern blotting, Northern blot, Cancer, Tumor, Posttranscriptional regulation, Ago2, Argonaute-2, RNAi, RNA interference

1. Introduction

Analysis of microRNA (miRNA) expression, regulation, and biogenesis is an essential tool for understanding its role in cancer. Genome-wide profiling uncovered a general downregulation of miRNAs in tumors compared to normal tissue (1). Processing defects could result in equivalent precursor levels, coinciding with varying mature miRNA levels (2–5), and miRNA expression levels are even suitable as prognostic and diagnostic markers (2). Therefore, determination of miRNA expression levels is a central approach to elucidate their regulation and function. Northern blotting has been applied to study mRNA expression since the

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_7, © Springer Science+Business Media, LLC 2011

earliest days of miRNA research (6–9) and remains one crucial method to determine miRNA expression and size. Additionally, to obtain sequence information, a method to clone precursors and subsequently determine the 3′ end of miRNAs was established (10). Both methods are described in the following paragraphs accentuating two of the many roles of Argonaute proteins in miRNA biogenesis and regulation (Figs. 1 and 2).

Next to the RNases Drosha and Dicer, Argonaute (Ago) proteins are major players in miRNA biogenesis and function. Present in all higher eukaryotes, many organisms encode multiple members of these evolutionarily conserved, ubiquitously expressed proteins (11). Ago proteins are organized in three domains: while the PAZ domain anchors the characteristic two nucleotide (nt) 3′ overhang (12, 13), and the highly basic pocket of the MID domain specifically binds the 5′ phosphate of a small RNA (14, 15), the P-element induced wimpy testis (PIWI) domain contains a catalytic triad (Asp 597/Asp 669/His 807) that possesses endonuclease activity (16, 17). Despite the fact that both, human Ago2 and Ago3, contain this catalytic triad, only human Ago2 mediates slicing to cleave the target mRNA. Additional factors can modulate the activity of Ago proteins, for example, posttranslational modification or interaction with specific proteins (18). Ago proteins are not only involved in RNA-induced silencing complex (RISC) formation and target silencing, but they also play an important role in posttranscriptional regulation of miRNA biogenesis (as reviewed in Chapter 1) (5, 18). As components of the RISC, they execute target silencing via target cleavage, translational inhibition, or mRNA decay (19, 20). Additionally, Ago proteins affect miRNA biogenesis and their regulation at a posttranscriptional level. Ago2 cleaves the 3′ arm in the middle of some highly complementary pre-miRNA hairpins to generate the Ago2-cleaved precursor miRNA (ac-pre-miRNA), an intermediate processing that is processed by Dicer as efficiently as noncleaved miRNAs (10). This cleavage probably facilitates strand separation by affecting the thermodynamic stability as already shown for siRNAs (small interfering RNAs) or contributes to strand selection since consequently only the 5′ arm of the ac-pre-miRNA can mature into the 22-nt miRNA guide strand (21–24).

Ago2 can also increase the efficacy and specificity of target silencing. Since this effect is restricted to perfectly matched binding sites and does not influence endogenous RNAi activity, it has great potential to optimize RNAi experiments (25). Thus, insights into basic mechanisms of miRNA biogenesis can lead to improved methodology for the analysis of cancer cells. MiRNA biogenesis factors possess the potency to enhance the efficacy of RNAi. Coexpression of Exportin-5 or Ago2 increases RNAi activity mediated by ectopic miRNAs, siRNAs, or short hairpin RNAs (shRNAs) (Fig. 3) (25, 26). Hence, we also describe a method how to use Ago2 coexpression to enhance RNAi for functional studies of cancer cells.

2. Materials

2.1. Northern Blotting

2.1.1. Buffers and Reagents

1. Cells in culture, lysates, or total RNA (integrity verified, isolated using a method that preserves the miRNA fraction): To prevent degradation, lysates and total RNA should be stored at –80°C (see Note 1).
2. Trizol: Store below room temperature (2–8°C) and protect from light.
3. Diethylpyrocarbonate (DEPC)-treated H_2O: 1 ml DEPC is dissolved in 1 L ddH_2O over night at room temperature and subsequently autoclaved to inactivate DEPC. Store DEPC at 4°C and autoclaved DEPC-treated ddH_2O at room temperature (see Note 2).
4. ssDNA oligonucleotide fully complementary to mature miRNA: Guidelines for primer design, see Subheading 3.1.6 "probe design". Dissolve desalted oligonucleotide in nuclease-free ddH_2O and adjust concentration to 10 µM, store at –20°C.
5. T4 Polynucleotide kinase (see Note 3).
6. EasyTides [γ-^{32}P] Adenosine 5′-triphosphate (Perkin-Elmer): For working with radioactivity, please observe the guidelines and safety measures of your institution. Always ensure appropriate shielding and distance to the radioactive source and plan your experiment in advance to ensure minimal expositions. Store [γ-^{32}P]ATP at 4°C or according to the recommendations of the manufacturer (see Note 4).
7. Running buffer: 0.5× TBE. Dilute 50 ml 10× TBE in 1 l DEPC-ddH_2O. Store at room temperature (see Note 2).
8. RNA loading dye (3×): 8 M urea, 1× TBE, 30 mM EDTA, 20% glycerol, and Bromophenol blue (BPB) in nuclease-free ddH_2O. Aliquot and store at –20°C. BPB is a small, negatively charged tracking dye that moves to the anode more rapidly than nucleic acids.
9. Decade Marker (Ambion): Use according to manufacturer's manual. Store at –20°C.
10. Ethidium bromide (EtBr)-staining solution (1:2,500 EtBr in 0.5× TBE): Dissolve 4 µg EtBr per ml 0.5× TBE buffer. Store the light-sensitive solution in a dark container at room temperature. Always wear nitrile gloves to avoid contact with the cancerogenic material.
11. Hybridization buffer: ExpressHyb (Clontech). Prior to first usage, preheat buffer to 68°C to dissolve any precipitates and store at room temperature. Heat to 37°C before each prehybridization.
12. Set up 20× SSC: Dissolve 3 M sodium chloride (NaCl) and 0.3 M sodium citrate ($C_6H_5Na_3O_7$) in ddH_2O and adjust pH

to 7.0 using HCl, prior to DEPC treatment (as described above). Use this solution to prepare the wash buffers that contain sodium dodecyl sulfate (SDS) and consequently cannot be treated with DEPC. Store at room temperature (see Note 2).

13. Wash buffer #1 (2× SSC, 0.05% SDS): Dilute 100 ml 20× SSC and 5 ml 10% SDS with 895 ml ddH_2O. Do not treat with DEPC and store at room temperature (see Note 2).
14. Wash buffer #2 (0.1× SSC, 0.1% SDS): Dilute 5 ml 20× SSC and 10 ml 10% SDS with 985 ml ddH_2O. Do not treat with DEPC and store at room temperature (see Note 2).

2.1.2. Equipment

1. 15% TBE urea gel: Store at 4°C. Alternatively, urea gels can be prepared from 10× TBE, polyacrylamide, and the appropriate amount of urea dissolved in ddH_2O at 37°C. Polymerization is catalyzed by adding 100-µl tetramethylethylenediamine (TEMED) and 6-µl ammonium persulfate (APS) per 10 ml buffer.
2. Filter paper extra thick (7.5 cm × 10 cm).
3. Hybond-N^+ Nylon membrane.
4. MicroSpin G-25 columns.

2.1.3. Hardware

1. Chamber for semidry electroblotting.
2. Crosslinker, for example, UV-Stratalinker 1800 (Stratagene).
3. Heat block.
4. Hybridization oven.
5. Autoradiograpy cassette including intensifying screen.
6. Kodak BioMax MS Film.

2.2. Determining the 3′ End of Precursor miRNAs

1. Cells in culture, lysates, or total RNA (integrity verified): To prevent degradation, lysates and total RNA should be stored at –80°C.

2.2.1. RNA Isolation

2. miRVana-miRNA isolation kit to specifically enrich RNA fragments that are smaller than 200 nt. Store buffers at 4°C (see Note 8).

2.2.2. Reverse Transcription and Amplification

1. Linker-2 (IDT): Dissolve in ddH_2O to 10 µM and store at –20°C. [5′rAPPCACTCGGGCACCAAGGA/3′ddC].
2. Forward primer complementary to miRNA of interest including CACC-overhang at its 5′ end (necessary for directional TOPO cloning). Dissolve desalted oligonucleotide in ddH_2O and adjust concentration to 10 µM (see Note 9).
3. Reverse primer complementary to linker-2 for specific reverse transcription. Dissolve desalted oligonucleotide in ddH_2O and adjust concentration to 10 µM.

4. T4 RNA ligase I (see Note 3).
5. T4 RNA ligase I buffer depleted of ATP (50 mM Tris–HCl, 10 mM $MgCl_2$, 10 mM Dithiothreitol, pH 7.8). Store at –20°C (see Note 10).
6. Thermoscript Reverse Transcriptase (see Note 3).
7. Pfu Turbo Polymerase (see Note 3).
8. Dimethylsulfoxide (DMSO): This chemical compound is used in PCR to minimize interfering reactions that inhibit secondary structures in the primer or DNA template. However, usage also decreases the mutation rate (27). Store at –20°C and thaw at room temperature 20 min prior to usage.
9. Nuclease-free dNTPs (10 mM each). Store at –20°C.

2.2.3. Cloning

1. pcDNA3.1 Directional TOPO Expression kit (Invitrogen): Store the vector and the salt solution at –20°C. To guarantee efficient transformation, the provided competent *Escherichia coli* cells should be stored at –80°C and prior to the heat shock reaction, exposure to room temperature should strictly be avoided.
2. Restriction enzymes to test for positive insertion of the template (e.g., XhoI, *Hin*dIII): Choose a combination of enzymes that enclose the insertion site of the DNA template so that cleavage will cut out the insert and positive insertion can be visualized on an agarose gel (see Note 3).
3. 4% Agarose gel: Dissolve 4 g agarose in 100 ml 1× TBE in a microwave and add 4 µl EtBr before pouring the gel into an appropriate chamber (see Note 11).
4. 10 bp DNA ladder. Aliquot and store at –20°C.

2.3. Improving RNA Interference with Argonaute-2

2.3.1. Ago2 Expression Vector

For codelivery of Ago2 along with the RNAi reagent of choice (see Subheading 2.3.2), an expression plasmid for Ago2 is necessary (see Note 12). Additionally, lentiviral vectors have also been successfully used for the delivery of Ago2 into hard-to-transfect cell lines (25).

2.3.2. RNAi Reagent: shRNA/siRNA/dsRNA

As the specificity component for a RNA interference experiment, either a short hairpin RNA (shRNA), a small interfering RNA (siRNA), or a double-stranded (dsRNA) can be used to target a specific gene for knockdown (see Note 13). The shRNA is usually expressed from a plasmid or lentiviral construct using a Pol II or Pol III-driven promoter. The primary transcript of the shRNA is then processed by the endogenous miRNA machinery into the mature, active short RNA molecule. The siRNA is normally transfected as a short duplex of about 21 bp of RNA with or without any modifications.

2.3.3. Standard Transfection or Infection Reagents

Depending on the individual experimental design, cells and transfection reagents or a viral infection protocol will be necessary. Standard protocols and recommendations of the manufacturer can be applied.

3. Methods

3.1. Northern Blotting (Fig. 1)

3.1.1. Isolating Total RNA

Total RNA is isolated using the Guanidium thiocyanate–phenol–chloroform extraction method (Trizol) according to manufacturer's instruction (28). Cells are lysed in a monophasic solution of the chaotropic guanidine isothiocyanate and phenol to subsequently denature and inactivate proteins including RNases. This ensures the isolation of high-quality RNA and prevents degradation (see Note 1). Addition of chloroform generates a second phase into which DNA and proteins are extracted, so that the upper phase contains pure RNA that can subsequently be recovered by sequential ethanol precipitations at –20°C.

RNA concentration and integrity is determined by gel electrophoresis and photometry. As total RNA only contains ~5% mRNA but consists of ~90% rRNA, distinguished bands of 28S (~4.8 kb) and 18S-rRNA (~1.8 kb) in the ratio of 2:1 are detected if the RNA is not degraded. Information about protein or organic contaminations are determined via photometric determination. RNA absorbs at 260 nm whereas the absorption maximum of proteins lies at 280 nm and organic contaminants absorb at 230 nm. OD 260:280 ratios of 1.8–2 account for pure RNA while the OD 260:230 ratio should ideally be higher than 2.0 to exclude organic contaminations.

3.1.2. Preparing RNA Samples for Electrophoretic Separation in Urea Polyacrylamide Gels

Equal amounts of total RNA (ideally 5–30 μg) are resuspended in RNA loading buffer, heated to 80°C for 10 min, and subsequently kept on ice (see Notes 4 and 6). This ensures effective denaturation that removes secondary structures to separate the sample depending on the size of the RNA.

3.1.3. Size-Separation in Urea Polyacrylamide Gels

Formaldehyde agarose gels illustrate the well-established method of size separation for Northern blots. Since miRNAs are very small molecules (mature miRNAs consist of 20–23 nt, precursor miRNAs approximately 70 nt), a very stringent separation is needed. Therefore, 15% urea polyacrylamide gels with much smaller pores are the method of choice for Northern blots to detect miRNAs. To avoid secondary structures, urea is added as a denaturing reagent.

Prior to loading the samples on the gel, a prerun for 45 min at 125 V in running buffer is indispensable to get rid of the excessive urea which is then flushed out of the wells additionally. Subsequently, the RNA is loaded into the wells, and the run proceeds for additional 90 min at 125 V. In some cases, an RNA

ladder might be of help but as the RNA loading buffer contains Bromophenol blue (BPB), which is in concordance with the 10-nt front, it is also sufficient to stop the run when the BPB band reaches the end of the gel.

3.1.4. Checking for Equivalent Loading of Total RNA

The majority of experimental approaches aim at comparing miRNA levels of different samples. Therefore, loading equal amounts of total RNA in each well is of special importance. To demonstrate equivalent levels, the blot can be stained with a probe for a loading control (such as U6 snRNA or 5S rRNA, so-called housekeeping RNAs) (see Note 6). To perform a quick check prior to blotting, the gel is stained in an EtBr solution for 5 min to visualize the levels of tRNA (70–90 nt), 5SrRNA (120 nt), and 5.8S rRNA (160 nt) using UV light (Fig. 1) (see Note 1).

3.1.5. Blotting RNA onto a Nylon Membrane

While the polyacrylamide gel is running, two filter papers and one piece of membrane (9 cm × 6 cm) per gel are soaked in 0.5× TBE and chilled at 4°C to equilibrate it for the subsequent transfer. The negatively charged RNA on the gel is then transferred onto a positively charged nylon membrane by semidry electroblotting at 400 mA for 60 min. To avoid overheating the chamber, 25 V should not be exceeded. In this case, transferring at 250 mA for 3 h would be the method of choice.

Subsequently after the transfer, the membrane is crosslinked twice at optimal settings (=120 mJ and below) to covalently fix the RNA on the nylon membrane.

3.1.6. Probe Design and Probe Labeling

MiRNA expression is detected via radioactively labeled probes that are complementary to the mature miRNA. These probes are labeled at the 5′ end using T4 Polynucleotide kinase that catalyzes the transfer of the γ-phosphate from [γ-^{32}P]ATP to the 5′ hydroxyl group of the DNA probe. Therefore, 2 µl 10× PNK buffer, 20 pmol probe, and 20 µCi [γ-^{32}P]ATP are mixed with 10 U T4-Polynucleotide kinase, and the 20-µl reaction is incubated at 37°C for 60 min (see Notes 3 and 4). The kinase reaction is heat inactivated at 68°C for 10 min, and 30-µl nuclease-free H_2O is added before salts and excessive radioactivity are removed using G-25 microspin columns according to manufacturer's manual.

3.1.7. Hybridization

To ease the experimental setup, it is recommended to perform the hybridization steps in a 50-ml tube that can easily be put into hybridization bottles and afterward be discarded (see Note 5). Therefore, the blot is inserted with the RNA inside, not adhering to the wall of the tube. After 30 min of prehybridization in 5-ml ExpressHyb buffer at 37°C, the labeled probe is added without touching the blot. Hybridization takes place at 37°C for 60 min rotating in a hybridization oven (see Note 7).

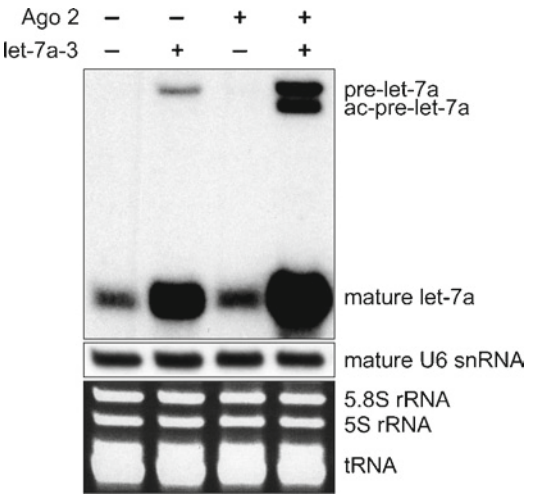


Fig. 1. *Northern blotting detects precursor and mature let-7a.* Northern blot analysis of 293 cells cotransfected with Ago2 give rise to increased expression of miRNAs. Elevated miRNA levels permit detection of an additional, smaller precursor band, the ac-pre-miRNA (Ago2-cleaved precursor microRNA). Reprobing of the blot for U6 snRNA served as a loading control. As additional loading control, EtBr staining was used to visualize tRNAs, 5S and 5.8S rRNAs under UV light.

3.1.8. Washing and Detection

Since an excess of labeled probe would arise in unspecific signals, the unbound radioactivity is removed by washing the blot twice with 25-ml wash buffer #1 for 5 min at 37°C rotating in the hybridization oven and once with 50-ml wash buffer #2 for 10 min at room temperature on a shaker (see Note 7). Afterward, the membrane is wrapped into saran wrap and put onto a film into an autoradiography cassette in the dark room. For weakly expressed miRNAs, an intensifying screen below the film can improve the signal. After a certain time of incubation at –80°C, the film is thawed at 37°C and exposed to detect the radiation signal. Incubation times can vary between 12 h and several weeks and depend on the expression level of the miRNA that should be detected. Alternatively, a phosphoimager can be used for more rapid detection. Incubation times of 30–60 min are then sufficient for effective visualization.

3.1.9. Stripping the Membrane to Remove Labeled Probe

In many cases, it is desired to strip and reprobe the blot with a further probe to detect an additional miRNA or U6 snRNA or 5S rRNA as loading controls in the same RNA samples (see Note 6). Therefore, 50 ml of 0.5% boiling SDS are added to the blot and the blot is left shaking at room temperature for 30 min. This reaction breaks the interaction between the labeled probe and the RNA, while the covalent bondage between RNA and membrane is not affected. The blot can now be stored at –20°C or be hybridized with another radioactively labeled probe after 30 min of prehybridization at 37°C in ExpressHyb buffer (see above).

3.2. Determining the 3′ End of Precursor miRNAs (Fig. 2)

3.2.1. Isolating RNA that is Specifically Enriched in RNAs Smaller than 200 nt

To clone precursor miRNAs, only a fraction of small RNAs (smaller than 200 nt) is required, not the total RNA of a cell. Since mature miRNAs are 20–23 nt and precursor miRNAs approximately 70 nt in length, larger RNA fragments would only disturb the reaction. Due to size exclusion, miRNAs and their precursors are highly enriched. To specifically isolate RNA fragments that are smaller than 200 nt, various commercially available kits can be used. Cells are first lysed, then phenol–chloroform extracted and subsequently purified using glass-fiber filters. For the first reaction, 25% ethanol is used. Only large RNAs are immobilized and the small RNA species are collected in the filtrate. In the second round, the ethanol concentration is increased to 55%, and the filtrate is passed through a second glass-fiber filter where the small RNAs become immobilized and can be eluted into H_2O after several wash steps.

3.2.2. Linker Ligation and cDNA Synthesis

As the 3′ end should be determined in this experiment, it is not possible to design a gene-specific reverse primer (as the sequence of the pre-miRNA end first needs to be elucidated). To solve this problem, an RNA linker that contains a 3′ terminal dideoxy-C base to prevent self-ligation, is ligated randomly to the pool of small RNAs at their 3′ ends. A maximum of 5 µg RNA, 5 µM Linker-2, T4 RNA Ligase buffer without ATP, 10% DMSO, and 40 U of T4 RNA Ligase 1 are mixed in a 20-µl reaction and incubated at 37°C for 1 h before the reaction is heat inactivated at 65°C for 15 min (see Note 10). To remove all reagents, the RNA is subsequently ethanol precipitated (–20°C for 10 min) and reverse transcribed into cDNA using Thermoscript, which has especially been engineered for reduced RNase H activity and higher thermal stability according to the manufacturer's manual.

3.2.3. PCR Amplification

To amplify the miRNA of interest, a PCR with a miRNA-specific forward primer and a reverse primer that is complementary to linker-2 is used. The forward primer is complementary to the 5′ arm of the miRNA plus few nucleotides from the stem loop so that a melting temperature of 68°C can be achieved (see Note 9).

For the PCR, the following amounts of reagents are used: 100 µM dNTP, 4% DMSO, 0.2 µM of each primer, 250 ng cDNA, and 1 U Pfu Turbo Polymerase. The following amplification conditions are recommended: 95°C 2:00/40×(95°C 0:20/61°C 0:30/72°C 0:30)/72°C 10:00/4°C.

The specificity of the PCR product is determined by running 10 µl of the sample on a 4% agarose gel (see Note 11).

3.2.4. Cloning the PCR Product into a Vector (pcDNA3.1D)

Sequencing a number of different clones is achieved by cloning the pre-miRNAs into a vector instead of sequencing the PCR product. The further experimental strategy will use the pcDNA3.1D

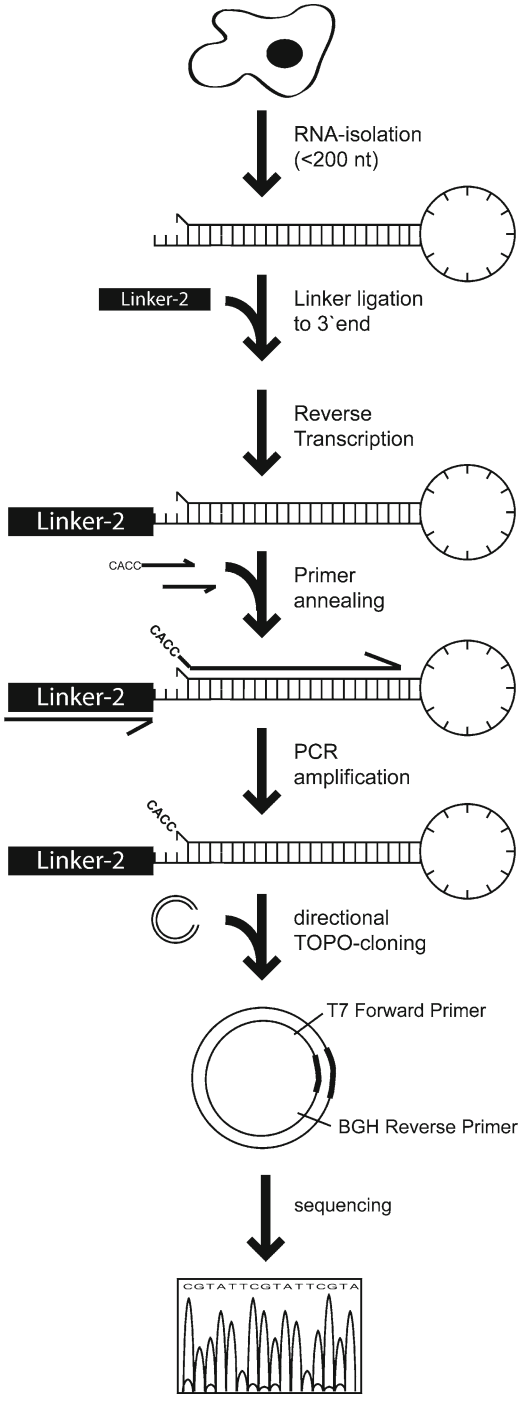


Fig. 2. *Precursor cloning determines the 3′ end of a miRNA.* To determine the 3′end of a miRNA of interest, RNA smaller than 200 nt is isolated prior to linker ligation. cDNA amplification is performed using a pre-miRNA specific forward primer in combination with a reverse primer that is complementary to the linker. Subsequently, the pre-miRNA is cloned into a vector to sequence various clones.

vector from Invitrogen. TOPO cloning enables cloning of blunt-ended PCR products directly into an expression vector in a ligation reaction without subcloning steps. As the vector contains a single-stranded GTGG overhang on the 5′ terminus and a 3′ blunt end, the four-nucleotide overhang invades the dsDNA of the PCR product and anneals to the CACC sequence at the 5′ end of the miRNA-specific forward primer, so that a Topoisomerase finally ligates the PCR product into the vector. The TOPO reaction is performed according to manufacturer's instruction and the vector is transformed into competent *E. coli* cells.

3.2.5. Test Digestion and Sequencing

After the vector is transformed into bacteria, DNA is precipitated from various colonies. Testing for correct integration is performed using restriction enzymes that cut the pcDNA3.1D vector close to the TOPO cloning site, for example, *Xho*I and *Hin*dIII. Using these two restriction enzymes, the predicted size of interest can be calculated as the sum of your pre-miRNA of interest plus 77 nucleotides from the vector backbone. Clones with the correctly predicted size can now be analyzed by sequencing using the T7 Forward primer together with the BGH Reverse Primer.

3.3. Improving RNA Interference with Argonaute-2 (Fig. 3)

3.3.1. Codelivery of Ago2 with RNAi Reagents (shRNA, siRNA, and dsRNA)

Ectopic Ago2 is introduced into the cell in parallel with the RNAi specificity component, the small RNA. The delivery method depends on the experimental setting of the analysis. If the shRNA plasmid or the siRNA duplex are transfected, an expression plasmid for Ago2 can be cotransfected (see Note 12). If a lentiviral shRNA construct is used, a lentiviral vector for Ago2 expression can be used for coinfection protocols. The relative amount of RNAi reagent vs. Ago2 does not appear to be critical: cotransfection of equivalent amounts of shRNA and Ago2 plasmids as well as coinfection of high-titer shRNA with lower titer Ago2 lentiviruses gave convincing results. The rate of RNAi enhancement depends on the shRNA/siRNA used, its baseline activity in the absence of additional Ago2, and the cell line. However, in most experiments, we have seen a strong increase in RNAi efficiency, whereas we have never seen a decrease in RNAi efficiency (partly unpublished data).

Importantly, two controls need to be set up in parallel: first, the experimental comparison should always happen between cells with ectopic Ago2 and the target-specific shRNA vs. cells with ectopic Ago2 and a control shRNA. As a second control, it is worth analyzing the difference between cells expressing ectopic Ago2 compared to cells expressing another, unrelated protein to validate whether or to what extent the overexpression of Ago2 affects the phenotype under investigation.

Notably, RNase-deficient Ago proteins – Ago1, Ago3, Ago4, or RNase-dead Ago2–mutants fail to enhance miRNA-mediated inhibition, restricting the RNAi enhancement to Ago2 in human cells (25). This phenomenon that can be used in every RNAi experiment may be of particular interest in high-throughput screening approaches using siRNA or shRNA libraries. Since these screens are mostly limited by evident false-positive or false-negative results (29), Ago2 cotransfection optimizes the RNAi experiment in two ways, by increasing the efficacy and specificity of target silencing. In contrast to Exportin-5, the effect of Ago2 is restricted to perfectly matched binding sites and does not affect endogenous RNAi activity, making it less prone to off-target effects (25, 26). Increased activity of RNAi allows a reduction in the concentration of the targeting construct and therefore minimizes nonspecific toxicity due to oversaturation of miRNA pathways by siRNA loads (30) and consequently could in the future potentially also enhance the potency of RNAi as a therapeutic utility.

3.3.2. Parameters Affecting Ago2 Activity and RNAi Efficiency

The RNAi-enhancing effect of Ago2 is probably the additive effect of three different mechanisms by which Ago2 could strengthen the efficacy of a given shRNA or siRNA. Firstly, Ago proteins have been shown to increase the expression of mature short RNAs most likely by stabilizing them. However, this effect has been observed for all Ago proteins, but only Ago2 enhances RNAi efficiency indicating that this cannot be the only mechanism. Secondly, Ago2 actively participates in the miRNA biogenesis pathway by cleaving the passenger strand and generating the ac-pre-miRNA (10). This is especially relevant in the shRNA setting, since these hairpins have a high degree of complementarity making them prone to Ago2-mediated cleavage. This also explains why shRNAs benefit more from Ago2 codelivery than siRNAs. Lastly, for a shRNA, siRNA, or miRNA, it appears to be important to which Ago protein it is bound since a short RNA perfectly complementary to a target can only induce efficient target RNA cleavage when bound to the slicer Ago2 (25).

While Ago2 has not affected most cellular pathways analyzed so far, when transiently overexpressed and analyzed within up to 96 h after Ago2 introduction, generation of a cell line stably overexpressing Ago2 does not appear beneficial, since we saw a selection for intermediate or low-expressing cells over time, indicating that Ago2 might slow proliferation in long-term cultures (unpublished results).

In summary, codelivery of Ago2 appears to be most beneficial under the following conditions: shRNA transfection or infection, transient overexpression of Ago2 in parallel with RNAi treatment, and an experimental readout up to 96 h after introduction of ectopic Ago2 into the cell.

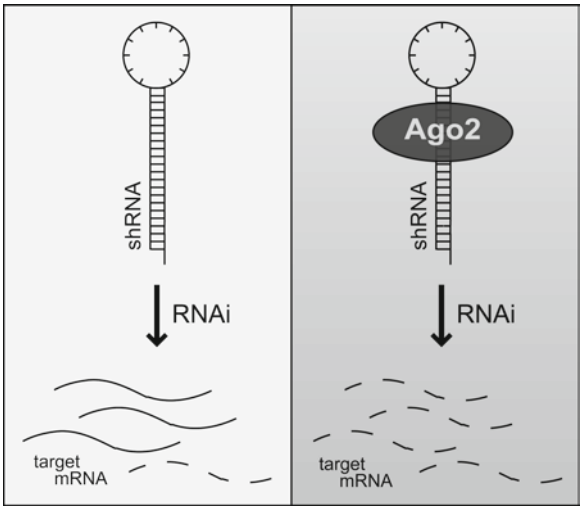


Fig. 3. *Ago2 cotransfection improves RNA interference*. Codelivery of Ago proteins that possess slicing activity (e.g., Ago2 in humans) improves RNAi efficacy and specificity of perfect complementary mRNA targets. Even though shRNA, siRNA, or dsRNA can be used as the specificity component of the RNAi experiment, shRNA transfection in combination with transient overexpression of Ago2 appears to achieve the best results.

4. Notes

1. We recommend testing the integrity of RNA before and after running the sample on urea polyacrylamide gels, since degraded RNA will be distributed throughout a broad range of molecular weight and consequently result in decreased intensity of the signal. Additionally, checking RNA levels prior to blotting also gives first information about equivalent loading of the blot.
2. To prevent degradation, we recommend the treatment of solutions with DEPC that inactivates RNases by covalently modifying histidine residues. Please keep in mind that DEPC cannot be used in combination with Tris and SDS since these reagents inactivate DEPC by reacting with it.
3. Always store enzymes at –20°C and remove exclusively for usage to maintain the 3D structure of the protein and consequently prevent degradation.
4. Half-life of [γ-^{32}P]ATP averages 14 days so that radioactivity decreases constantly. Therefore, we recommend using freshly prepared [γ-^{32}P]ATP if miRNAs with very low expression levels should be detected. Additionally, high amounts of RNA should be loaded onto the gel (20–25 μg) if possible.

5. We recommend to perform the hybridization steps in a 50-ml tube that can be put into a hybridization bottle. Instead of cleaning the hybridization bottles, 50-ml tubes including radioactive material can easily be discarded.
6. As a loading control, the membrane is stained with U6 snRNA or 5S rRNA. Since the expression of these genes is very high, only 10-pmol oligo and 5 μCi [γ-^{32}P]ATP are used to label the probe.
7. The above-mentioned conditions were optimized for our purposes. However, modification of the washing steps or changes in the hybridization temperature can be used to improve your personal miRNA Northern blot protocol. Increased hybridization temperature will decrease both the background signal and the binding capacity, whereas decreased hybridization temperature leads to increased signal intensity coinciding with higher background levels.
8. To clone precursor miRNAs, only a fraction of small RNAs (smaller than 200 nt), not the total RNA of a cell, is required. Since mature miRNAs are 20–23 nt and precursor miRNAs approximately 70 nt in length, larger RNA fragments would only disturb the reaction. Due to size exclusion, miRNAs and their precursors are highly enriched.
9. The miRNA-specific forward primer to amplify the miRNA of interest should never exceed the region where the potential 3′ end of the miRNA might be located. Otherwise, sequencing will give wrong or no information about the 3′ end.
10. For the linker-ligation reaction, a buffer that is depleted of ATP has to be used to prevent circularization of the miRNAs in the presence of ATP.
11. Due to the high Agarose concentrations, it is recommended to use a capacious container as overboiling might occur very rapidly.
12. The human Ago2 cDNA can be obtained from Addgene (http://www.addgene.org), where it has been kindly deposited by the Tuschl lab in form of a pIRESneo expression plasmid. This vector harbors Ago2 under the control of a CMV-driven Polymerase II (Pol II) promoter and is fused to an internal ribosome entry site (IRES) driving the expression of a resistance marker against Neomycin. Any other expression plasmid leading to strong expression of human Ago2 is also suitable.
13. The use of long dsRNA is only possible in invertebrates such as *Drosophila melanogaster* since dsRNA would evoke a Protein Kinase R (PKR) and Interferon (IFN) response in mammalian cells.

References

1. Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R., and Golub, T. R. (2005) MicroRNA expression profiles classify human cancers, *Nature 435*, 834–838.
2. Calin, G. A., Ferracin, M., Cimmino, A., Di Leva, G., Shimizu, M., Wojcik, S. E., Iorio, M. V., Visone, R., Sever, N. I., Fabbri, M., Iuliano, R., Palumbo, T., Pichiorri, F., Roldo, C., Garzon, R., Sevignani, C., Rassenti, L., Alder, H., Volinia, S., Liu, C. G., Kipps, T. J., Negrini, M., and Croce, C. M. (2005) A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia, *N Engl J Med 353*, 1793–1801.
3. Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., and Croce, C. M. (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers, *Proc Natl Acad Sci U S A 101*, 2999–3004.
4. Visone, R. and Croce, C. M. (2009) MiRNAs and cancer, *Am J Pathol 174*, 1131–1138.
5. Winter, J., Jung, S., Keller, S., Gregory, R. I., and Diederichs, S. (2009) Many roads to maturity: microRNA biogenesis pathways and their regulation, *Nat Cell Biol 11*, 228–234.
6. Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs, *Science 294*, 853–858.
7. Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*, *Science 294*, 858–862.
8. Lee, R. C. and Ambros, V. (2001) An extensive class of small RNAs in *Caenorhabditis elegans*, *Science 294*, 862–864.
9. Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14, *Cell 75*, 843–854.
10. Diederichs, S. and Haber, D. A. (2007) Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression, *Cell 131*, 1097–1108.
11. Hutvagner, G. and Simard, M. J. (2008) Argonaute proteins: key players in RNA silencing, *Nat Rev Mol Cell Biol 9*, 22–32.
12. Joshua-Tor, L. (2006) The Argonautes, *Cold Spring Harb Symp Quant Biol 71*, 67–72.
13. Parker, J. S., Roe, S. M., and Barford, D. (2006) Molecular mechanism of target RNA transcript recognition by Argonaute-guide complexes, *Cold Spring Harb Symp Quant Biol 71*, 45–50.
14. Parker, J. S., Roe, S. M., and Barford, D. (2005) Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex, *Nature 434*, 663–666.
15. Yuan, Y. R., Pei, Y., Ma, J. B., Kuryavyi, V., Zhadina, M., Meister, G., Chen, H. Y., Dauter, Z., Tuschl, T., and Patel, D. J. (2005) Crystal structure of *A. aeolicus* argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage, *Mol Cell 19*, 405–419.
16. Rivas, F. V., Tolia, N. H., Song, J. J., Aragon, J. P., Liu, J., Hannon, G. J., and Joshua-Tor, L. (2005) Purified Argonaute2 and an siRNA form recombinant human RISC, *Nat Struct Mol Biol 12*, 340–349.
17. Song, J. J., Smith, S. K., Hannon, G. J., and Joshua-Tor, L. (2004) Crystal structure of Argonaute and its implications for RISC slicer activity, *Science 305*, 1434–1437.
18. Peters, L. and Meister, G. (2007) Argonaute proteins: mediators of RNA silencing, *Mol Cell 26*, 611–623.
19. Eulalio, A., Huntzinger, E., and Izaurralde, E. (2008) Getting to the root of miRNA-mediated gene silencing, *Cell 132*, 9–14.
20. Filipowicz, W., Bhattacharyya, S. N., and Sonenberg, N. (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet 9*, 102–114.
21. Leuschner, P. J., Ameres, S. L., Kueng, S., and Martinez, J. (2006) Cleavage of the siRNA passenger strand during RISC assembly in human cells, *EMBO Rep 7*, 314–320.
22. Matranga, C., Tomari, Y., Shin, C., Bartel, D. P., and Zamore, P. D. (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes, *Cell 123*, 607–620.
23. Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H., and Siomi, M. C. (2005) Slicer function of Drosophila Argonautes and its involvement in RISC formation, *Genes Dev 19*, 2837–2848.
24. Rand, T. A., Petersen, S., Du, F., and Wang, X. (2005) Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation, *Cell 123*, 621–629.

25. Diederichs, S., Jung, S., Rothenberg, S. M., Smolen, G. A., Mlody, B. G., and Haber, D. A. (2008) Coexpression of Argonaute-2 enhances RNA interference toward perfect match binding sites, *Proc Natl Acad Sci USA 105*, 9284–9289.
26. Yi, R., Doehle, B. P., Qin, Y., Macara, I. G., and Cullen, B. R. (2005) Overexpression of exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs, *RNA 11*, 220–226.
27. Chakrabarti, R. and Schutt, C. E. (2001) The enhancement of PCR amplification by low molecular-weight sulfones, *Gene 274*, 293–298.
28. Chomczynski, P., Mackey, K., Drews, R., and Wilfinger, W. (1997) DNAzol: a reagent for the rapid isolation of genomic DNA, *Biotechniques 22*, 550–553.
29. Echeverri, C. J. and Perrimon, N. (2006) High-throughput RNAi screening in cultured cells: a user's guide, *Nat Rev Genet 7*, 373–384.
30. Grimm, D., and Kay, M. A. (2006) Therapeutic short hairpin RNA expression in the liver: viral targets and vectors, *Gene Ther 13*, 563–575.

Chapter 8

miRNA Profiling on High-Throughput OpenArray™ System

Elena V. Grigorenko, Elen Ortenberg, James Hurley, Andrew Bond, and Kevin Munnelly

Abstract

Micro RNA (miRNAs) are a class of 17–25 nucleotides noncoding RNAs that have been shown to have critical functions in a wide variety of biological processes. Measuring quantity of miRNAs in tissues of different physiological and pathological conditions is an important first step to investigate the functions of miRNAs. To this date, the number of identified miRNA consists of around 850 different species, and more sequence-predicted miRNA genes are awaiting experimental confirmation. The need for high-throughput technologies allowing to profile all known miRNAs with power similar to microarray and precision/specificity of qPCR is evident. The example of such system based on high-density array of nanoliter PCR assays is described here. Functionally equivalent to a microtiter plate, a single OpenArray™ nanoplate makes possible to do up to 3,072 real-time PCRs at a single experiment. Methods for miRNA profiling using the dual-label probe chemistry (Taqman®) are outlined in this chapter, and experimental data illustrating system performance are provided.

Key words: miRNA, Taqman®, Polymerase chain reaction, Nanofluidic, High-throughput, Expression profiling, Quantitative PCR

1. Introduction

MicroRNAs (miRNAs) are small (~17–25 nucleotides) noncoding RNAs that bind to mRNAs to regulate protein expression, either by blocking translation and/or by promoting degradation of the mRNA target [1]. They have been found to be involved in numerous cellular processes such as cell fate determination, cell proliferation and cell death [2, 3]. Specific profiles of miRNA in various types of tumors and presence of miRNAs in biological fluids appear to be promising as potential biomarkers for tumorogenesis and patient survival [4, 5, 8]. The number of known miRNAs has rapidly increased in recent years, yet current methods for global

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_8, © Springer Science+Business Media, LLC 2011

profiling of miRNA expression are still lacking and required throughput could be expensive for profiling of large number of samples against full miRNA transcriptome.

The real-time PCR detection of miRNA expression offers several distinct advantages over methods based on hybridization including quantification of miRNA over the large dynamic range and assay specificity with single-base discrimination detecting only mature miRNA and sensitivity for the detection of miRNA in a total RNA preparation. The Life Technologies, Inc (Carlsbad, CA) offers a library of hundreds of predesigned Taqman® miRNA assays suitable for global miRNA expression profiling. The principle of Taqman® miRNA assays is similar to conventional Taqman® RT-PCR. A major difference is the use of novel stem-loop primer during reverse transcription reaction. Stem-loop primers were first reported in 2005 [6] and they have several advantages: (1) the stem-loop facilitates annealing a short reverse sequence to the 3′ end of the miRNA that in turn provides a better discrimination between similar miRNAs, (2) its double-stranded structure does not allow the hybridization of RT primer to miRNA precursors and other long RNAs, (3) base stacking of the stem enhances stability of miRNA–DNA complexes thus improving RT efficiency for relatively short RT primer, and lastly, (4) unfolded stem-loop structure presents a longer RT product that is more amendable for real-time Taqman® assay and internal probe design.

Current strategies to implement qPCR for global miRNA expression analysis using high-density 1,536-well microtiter plate format must overcome several technical challenges. Changes in PCR stoichiometry from evaporation during thermal cycling and nonspecific surface interactions in reduced PCR volume can lead to poor PCR efficiency and precision. Achieving high densities of physically independent PCR volumes requires stringent fluidic isolation between adjacent reaction containers to prevent cross contamination during temperature cycling. Of equal importance are precise methods for transfer of a small volume of PCR mix between individual wells of 1,536-well microplate in a humidity-controlled environment to prevent liquid evaporation during the preparation of plate for PCR. In addition, maintaining dynamic range and precision of qPCR when scaling down to nanovolumes requires a proportional increase in template concentration to ensure an equivalent amount of template in the reduced reaction volume.

We have addressed these shortcomings with an approach based on through-hole nanofluidic arrays [7]. A stainless steel (317 stainless steel) platen, the size of a microscope slide (25 mm × 75 mm × 0.3 mm), is photolithographically patterned and etched to form a rectilinear array of 3,072 through-holes. The through-holes are grouped in 48 subarrays of 64 holes each and spaced on a 4.5-mm pitch equal to that of wells in a 384-well microplate (Fig. 1). A series of vapor and liquid deposition steps covalently attaches a

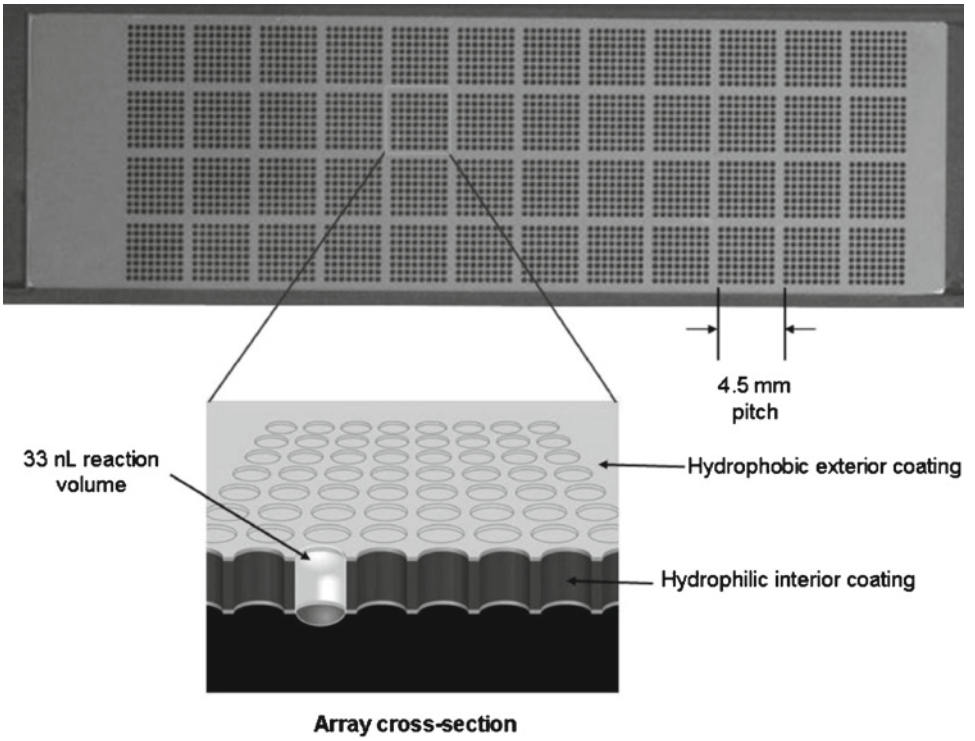


Fig. 1. A stainless steel platen (317 stainless steel) of the size of a microscope slide (25 mm × 75 mm × 0.3 mm) is photolithographically patterned and wet-etched to form a rectilinear array of 3,072 micro-machined, 320 μm diameter holes of 33 nL each. The 48 groups of 64 holes are spaced at 4.5 mm to match the pitch of the wells in a 384-well microplate. A PCR compatible poly(ethylene glycol) hydrophilic layer is amine-coupled to the interior surface of each hole, and a hydrophobic fluoroalkyl layer is vinyl-coupled to the exterior surface of the platen, resulting in the retention in individual, isolated containers of PCR reagents and sample introduced onto the array.

PCR compatible poly(ethylene glycol) (PEG) hydrophilic layer amine-coupled to the interior surface of each through-hole, and a hydrophobic fluoroalkyl layer to the exterior surface of the platen. The differential hydrophilic–hydrophobic coating facilitates precise loading and isolated retention of fluid in each through-hole. Assay primer pairs are transferred into individual through-holes by an array of 48 slotted pins manipulated by a four-axis robot in an environmentally controlled chamber to prevent evaporative loss during loading. Once a platen is fully populated with primer pairs, the solvent is evaporated in a controlled manner leaving the primers immobilized in a PEG matrix on the inside surface of each through-hole. The array loaded with primer is stored in a sealed bag at –20°C, and it is ready for sample addition.

Real-time PCR occurs in a computer-controlled imaging thermal cycler whose essential components are two pairs of off-axis, high-energy light emitting diode (LED) excitation sources, a thermoelectric flat block holding up to three encased arrays, two emission filters in a computer-controlled filter wheel, and a

thermoelectrically cooled CCD camera. Under software control, the real-time method for 9,216 PCR amplifications using dual-labeled probes such as Taqman® is implemented in less than 2.5 h. Postacquisition data processing generates fluorescence amplification and melt curves for each through-hole in the array, from which cycle threshold (C_T) is computed. All data are stored in a flat file (*.csv) format for ready export to a database or third-party software, for further analysis.

2. Materials

2.1. Reagents

1. Human heart, brain, and lung RNA are isolated from normal tissue (Life Technologies, Carlsbad, CA).
2. MegaPlex stem-loop reverse transcription (RT) primers (Life Technologies, Carlsbad, CA).
3. Synthetic miRNA templates for let-7 family isoforms were supplied by IDT DNA Technologies, Inc (Coralville, IA).
4. Taqman® miRNA RT kit (Life Technologies, Carlsbad, CA).
5. Universal Taqman® Mastermix (Life Technologies, Carlsbad, CA).
6. REmix (5×), an additive provided by BioTrove and required for performance of dual-labeled assays on BioTrove OpenArray™ platform.
7. miRNA-specific Taqman® assay that includes one tube of miRNA-specific RT primer and one tube containing the mix of miRNA-specific forward and reverse primer and miRNA-specific Taqman® MGB probe (Life Technologies, Carlsbad, CA).
8. OpenArray™ DLP qPCR Kit (BioTrove Inc, Woburn, MA).
   - OpenArray™ plates (nanotiter plates) preloaded with user supplied Taqman® miRNA assays and probes.
   - OpenArray™ qPCR cases.
   - Immersion fluid.
   - Case sealing glue.
   - Plate-file CDROM specifying location of the miRNA Taqman® assays in the nanotiter plate.
9. NT Imager (BioTrove Inc.)

2.2. Additional Materials Supplied by Biotrove

1. Vertical slide holder.
2. Autoloader (Fig. 2).
3. Case sealing station.
4. Autoloader supplies.

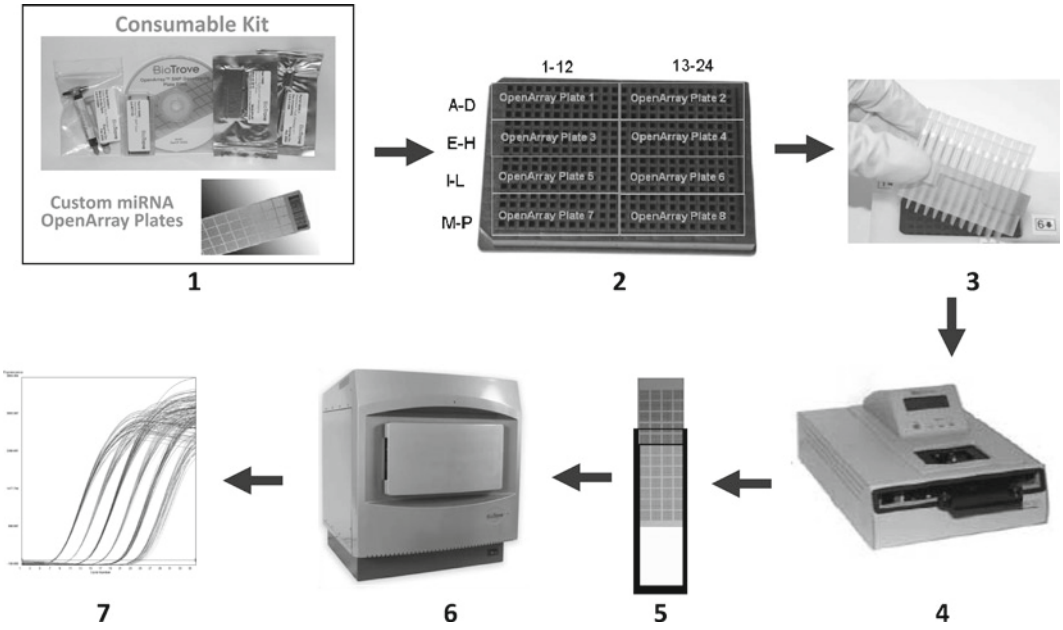


Fig. 2. The OpenArray workflow procedure. Customers receive OpenArray™ plate kit consisting of preloaded miRNA custom plates and OpenArray™ consumables (1), prepare sample plate (2), load samples on OpenArray™ plate using Autoloader (3, 4), insert OpenArray™ plate into a case which is sealed with glue (5), proceed with plate cycling (6), and analyze results (7).

5. OpenArray™ plate holder.
6. OpenArray™ tip block.
7. Box of tips for AutoLoader (384 plasma-treated 20 µL FinnPipette tips).

2.3. Additional Materials Not Supplied by Biotrove

1. FinnPipette pipettor (16-channel), 5–50 µL (VWR).
2. MatriCal plates, 384-well, black, low volume, polypropylene (MatriCal).

3. Methods

3.1. Testing miRNA Assay Specificity Using Synthetic Oligonucleotide Template

The existence of multiple miRNA isoforms within same family presents a significant challenge in miRNA quantification. To determine whether primers used in the study can distinguish between multiple miRNA isoforms, primer specificity should be tested on a synthetic template. The target sequence for a specific assay can be found in online academic or industrial databases (for example http://www.mirbase.org or http://www.appliedbiosystems.com). The let-7 family was chosen as a model for this experiment since its members have mismatches of one or more nucleotides (Table 1) in a sequence.

Table 1
Isoforms of human let-7 family

Let-7a	UGAGGUAGUAGGUUGUAUAGUU
Let-7b	UGAGGUAGUAGGUUGUGUGGUU
Let-7c	UGAGGUAGUAGGUUGUAUGGUU

Note: single nucleotide difference as highlighted letters between three let-7 isoforms.

3.1.1. Reverse Transcription Reaction Using Synthetic Oligonucleotide Template

1. Allow to thaw Taqman® miRNA Reverse Transcription Kit components and reverse primers on ice.
2. Prepare RT mix according to the manufacturer protocol with the final volume of 15 μL adjusted to the number of samples to be reverse transcribed plus 20% of overage for pipetting errors. For a single reaction, combine RT master mix consisting of 0.15 μL of 100 mM dNTPs, 1 μL of Multiscribe reverse transcriptase, 1.5 μL of 10× RT buffer, 0.19 μL of RNase inhibitor, 4.16 μL of water, and 3 μL of reverse miRNA Taqman® primer and 5 μL of synthetic miRNA template at concentration of 100 pg per reaction.
3. Mix gently, spin to bring solution at the bottom of the tube and incubate in thermocycler using the following program: 16°C – 30 min, 42°C – 30 min, 85°C – 5 min, 4°C – indefinitely.
4. After RT program is completed, take the tube out of the thermocycler, spin it, and put the tube on ice.
5. The sample is now ready for the addition of PCR reagents and loading into OpenArray™ plate or it can be stored at –20°C.

3.1.2. Preparing Sample Plate for OpenArrays Loading

1. Thaw reverse-transcribed miRNA samples on ice, spin tubes at 1,000 × *g* for 1 min. Place sample on ice to keep cool.
2. Prepare PCR master mix by combining 132 μL of GeneAmp PCR Master mix, 52 μL of Remix, and 15.8 μL of PCR-grade water. Please note that these reagents' volumes are required for one OpenArray plate. Adjust volumes accordingly if you will be using more than one plate (see Note 3).
3. Pipette 3.8 μL of PCR master mix into 48 adjacent wells of 384-well MatriPlate.
4. Pipette 1.2 μL of reverse-transcribed synthetic miRNA template, mix with PCR mix by gently pipetting twice.
5. Microplate layout: Sample and master mix are dispensed into wells A1–A12, B1–B12, C1–C12, and D1–D12 where each well on microtiter plate corresponds to a subarray in the OpenArray plate (Fig. 2, point 2).

6. Cover the plate with aluminum foil sealing tape, and centrifuge the plate for 1 min at 500 × *g* to eliminate bubbles. Plate the plate on ice to keep samples cold.

3.1.3. OpenArray Loading

1. With forceps, peel off the foil exposing the 48 samples to be loaded onto the OpenArray plate.
2. Place the plate guide exposing the well from which samples will be transferred into OpenArray plate. To load the first OpenArray plate, the mask exposes only wells A1–A12, B1–B12, C1–C12, and D1–D12.
3. A tip block is prepared by placing 20 μL specially treated Finn Pipette tips into each hole of the tip block with 16-channel pipettor. Inserting 12 pipette tips at a time fills all 48 pipette tips positions in four insertion operations (Fig. 2, point 3).
4. The tip block is inserted into the plate guide, so that each individual pipette tip draws up fluid from each of 48 contiguous wells in the MatriPlate. One minute of gentle vertical movement of the tip block relative to the microplate ensures that at least 4 μL of liquid (sample and PCR master mix) is drawn into each pipette tip.
5. Insert OpenArray Plate into the plate holder and place it into Autoloader (Fig. 2, point 4).
6. Insert the tip block into Autoloader. The Autoloader brings the pipette tip block into a contact with OpenArray plate and while plate moves relative to the tip, the liquid from pipette tips dispenses into the through-holes of OpenArray plate.
7. Fill the case three-fourths with immersion fluid. This is to prevent thermal evaporation of PCR mix during thermal cycling (Fig. 2, point 5).
8. Remove the plate from Autoloader, insert into the cassette, and pipette UV-curable epoxy into the case to seal. Insert cassette into the case sealer where the glue is exposed to UV light to cure the epoxy forming a hermetic seal.

3.1.4. Thermal Cycler Protocol and Imaging

1. Clean the glass surface of sealed-OpenArray plates with 70% ethanol by squirting small amount of liquid on glass surface and wipe excess with lint-free KimWipes.
2. Attach the plate frames with the adhesive to the case prior to thermal cycling. Make sure that frames are attached to the most black side of the case. The frames keep case in tight contact with thermal block to ensure thermal uniformity during cycling.
3. The OpenArray thermal cycling protocol consists of 94°C – 10 min; 59°C – 70 s, 94°C – 60 s for 40 cycles. The images

are collected at the end of each cycle PCR and this information is stored into an appropriate folder.

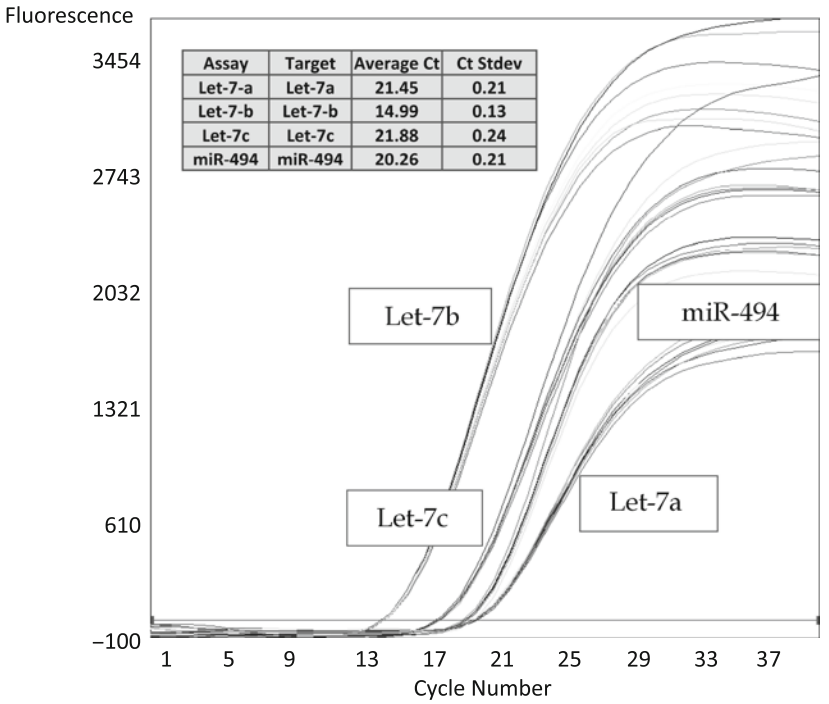
4. Figure 3 shows fluorescence amplification curves for three members of let-7 family and unrelated miR-494.

3.2. Identification and Quantification of miRNAs in Total RNA Samples

The quantification of miRNA in total RNA samples using Taqman® miRNA assays is usually done by two steps: reverse transcription reaction followed by PCR (Note 1). In RT step, the cDNA is reversely transcribed from a total RNA using a mixture of stem-loop MegaPlex RT primers. Instead of doing RT reaction with a single miRNA-specific primer, the mix of MegaPlex primers allows to perform RT reaction for large number of miRNAs. In addition, the usage of these primers (1) improves the workflow, (2) reduces reagent cost, (3) minimizes sample input, (4) allows normalization to control genes detected in the same reaction as miRNA and lastly, allows to perform global miRNA expression analysis (see Note 2).

3.2.1. Reverse Transcription Reaction Using Total RNA

1. Allow to thaw Taqman® miRNA Reverse Transcription kit components and MegaPlex RT primers on ice.
2. Prepare RT mix according to the manufacturer protocol with the final volume of 15 μL adjusted to the number of samples



Assay	Target	Average Ct	Ct Stdev
Let-7-a	Let-7a	21.45	0.21
Let-7-b	Let-7-b	14.99	0.13
Let-7c	Let-7c	21.88	0.24
miR-494	miR-494	20.26	0.21

Fig. 3. The example of real-time fluorescence amplification curves using miRNA synthetic templates for amplification. Real-time fluorescence amplification curves showing a specific detection of expression of closely related members of let-7a family. Six independent reactions were performed for each miRNA assay, and the *inlet table* shows the precision of assay performance of technical replicates on OpenArray system.

to be reverse transcribed plus 20% of overage for pipetting errors. For a single reaction, combine RT master mix consisting of 0.15 µL of 100 mM dNTPs, 1 µL of Multiscribe reverse transcriptase, 1.5-µL of 10× RT buffer, 0.19-µL of RNase inhibitor, 6.16 µL of water, and 1 µL of MegaPlex stem-loop RT primer mix and 5 µL of total RNA template at concentration of 100 ng per reaction (see Notes 1 and 3).

3. Mix gently, spin to bring solution at the bottom of the tube, and incubate in thermocycler using the following program: 16°C – 30 min, 42°C – 30 min, 85°C – 5 min, 4°C – indefinitely.
4. After RT program is completed, take tube out of thermocycler, spin it, and put tube on ice.
5. The sample is now ready for the addition of PCR reagents and loading into OpenArray™ plate or it can be storedat –20°C.

3.2.2. Preparing Sample Microplate and OpenArray Loading

Follow the same procedure described in Subheading 3.1.3 for the preparation of sample plate and Subheading 3.1.4 for OpenArray loading procedure, also see Fig. 2.

3.2.3. Thermal Cycler Protocol and Imaging

1. The same cycling protocol is used for the profiling of reverse-transcribed total RNA using MegaPlex step-loop primer mix (see Subheading 3.1.4).
2. Figure 4 shows the tissue-specific expression of let-7b and let-7c in heart, lung, and brain RNA similar to results reported in the literature [9, 10](see Note 4).

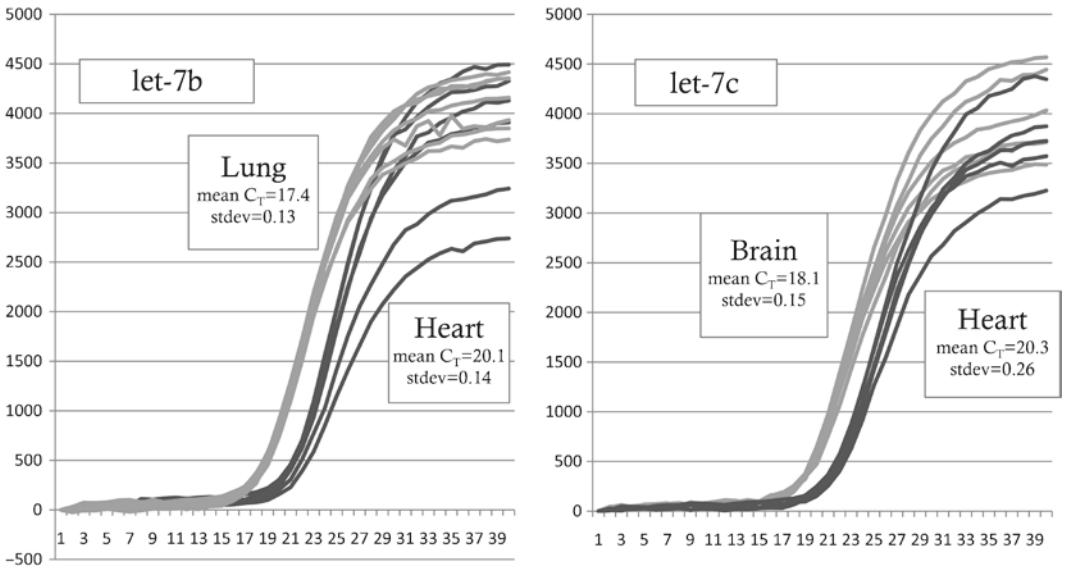


Fig. 4. Tissue-specific expression of let-7b and let-7c miRNA in normal heart, lung, and brain tissues. Five independent reactions were performed for each miRNA assay, and the *inlet table* shows the precision of assay performance of technical replicates on OpenArray system.

4. Notes

1. Greater than 1 ng of cDNA per Taqman assay in the OpenArray plate is recommended, consistent with the Life Technologies' recommendation of 1–10 ng of cDNA for the same assay in a microplate. To achieve this goal, a starting RNA concentration of 200 ng/μL is required.
2. If detection of low abundant RNA is required or starting RNA material is limited, the preamplification of cDNA with miRNA-specific primers can be done before PCR. For these purposes, it is recommended to use Taqman® PreAmp Master (Life Technologies, Inc) mix in a combination with MegaPlex Preamp primers. The number of preamplification PCR cycles can be determined empirically and it is based on starting cDNA concentration. The dilution of preamplification products is necessary for downstream qPCR, and testing of dilution series is recommended for any new sample type.
3. Volume of PCR master mix is sufficient to load one OpenArray plate plus an additional 10% volume to account for pipetting errors.
4. RNA quality and amount per reaction can affect miRNA Taqman® assay performance. When isolating RNA from FFPE tissues, use appropriate RNA isolation kits.

References

1. Nilsen TW (2007) Mechanisms of microRNA-mediated gene regulation in animal cells. Trends Genet, 23:243–249.
2. Ambros V (2004) The functions of animal microRNAs. Nature, 431:350–355.
3. Kloosterman WP, Plasterk RH (2006) The diverse functions of microRNAs in animal development and disease. Dev Cell, 11:441–450.
4. Hunter MP, Ismail N, Zhang X, Aguda BD, Lee EJ, Yu L, Xiao T, Schafer J, Lee ML, Schmittgen TD, Nana-Sinkam SP, Jarjoura D, Marsh CB (2008) Detection of microRNA expression in human peripheral blood microvesicles. PLoS, 3:e3694.
5. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, Jacks T, Horvitz TH, Golub TR (2005) MicroRNA profiles classify human cancers. Nature, 435:834–838.
6. Vass L, Kelemen JZ, Feher LZ, Lorincz Z, Kulin S, Xseh S, Dorman G, Puskas LG (2009) Toxicogenomics screening of small molecules using high-density, nanocapillary real-time PCR. Int J Mol Med, 23:65–74.
7. Dixon JM, Lubomirski M, Amaratunga D, Morrison TB, Brenan CJ, Ilyin SE (2009) Nanoliter high-thorughput RT-qPCR: a statistical analysis and assessment. Biotechniques, 46:ii–viii.
8. Wang K, Zhang S, Marzolf B, Troisch P, Brightman A, Hu Z, Hood LE, Galas DJ (2009) Circulating miRNAs, potential biomarkers for drug-induced liver injury. Proc Natl Acad Sci USA, 11:4402–4407.
9. Williams AE, Larner-Svensson H, Perry MM, Campbell GA, Herrick SE, Adcock IM, Erjefalt JS, Chung KF, Lindsay MA (2009) MicroRNA expression profiling in mild asthmatic human airways and effect of corticosteroid therapy. PLoS, 4:e5889.
10. Sun Y, Koo S, White N, Peralta P, Esau C, Dean N, Perera RJ (2004) Development of a micro-array to detect human and mouse microRNAs and characterization of expression in human organs. Nucleic Acids Res, 32:e188.

Chapter 9

Silicon Nanowire Biosensor for Ultrasensitive and Label-Free Direct Detection of miRNAs

Guo-Jun Zhang

Abstract

MicroRNA (miRNA), a large and growing class of 18–24-nucleotide long, noncoding RNA molecules in all known animal and plant genomes, is a key player in gene regulation. The functions of miRNA are yet to be understood with respect to how and where it is produced and the changes within an organism associated with variations in miRNA expression level. The expression profiles serve as molecular diagnostics for diseases and new targets in drug discovery. Consequently, highly sensitive and selective detection of miRNA is playing a significant role in understanding miRNA functions. Existing major methods of detecting miRNA are dependent on hybridization, in which a target miRNA molecule is hybridized to a complementary probe molecule. Recently developed detection methods introduce nanomaterials to the hybridized duplex to enhance the sensitivity. However, all of them are indirect, involving labeling or conjugating process. To overcome the above-mentioned issues, we have demonstrated a highly sensitive and label-free direct detection method for miRNA by using peptide nucleic acids (PNAs)-functionalized silicon nanowires (SiNWs) biosensor. The sensor is capable of detecting target miRNA as low as 1 fM (10^{-15} M), as well as identifying fully matched versus mismatched miRNA sequences. More importantly, the SiNW biosensor enables miRNA detection in total RNA extracted from HeLa cells. The developed detection method shows potential applications in label-free, early detection of miRNA as a biomarker in cancer diagnostics with very high sensitivity and good specificity.

Key words: Silicon nanowire, Biosensor, Label-free detection, Hybridization, Peptide nucleic acid, MicroRNA

1. Introduction

MicroRNAs (miRNAs) are noncoding RNA molecules ranging in size from 18 to 24 nucleotides, which regulate gene expression in both plants and animals. They play an important role in developmental and cell biology, including stem cell differentiation and development. In addition, miRNAs are clinically important biomarkers for many cancer types because miRNA expression can be

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_9, © Springer Science+Business Media, LLC 2011

correlated with cancer type and stage. It is therefore critical to develop robust detection methods for miRNAs with high sensitivity, selectivity, as well as simplicity (1).

Owing to its small size, detection of miRNA is challenging. Many methods for detection of miRNA have been developed including northern blotting (2, 3), microarray (4, 5), polymerase chain reaction (PCR) (6), surface plasmon resonance (SPR) (7), as well as a variety of nanoparticles-based (8), conducting polymer nanowires-based (9), and bioluminescence-based techniques (10). Northern blot allows gene expression quantification and miRNA size determination. However, the low sensitivity and time-consuming laborious procedures of northern blot make it difficult for routine miRNA analysis (2, 3). Nucleic acid assays combine amplification by polymerase chain reaction (PCR) and detection using fluorophores as labels (4, 5). Microarray-based detection methods also require labeling to visualize the hybridization event (6). To improve detection sensitivity of miRNA, a SPR-based method is capable of detecting miRNA in total RNA samples down to femtomolar concentrations (7). Moreover, some conjugates like nanoparticles (8), enzymes (10) are introduced to the hybridized duplex, which enhances the signals. The sensitivity is thus improved by using these strategies. Very recently, a nanogapped microelectrode-based biosensor array has been fabricated for ultrasensitive electrical detection of miRNAs (9).

To date, all reported methodologies require labeling and conjugating steps and are thus time-consuming and indirect. Therefore, there is a demand for miRNA detection methods that are sensitive, direct, simple, rapid, and label-free for measurement of miRNA directly from a cellular extract. Direct and label-free electrical readout systems provide an extremely attractive sensing modality, which is widely applicable for miRNA detection. Silicon nanowires (SiNWs) are ultrasensitive biosensors that are capable of detecting nucleic acids (11–17) and proteins (18–20). In this chapter, we describe a new method of directly detecting miRNA excluding labeling and conjugating processes using peptide nucleic acids (PNAs)-functionalized silicon nanowires (SiNWs) biosensor with high sensitivity and good specificity. As PNA does not have an anionic phosphate backbone, the hybridization between PNA and miRNA eliminates repulsion, resulting in increased melting temperature and subsequently enhanced hybridization efficiency. The SiNW biosensors functionalized with PNA are capable of discriminating between single base differences, as in single nucleotide polymorphisms (SNPs). Furthermore, miRNA in total RNA extracted from cancer cell lines are detectable by the developed SiNW biosensors. Due to high surface-to-volume ratio that the nanowire dimensions confer, the sensitivity is greatly enhanced and the detection limit can be lowered to femtomolar concentrations.

2. Materials

2.1. Reagents

1. Silicon-on-insulator (SOI) wafers with 145 nm buried oxide layer are used.
2. Piranha solution (see Note 1).
3. Buffered hydrofluoric (HF) acid (see Note 2).
4. 3-aminopropyltriethoxysilane (99%, Sigma-Aldrich).
5. Glutaraldehyde (25% in H_2O, Sigma-Aldrich).
6. PNA probes are synthesized by Eurogentec (Herstal, Belgium).
7. MiRNAs are synthesized by 1st BASE Oligos (Singapore).
8. 20×SSC buffer (1st BASE Oligos, Singapore).
9. Trifluoroacetic acid (TFA, Sigma-Aldrich).
10. Diethyl pyrocarbonate (DEPC, 99%, Aldrich).
11. TRIzol reagent (Invitrogen, Carlsbad, CA).
12. YM-50 Montage spin column (Millipore Corp., Billerica, MA).
13. RNaseZap (Ambion, TX).

2.2. Equipment

1. Oxide growth (Furnace, SEMCO).
2. DUV lithography (Clean Track 8 and NIKKON Scanner, TEL).
3. Fin etch (RIE, Precision 5000 Mark II, Applied Materials).
4. Ion implant (E500 HP Implanter, VARIAN).
5. Implant activation (RTP Systems AST 3000, STEAG).
6. Metal deposition (Endura HP PVD, Applied Materials).
7. Metal etch (Centura Metal Etcher, Applied Materials).
8. Oxide deposition (Centura PECVD System, Applied Materials).
9. Dry oxide etch (Oxide Etcher, TEL).
10. Probe station (Alessi REL-6100, Cascade Microtech).

2.3. DEPC-Treatment of Solutions

1. Add 0.1% (v/v) (DEPC) to water/buffer solutions involved in the experiment (see Note 3).
2. Store them overnight.
3. DEPC must then be completely destroyed by autoclaving (see Note 4).

2.4. Preparation of 2% 3-Aminopropyltriethoxysilane Solution

1. Prepare 10 ml of ethanol/H_2O/3-aminopropyltriethoxysilane (98.5:0.5:1 (v/v)) by mixing 9.85 ml of ethanol, 50 µl of ultrapure H_2O, and 100 µl of 3-aminopropyltriethoxysilane (see Note 5).
2. Do not use the mixture if it turns turbid.

3. The freshly prepared solution is used immediately.

2.5. Preparation of Glutaraldehyde Solution

1. Prepare 10 ml of 2.5% glutaraldehyde by dissolving 1 ml of 25% glutaraldehyde water solution in 9 ml of D.I. water.
2. Use the solution immediately.

2.6. PNA Preparation

1. The PNA sequence is N-AACCACACAACCTACTAC-CTCA-C.
2. Dissolve in a calculated amount of D.I. water including 0.1% TFA based on the PNA amount provided in the certificate as stock (see Note 6).
3. Buffer preparation: dilute stock 20×SSC buffer to 1×SSC buffer, store at room temperature.
4. Dilute PNA to 10 μM in the 1×SSC buffer.
5. Store in aliquots at –20°C.

2.7. Preparation of miRNA Oligos

1. Let-7b is a complementary target, whose sequence is 5′-UGAGGUAGUAGGUUGUGUGGUU-3′.
2. Let-7c is a one-base mismatched target, whose sequence is 5′-UGAGGUAGUAGGUUGUAUGGUU-3′.
3. Control is a non-complementary target, whose sequence is 5′-AUGCAUGCAUGCAUGCAUGCAA-3′.
4. Dissolve in a calculated amount of D.I. water based on the miRNA amount provided in the certificate as stock.
5. Buffer preparation: dilute stock 20×SSC buffer to 0.01×SSC buffer, store at room temperature.
6. Dilute miRNA to various concentrations in the 0.01×SSC buffer.
7. Store in aliquots at –20°C.

2.8. Extraction of Total RNA from Hela Cells

1. Total RNA from human HeLa cells is extracted using TRIzol reagent according to the manufacturer's recommended protocol.
2. MiRNAs in the total RNA are enriched using an YM-50 Montage spin column.
3. RNA concentration is determined by UV-vis spectrophotometry.
4. Store in aliquots at –80°C.

3. Methods

SiNW sensors are typical FET-based devices, which contains source, drain, and gate electrodes. Figure 1 shows a schematic

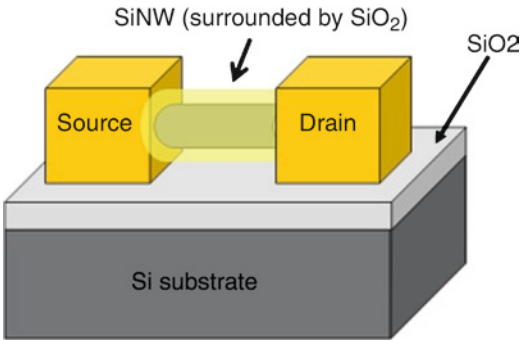


Fig. 1. Schematic of a single SiNW biosensor consisting of source (S) and drain (D) contacts and an individual SiNW surrounded by SiO_2 (reproduced from ref. (17) with permission from Elsevier Science).

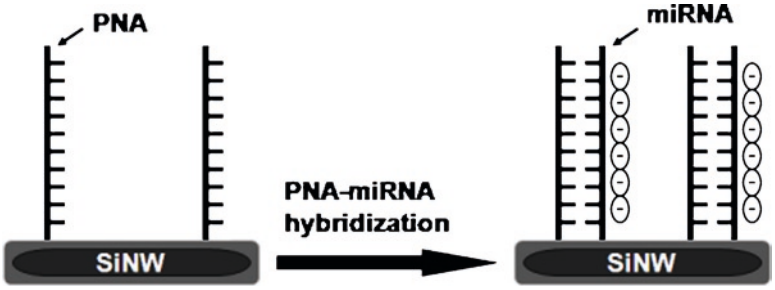


Fig. 2. Schematic illustration of the label-free direct hybridization assay developed for ultrasensitive detection of miRNA (reproduced from ref. (17) with permission from Elsevier Science).

diagram of a typical SiNW biosensor. The sensing mechanism by SiNW can be understood in terms of the change in charge density which induces a change in electric field at the SiNW surface. For example, binding of biomolecules with negative charge to the surface of *n*-type FET leads to an increase in device resistance.

Ultrasensitive, direct, and label-free detection for microRNA plays an important role. Figure 2 illustrates the working principle of the SiNW biosensor for detection of miRNA. PNA is covalently immobilized on the electrically addressable SiNW surface via conventional silane chemistry. Electrical biosensing by SiNW is based on change in resistance of the SiNWs due to depletion of charge carriers in its "bulk" when negatively charged miRNA originating from the phosphate groups on the miRNA backbone are bound to the PNA-functionalized surface via hybridization. When the targeted species, miRNAs complementary to the immobilized PNA, are present, resistance change occurs, whereas when they are noncomplementary, resistance change is minimal. As each wire is provided with independent metal contacts, resistance

can be individually measured. The miRNAs for hybridization were let-7b (complementary), let-7c (one-base mismatched), and control (noncomplementary).

3.1. Fabrication of Silicon Nanowire Device

1. The silicon wafers are doped with *n*-type (phosphorous) impurities using an ion implanter where implant dose is varied from 1×10^{13} to 1×10^{15} cm^{-2} and energy from 30 to 50 keV.
2. The dopants are then activated in rapid thermal annealing furnace and nanowire-fins are patterned using standard DUV lithography in the array format.
3. Silicon is etched in reactive ion etcher and resulting fins (60–80 nm wide) are oxidized in O_2 at 900°C for 2–6 h to realize nanowire array.
4. The two ends of the nanowires are further doped to obtain *n*+ regions, followed by connecting to contact metal and alloying to realize Ohmic contacts.
5. The device is passivated by silicon nitride film except for the active nanowire sensor area and metal pads.
6. Etching in nanowire area is carried out using a combination of dry and wet release.
7. Figure 3 shows an optical image of a SiNW sensor chip having two portions of SiNW arrays and 100 SiNWs in each portion. The individual SiNW has a dimension of ~50 nm in diameter and a length of 100 μm.

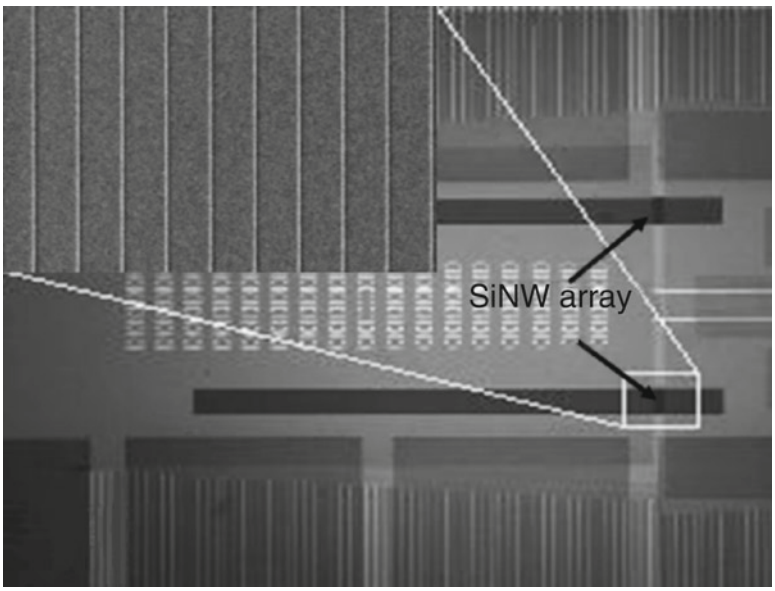


Fig. 3. Optical image of a SiNW sensor chip. Each chip consists of two portions of individually addressable SiNW arrays. Inset image shows high-resolution SEM image of ten nanowires in each portion. The individual SiNW has a dimension of ~50 nm in diameter and a length of 100 μm (reproduced from ref. (17) with permission from Elsevier Science).

3.2. Surface Functionalization of Silicon Nanowire

1. Immerse the chips in solution of 2% 3-aminopropyltriethoxysilane for 2 h.
2. Wash the chips with absolute ethanol for three times and blow-dry.
3. Immerse the chips in 2.5% glutaraldehyde in water for 1 h.
4. Wash the chips with water and blow-dry.
5. The chips are used immediately.

3.3. Immobilization of PNA on the Nanowire Surface

1. Apply 20 µl of 10 µM PNA in 1× SSC on the nanowire surface.
2. Place the chips in a humid atmosphere at room temperature overnight (see Note 7).
3. Wash the chips with the same buffer, rinse with water before measurement, and blow-dry.

3.4. Sequence Specificity of the PNA-Functionalized SiNW Biosensor

1. Apply the furnished acrylic wells to the NW region after carefully peeling off the cover of double-side tape (see Notes 8 and 9).
2. The chip surfaces are decontaminated by RNaseZap, followed by a wash with 70% ethanol (see Note 10).
3. 20 µl of 0.01×SSC buffer solution is added to the well as buffer medium.
4. The SiNW resistances are measured with the probe station (see Note 11).
5. Resistances are measured at 0.1 V.
6. Rinse the chip with water after measurement.
7. Apply 20 µl of 1 nM miRNA in 0.01×SSC on the surface.
8. Place the chips in a humid atmosphere at room temperature for 1 h.
9. Wash the chips with the same buffer, rinse with water, and blow-dry.
10. 20 µl of 0.01×SSC buffer solution is added to the well as buffer medium.
11. Resistances are measured at 0.1 V again.
12. The resistance change of SiNWs before and after PNA-miRNA hybridization is recorded.
13. Data are analyzed as an average of the responses from 15 measurements.
14. The high sequence specificity is demonstrated by hybridizing the PNA-functionalized SiNW sensor to the three different types of targets (let-7b, let-7c, and control, respectively). An example of the results of the sequence specificity produced is shown in Fig. 4.

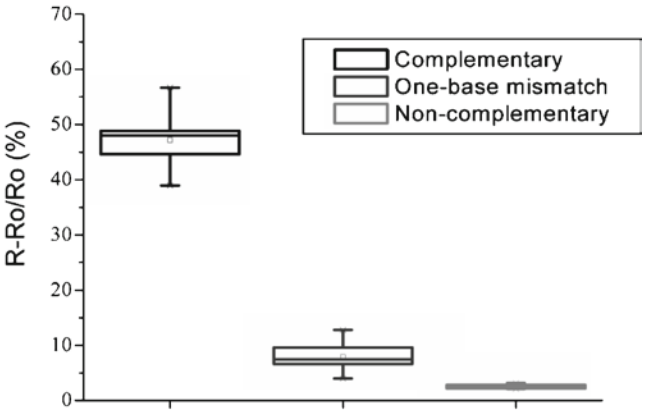


Fig. 4. Hybridization specificity demonstrated by response of the PNA-functionalized SiNW biosensors to fully complementary (let-7b), one-base mismatched (let-7c), and noncomplementary miRNA sequences. The specific resistance change was obtained from hybridization of let-7b to the SiNW device immobilized with PNA. As can be seen, a significant change (~47.2%) was observed when 1 nM let-7b was used, whereas only a negligible change was obtained when same concentration of noncomplementary miRNA was applied to the SiNW device. To evaluate the capability of the SiNW device for discrimination of single base mismatch in miRNAs, let-7c was tested at 1 nM. The increase in resistance for let-7c was ~7.9% which is much lower than that of let-7b. The high specificity suggests that the SiNW device allows for label-free discrimination between the fully matched and mismatched miRNAs, offering unique advantage over other technologies which require labeling and additional tags (reproduced from ref. (17) with permission from Elsevier Science).

3.5. Detection Response of the PNA-Functionalized SiNW Biosensor

1. Apply the acrylic wells to the sensor surface carefully after PNA immobilization.
2. The chip surfaces are decontaminated by RNaseZap, followed by a wash with 70% ethanol.
3. 20 µl of 0.01×SSC buffer solution is added to the well as buffer medium.
4. Resistances are measured at 0.1 V.
5. Rinse the chip with water after measurement.
6. Different concentrations of miRNA are applied to the PNA-functionalized SiNW sensor.
7. Various resistances are measured at 0.1 V.
8. The detection response of the PNA-functionalized SiNW biosensor is proportional to the concentrations of the complementary target miRNA. Examples of the response as a function of concentrations of miRNA are shown in Fig. 5.

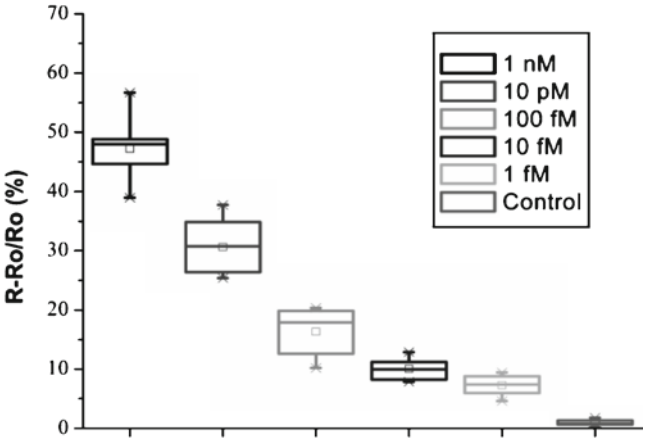


Fig. 5. Response of the PNA-functionalized SiNW biosensors to the complementary miRNA (let-7b) of varying concentrations. Different known concentrations of miRNA were tested with different groups of SiNWs. The resistance change before and after PNA-miRNA hybridization primarily depends on the amount of charge layer contributed by miRNA. The more the target miRNA molecules hybridized, the more negative charges accumulated on the SiNW surface, and thus the higher the resistance increase. As described above, an obvious resistance increase was obtained when 1 nM let-7b was hybridized to the PNA-functionalized SiNW device. It was observed that resistance change drops as a function of varying concentrations of let-7b. A 7.3% response was still observed while 1 fM let-7b was employed, which is distinguishable from the control signals. This indicates that ultralow concentrations of miRNA can effectively be detectable down to 1 fM with the device used in the work without labeling/tagging (reproduced from ref. (17) with permission from Elsevier Science).

3.6. Detection of miRNA in Hela Cells

1. Apply the acrylic wells to the sensor surface carefully after PNA immobilization.
2. The chip surfaces are decontaminated by RNaseZap, followed by a wash with 70% ethanol.
3. 20 µl of 0.01×SSC buffer solution is added to the well as buffer medium.
4. Resistances are measured at 0.1 V.
5. Rinse the chip with water after measurement.
6. Aliquots of the total RNA are diluted with 0.01×SSC and subsequently applied to a PNA-functionalized SiNW device for 1 h.
7. Resistances are measured at 0.1 V.
8. The results obtained with resistance change are normalized with respect to the total RNA.
9. The concentration of let-7b detectable in the total RNA extracted from Hela cells is determined.

4. Notes

1. Piranha solution is very dangerous, being both strongly acidic and a strong oxidizer. It must be handled extremely carefully.
2. HF is an extremely hazardous liquid and vapor and is highly corrosive to eyes and skin. Goggles and gloves must be used during operation.
3. Unless stated otherwise, all buffer solutions should be prepared in water.
4. All buffer solutions and water are treated with diethylpyrocarbonate by mixing 0.1% of diethylpyrocarbonate with the solutions and storing them overnight prior to autoclaving.
5. A glass petri dish must be used for this step. EtOH will dissolve plastic petri dishes and cause chips to adhere to the dish.
6. All the tips used in the experiment are sterilized prior to use.
7. Put a wet paper on the bottom of a Petri dish, place the SiNW chip on the top of the paper, and seal the dish with Parafilm.
8. Rectangular wells (inner size: ~0.8×0.4 cm^2) are cut by laser machine after double-side tape is fixed on the bottom of the acrylic plate. Ensure that wells are properly adhered to the surface of the NW chip to prevent any leakage of the solution.
9. Check for leakage with DI water after pasting the wells on the surface.
10. Treatment is conducted in a Class II Biohazard Safety Cabinet.
11. SiNW terminals: (1) Drain connected to voltage source; (2) Source connected to common ground (any one of the common pads); (3) Substrate connected to common ground.

References

1. Cissell, K.A., Shrestha, S., and Deo, S.K. (2007) MicroRNA detection: challenges for the analytical chemist. Anal. Chem. **79**, 4754–4761.
2. Valoczi, A., Hornyik, C., Varga, N., Burgyan, J., Kauppinen, and S., Havelda, Z. (2004). Sensitive and specific detection of microRNAs by northern blot analysis using LNA-modified oligonucleotide probes. Nucleic Acids Res. **32**, e175.
3. Ramkissoon, S.H., Mainwaring, L.A., Sloand, E.M., Young, N.S., and Kajigaya, S. (2006) Nonisotopic detection of microRNA using digoxigenin labeled RNA probes. Mol. Cell. Probes **20**, 1–4.
4. Liu, C.G., Calin, G.A., Meloon, B., Gamliel, N., Sevignani, C., Ferracin, M., Dumitru, D.C., Shimizu, M., Zupo, S., Dono, M., Alder, H., Bullrich, F., Negrini, M., and Croce, C.M. (2004) An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. Proc. Natl. Acad. Sci. U. S. A. **101**, 9740–9744.
5. Thomson, J.M., Parker, J., Perou, C.M., and Hammond, S.M. (2004) A custom microarray platform for analysis of microRNA gene expression. Nat. Methods **1**, 47–53.
6. Raymond, C.K., Roberts, B.S.; Garrett-Engele, P., Lim, L.P., and Johnson, J.M. (2005) Simple, quantitative primer-extension PCR assay for direct monitoring of micro

RNAs and short-interfering RNAs. RNA **11**, 1737–1744.

7. Fang, S., Lee, H.J., Wark, A.W., and Corn, R.M. (2006) Attomole detection of microRNAs by nanoparticle-amplified SPR imaging measurements of surface polyadenylation reactions. J. Am. Chem. Soc. **128**, 14044–14046.
8. Gao, Z.Q. and Yang, Z.C. (2006) Ultrasensitive detection of MicroRNA using electrocatalytic nanoparticle tags. Anal. Chem. **78**, 1470–1477.
9. Fan, Y., Chen, X.T., Trigg, A.D., Tung, C.-H., Kong, J.M., and Gao, Z.Q. (2007) Detection of MicroRNAs using target-guided formation of conducting polymer nanowires in nanogaps. J. Am. Chem. Soc. **129**, 5437–5443.
10. Cissell, K.A., Rahimi, Y., Shrestha, S., Hunt, E.A., and Deo, S.K. (2008) Bioluminescence-based detection of microRNA, miR21 in breast cancer cells. Anal. Chem. **80**, 2319–2325.
11. Hahm, J.-I. and Lieber, C.M. (2004) Direct ultrasensitive electrical detection of DNA and RNA sequence variations using nanowire nanosensors. Nano Lett. **4**, 51–54.
12. Li, Z., Chen, Y., Li, X., Kamins, T.I., Nauka, K., and Williams, R.S. (2004) Sequence-specific label-free DNA sensors based on silicon nanowires. Nano Lett. **4**, 245–247.
13. Bunimovich, Y.L., Shin, Y.S., Yeo, W.-S., Amori, M., Kwong, G., and Heath, J.R. (2006) Quantitative real-time measurements of DNA hybridization with alkylated nonoxidized siliocn nanowires in electrolyte solution. J. Am. Chem. Soc. **128**, 16323–16331.
14. Gao, Z.-Q., Agarwal, A., Trigg, A.D., Singh, N., Fang, C., Tung, C.-H., Fan, Y., Buddharaju, K.D., and Kong, J.-M. (2007) Silicon nanowire arrays for label-free detection of DNA. Anal. Chem. **79**, 3291–3297.
15. Zhang, G.-J., Chua, J., Chee, R.E., Agarwal, A., Wong, S.M., Buddharaju, K.D., and Balasubramanian, N. (2008) Highly sensitive measurements of PNA–DNA hybridization using oxide-etched silicon nanowire biosensors. Biosens. Bioelectron. **23**, 1701–1707.
16. Zhang, G.-J., Zhang, G., Chua, J., Chee, R.E., Wong, E.H., Agarwal, A., Buddharaju, K.D., Singh, N., Gao, Z.Q., and Balasubramanian, N. (2008) DNA sensing by silicon nanowire: charge layer distance dependence. Nano Lett. **8**, 1066–1070.
17. Zhang, G.-J., Chua, J., Chee, R.E., Agarwal, A., and Wong, S.M. (2009) Label-free direct detection of MiRNAs with silicon nanowire biosensors. Biosens. Bioelectron. **24**, 2504–2508.
18. Zheng, G., Patolsky, F., Cui, Y., Wang, W.U., and Lieber, C.M. (2005) Multiplexed electrical detection of cancer markers with nanowire sensor arrays. Nat. Biotechnol. **23**, 1294–1301.
19. Stern, E., Klemic, J.F., Routenberg, D.A., Wyremebak, P. N., Turner-Evans, D.B., Hamilton, A.D., LaVan, D.A., Fahmy, T.M., and Reed, M.A. (2007) Label-free immunodetection with CMOS-compatible semiconducting nanowires. Nature **445**, 519–522.
20. Chua, J., Chee, R.E., Agarwal, A., Wong, S.M., and Zhang, G.-J. (2009) Label-free electrical detection of cardiac biomarker with CMOS-compatible silicon nanowire sensor arrays. Anal. Chem. **81**, 6266–6271.

Chapter 10

High-Throughput and Reliable Protocols for Animal MicroRNA Library Cloning

Caide Xiao

Abstract

MicroRNAs are short single-stranded RNA molecules (18–25 nucleotides). Because of their ability to silence gene expressions, they can be used to diagnose and treat tumors. Experimental construction of microRNA libraries was the most important step to identify microRNAs from animal tissues. Although there are many commercial kits with special protocols to construct microRNA libraries, this chapter provides the most reliable, high-throughput, and affordable protocols for microRNA library construction. The high-throughput capability of our protocols came from a double concentration (3 and 15%, thickness 1.5 mm) polyacrylamide gel electrophoresis (PAGE), which could directly extract microRNA-size RNAs from up to 400 μg total RNA (enough for two microRNA libraries). The reliability of our protocols was assured by a third PAGE, which selected PCR products of microRNA-size RNAs ligated with 5′ and 3′ linkers by a miRCat™ kit. Also, a MathCAD program was provided to automatically search short RNAs inserted between 5′ and 3′ linkers from thousands of sequencing text files.

Key words: Total RNA abundance, MicroRNA, Polyacrylamide gel electrophoresis, 5′ Linker, 3′ Linker, T–A cloning, Plasmid

1. Introduction

A group of nonprotein-coding, endogenous, and short RNA molecules (18–25 nucleotides) were first discovered in *Caenorhabditis elegans* (1). During the last decade, it was found that this kind of short RNA molecules had been identified in virtually every metazoan and plant species examined. This kind of short RNA molecules is called microRNA (miRNA). MicroRNAs are important regulatory molecules, which control gene expression at the posttranscriptional levels. Mature miRNAs can bind by partially complementary sequences to target messenger RNAs (mRNAs)

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_10, © Springer Science+Business Media, LLC 2011

and negatively regulate their expression by preventing translation into proteins. Interestingly, a miRNA can bind up to 200 genes, and each mRNA could have recognition sites for more than one miRNA (2). About 30% of the human protein-coding genes are negatively regulated by miRNA (3), which suggests that miRNAs are the biggest regulators in gene expression (4).

Identifying miRNA targets is a very important step to study miRNA functions. Although computational approaches could be used as a powerful strategy to predict miRNAs (4–6) based on sequence complementarities between miRNAs and their targets, experimental construction of miRNA libraries was the most important step to identify miRNAs from organisms with unknown genome sequences (7, 8) and tumor tissues. In this chapter, high-throughput and reliable protocols are presented for miRNA libraries construction in six steps shown in Fig. 1. In step 1, TRIzol® reagent (Invitrogen (9)) was used to extract total RNA from animal tissues, and the amount of total RNA needed for a miRNA library construction was discussed. In step 2, a double concentration denaturing polyacrylamide gel electrophoresis (PAGE) was applied to separate short pieces of RNA (~21 nt) from total RNA, and the maximum load of a channel in the gel electrophoresis was discussed. In steps 3, 4, and 5, protocols from a miRCat® kit (Integrated DNA Technologies, Inc. (10)) were followed to clone small RNAs and eachwas linked to a 5′ linker as a cap and a 3′ linker as a tail. In step 6, gel electrophoresis was used for the third time to select the short pieces of RNA (~21 nt)

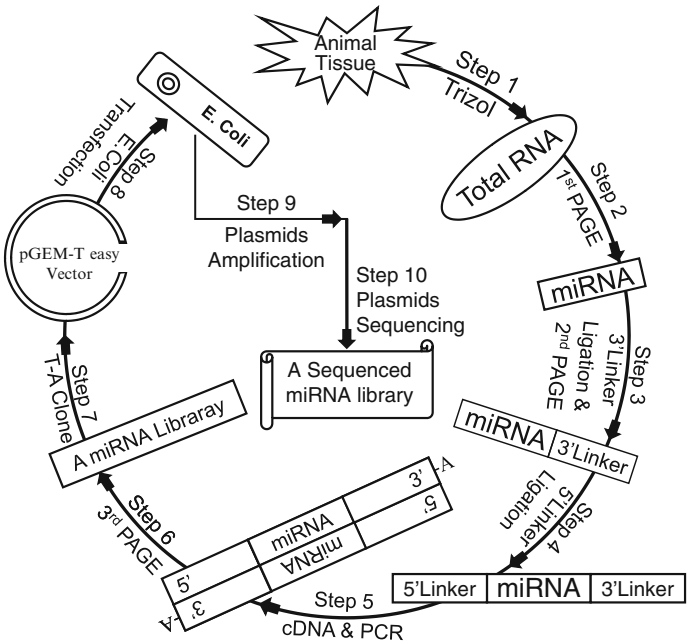


Fig. 1. Ten steps for an animal microRNA library construction. It took about 2 days to finish work from step (1) to step (6).

with 5′ and 3′ linkers from the PCR products. Through the three PAGE selections, all short pieces of RNAs selected were microRNA-size RNAs. After these samples are sequenced, you will have thousands of text files about your samples. Looking through each text file to find a possible miRNA sequence either by eye or by Microsoft WORD or by BioEdit is a tedious task. At the end of the chapter, a mathCAD® program is provided for automatically searching possible miRNAs from thousands of sequence text files.

2. Materials

2.1. Total RNA Extraction from Tissues

2.1.1. Instruments

1. A tissue homogenizer.
2. A high-speed centrifuge.
3. A vortexer.
4. Three refrigerators (+4, −20, and −80°C).
5. KimWipers® paper.
6. Nanadrop® spectrophotometer.

2.1.2. Chemicals

1. Trizol reagent (Invitrogen).
2. Chloroform.
3. Isopropanol.
4. 75% ethanol (−20°C).
5. DNAse & RNAse-free water.
6. 4 M lithium chloride (for egg RNA).

2.2. Denaturing PAGE

2.2.1. Instruments

Bio-RAD® mini-protean tetra cell system (11) with a 5-well, 1.5 mm comb (module 165-3363) and a 10-well, 0.75 mm comb (module 165-3354).

2.2.2. Chemicals

1. 10× TBE buffer (Tris–Borate–EDTA).
   - 108 g Tris base (890 mM).
   - 55 g boric acid (890 mM).
   - 9.3 g Na_4EDTA (20 mM).

   Add ddH_2O to 1.0 L. The pH is 8.3 and requires no adjustment.
2. 50× TAE buffer (Tris-acetate buffer).
   - 242 g Tris base.
   - 57.1 mL glacial acetic acid (40 mM).
   - 37.2 g $Na_2EDTA{\cdot}2H_2O$ (20 mM).
   - Add ddH_2O to 1 L.

3. 40% gel solution (38% acrylamide with 2% bisacrylamide).
   - 38.0 g acrylamide (ultrapure).
   - 2.0 g *N*,*N*′ methylene-bis-acrylamide (ultrapure).

   Add ddH_2O to 60 mL in a glass bottle and heat the bottle in a water bath (<55°C) to dissolve the chemicals. Adjust to 100 mL by adding ddH_2O. Cover the bottle with aluminum foil and store it at 4°C.
4. 6× Formamide loading buffer.
5. 10% w/v Ammonia persulfate solution (APS).
   - 100 mg ammonia persulfate.
   - Add ddH_2O to 1.0 mL.
6. GelStar nucleic acid stain (Cambrex BioScience, Cat. no. 50535).

2.3. A MicroRNA Cloning Kit: miRCat™

Table 1 shows the reagents and their tube numbers of a miRCat kit. The small letters and capital letters in the oligonucleotide sequences of the tube 1–5 represent ribonucleotides in RNA and deoxyribonucleotides in DNA, respectively. Reagents in the kit are sufficient for the generation of ten small RNA libraries. The internal RNA control (miSPIKE™) in tube 3 is a 21-mer RNA designed specifically to assist in small RNA cloning. Because this RNA oligonucleotide lacks a 5′ phosphate, it cannot be 5′ linked and will not participate in subsequent ligation steps. The oligonucleotides serve mainly as a marker for the locations of small RNAs in electrophoresis gels and a mass carrier/coprecipitant for small RNAs. The miSPIKE in tube 3, two cloning linkers in tube 1, 2, and two PCR primers in tube 4, 5 are dry powder. Before opening these tubes, they should be centrifuged (<800 × *g*) so that the dry reagents go down to the bottom of the tubes. According to Table 2, add liquid from tube 7 to rehydrate these agents.

2.4. Additional Agents

1. PCR reagents.
2. PCR amplicon cloning kit (pGEM T-EASY, Promega – Cat. no. A1380).
3. 50 bp DNA Ladder (BioLab Inc., N3236S).
4. DTR desalting columns (Edge Biosystems, Cat. no. 42453).
5. Super optimal broth (SOB medium).
6. Agarose gel (1.5%) Electrophoresis: weigh 0.75 g agarose, add 1×TAE buffer 50 mL, and heat in microwave oven for 2 min. When the gel solution cools down to around 60°C, add 10 µL×1 µg/mL EtBr, and cast the gel horizontally with a comb. Add samples to wells and run the gel at 100 V for about 1 h in 1×TAE buffer. Check electrophoresis results by UV light.
7. QIAprep Spin Miniprep Kit.

Table 1
Reagents of miRCat™ kit

Tube no.	Contents
1	3′ Cloning linker: 5′-rAppCTGTAGGCACCATCAAT/3ddC-3′
2	5′ Cloning linker: 5′-TGGAATucucgggcaccaaggU-3′
3	Internal control RNA: 5′-cucaggatggcggagcggucu-3′
4	Forward PCR primer: 5′-TGGAATTCTCGGGCACC-3′
5	Reverse transription/PCR primer: 5′-GATTGATGGTGCCTACAG-3′
6	IDT RNase/D Nase/pryrogen-free water
7	IDTE (pH 7.5)
8	10× Ligation buffer
9	Ligation enhancer
10	10 mM ATP
11	10 mg/mL glycogen
12	3 M NaOAc (pH 5.2)
13	T4 RNA ligase (5 U/μL)
14	T4 DNA ligase (30 U/μL)

Table 2
Rehydration of stock oligonucleotides in the miRCat™

Tube no.	Contents	Add liquid from tube 7 (μL)	Final concentration (μM (pMole/μL))
1	3′ Cloning linker	40	50
2	5′ Cloning linker	40	50
3	Internal control RNA	15	10
4	Forward PCR primer	40	10
5	Reverse transcription/PCR primer	40	10

3. Methods

3.1. Using TRIzol® to Extract Total RNA from Animal Tissues

For a miRNA library construction from a specific tissue, the first question is how many grams of the tissue are needed. The answer to this question is determined by two factors. The first one is the minimum quantity of total RNA needed for a miRNA library cloning by a specific method, and the second one is the total RNA

abundance of the tissue. The total RNA abundance of a tissue is defined as the amount of total RNA extracted from 1 g of tissue. According to our experience on purification of miRNA by PAGE, 200–300 μg total RNA is enough for the generation of an miRNA library. Figure 2 shows the abundance of total RNA in nine tissues of a kind of fish: rainbow trout (*Oncorhynchus mykiss*). The liver tissue has the highest abundance of total RNA. From 1 g of liver tissue, 9,309.8 μg of total RNA could be extracted by Trizol reagent; the fish oocyte (eggs) tissue has the lowest abundance of total RNA. Only 52.0 μg of total RNA could be obtained from 1 g of fish eggs. According to the abundance data shown in the figure, the weight of a specific tissue for an miRNA library could be found from the ratio of 300 μg total RNA to the abundance of the tissue. For example, to construct a rainbow trout fish oocyte miRNA library, we have to use at least 5 g of fish eggs. To construct an miRNA library from several tissues, we have to pool different volumes of total RNA from different tissues, so that each tissue contributes the same amount total RNA in the pool. For many kinds of tissues, their total RNA could be extracted by using Trizol reagent with protocols suggested by Invitrogen or other companies. Because fish oocytes have a high content of glycoprotein, fat, and polysaccharides, a special protocol must be used for oocyte total RNA purifications. The protocols presented in this chapter for miRNA library generation had been successfully used to generate a fetal cow (*Bos taurus*) ovary miRNA library, a rainbow trout somatic miRNA library, and a rainbow trout oocyte (eggs) miRNA library.

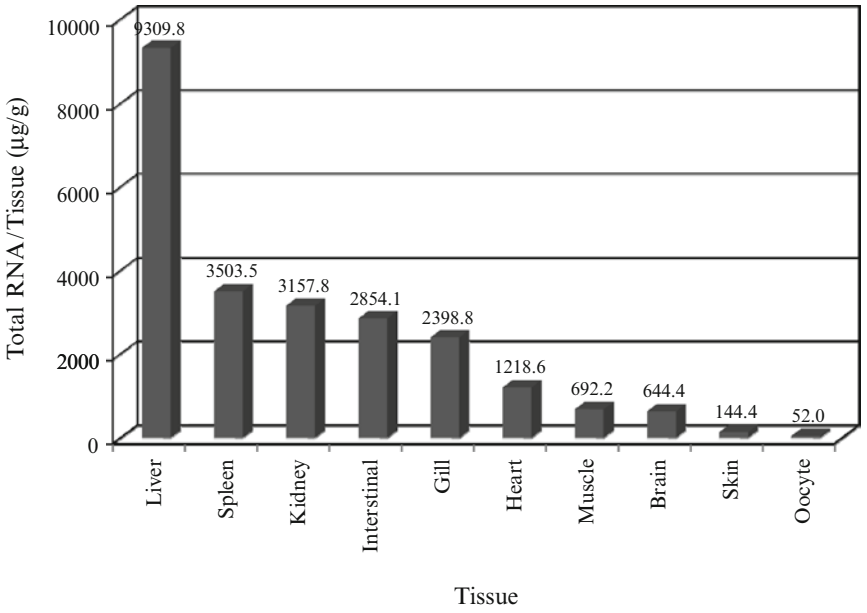


Fig. 2. The abundance of total RNA in nine tissues of rainbow trout fish.

3.1.1. Homogenization

1. Set up a homogenizer and place a small beaker under the blade of the homogenizer. Prepare three 50-mL conical plastic tubes per sample with MilliQ water (30–40 mL). Clean the blade of the homogenizer in the water by running the homogenizer three times in the water of the three tubes.
2. Add Trizol to a 50-mL conical plastic tube and place the tube on ice.
3. Take an animal tissue out of a –80°C refrigerator and quickly put the frozen tissue into the Trizol tube. Usually the ratio of the weight of a tissue to the volume of Trizol should be 1.0 g: 10 mL. But for fish egg total RNA extraction, the ratio could be changed to 2.0 g: 10 mL.
4. Homogenize the tissue as soon as possible in a hood at room temperature.
5. Incubate the homogenized samples for 5 min at room temperature to permit the complete dissociation of nucleoprotein complexes.

3.1.2. Phase Separation

1. Add 0.2 mL of chloroform per 1 mL of Trizol reagent to the tube.
2. Cap sample tubes securely. Shake tubes vigorously by hand for 15 s and incubate them at room temperature for 2–3 min.
3. Centrifuge the samples at no more than 12,000 × *g* for 15 min at 4°C (the tube holding the liquid must be strong enough to tolerate the centrifugation force). Following centrifugation, the mixture separates into a lower, red phenol-chloroform phase, an interphase, and an upper, colorless aqueous phase. RNA remains exclusively in the aqueous phase. The volume of the aqueous phase is about 60% of the volume of Trizol reagent used for homogenization.

3.1.3. RNA Precipitation

1. Transfer the aqueous phase to a fresh plastic tube.
2. Precipitate the RNA from the aqueous phase by mixing with isopropyl alcohol. Use 0.5 mL of isopropyl alcohol per 1 mL of Trizol reagent used for the initial homogenization.
3. Incubate samples at room temperature for 10 min and centrifuge at no more than 12,000 × *g* for 10 min at 4°C. The RNA precipitate, often invisible before centrifugation, forms a gel-like pellet on the sides and bottom of the tube. Because of glycoprotein in fish egg, a very large visible pellet could be seen for fish egg total RNA.

3.1.4. RNA Wash

1. Pour off the isopropanol and dry the tubes briefly by placing them upside down on KimWipers®.
2. To the fish egg total RNA precipitated by isopropyl alcohol, add 5 mL of 4 M lithium chloride and keep on ice for 30 min;

resuspend the pellet in the solution by vortexing and then transfer the liquid to four 1.5-mLtubes and centrifuge at 12,000 × *g* at 4°C for 10 min. Pure RNA pellets are visible but smaller. Pour off lithium chloride.

3. Wash the RNA pellet once with 75% ethanol, adding at least 1 mL of 75% ethanol per 1 mL of Trizol reagent used for the initial homogenization.
4. Mix the sample by vortexing and centrifuge at no more than 7,500 × *g* for 5 min at 4°C.

3.1.5. Redissolving Total RNA

1. At the end of the procedure, briefly dry the RNA pellet (air-dry or vacuum-dry for 5–10 min). Do not dry the RNA by centrifugation under vacuum. It is important not to let the RNA pellet dry completely as this will greatly decrease its solubility. Partially dissolved RNA samples have an A_{260}/A_{280} ratio < 1.6. Figure 3 is a screen snap of a NanoDrop spectrophotometer measuring the total RNA quality from rainbow trout eggs. The figure shows the concentration of the total RNA is 3,216.1 ng/mL and its A_{260}/A_{280} ratio is 2.04.
2. Dissolve RNA in RNase-free water by passing the solution a few times through a pipette tip and incubating for 10 min at 58°C, and then store at –80°C.

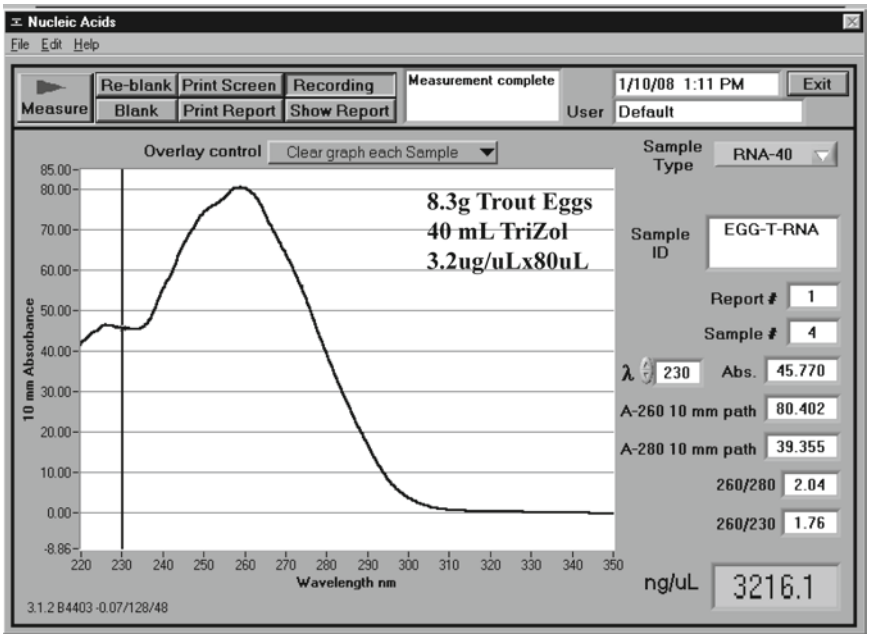


Fig. 3. A screen printing of a Nanadrop® spectrophotometer, which was used to measure total RNA concentrations and to check the total RNA quality.

3.2. From Total RNA to miRNA: First PAGE

We found that a single 15% denaturing (7 M urea) polyacrylamide gel could not be directly used to extract miRNA from large amount of total RNA (>100 μg). Because large RNA molecules would get trapped in the gel and block the pores in the gel, samples would not move in the electric field, even the voltage producing the electrical field was very high. Different-sized RNA molecules moving in a gel lane are very similar to different-sized vehicles moving on a highway. If there are too many heavy trucks running on all lanes of a highway, then small cars would not be able to run ahead of heavy trucks. We found that a part of 3% and a part of 15% denaturing polyacrylamide gel could handle large amount of total RNA in a Bio-RAD® mini-protean tetra cell system (Fig. 4a). In our experiments, a 5-well casting module 165-8019 was used for gel preparation. The thickness of a gel casted in the module was 1.5 mm. The length of the gel in RNA molecules moving direction was 65.0 mm, and the width of the gel was 84.0 mm. The depth of the five wells was 10.0 mm, and the width of the wells was 13.0 mm. The interval of two neighbor wells was 3.0 mm. With this system, our experiments showed that the maximum total RNA load of each well in the system could be up to 80 μg. Total RNA molecules went through the 3% gel of about 22 mm length at first. Since the 3% gel had larger pores, RNA molecules could be separated according to their sizes in the direction of the electric field on the way of the 22 mm gel. RNA molecules with very large sizes were trapped in the 3% gel, and RNA molecules with medium and small sizes could enter the next 15% gel of about 45 mm length. The following are the protocols for how to cast a gel with double concentrations.

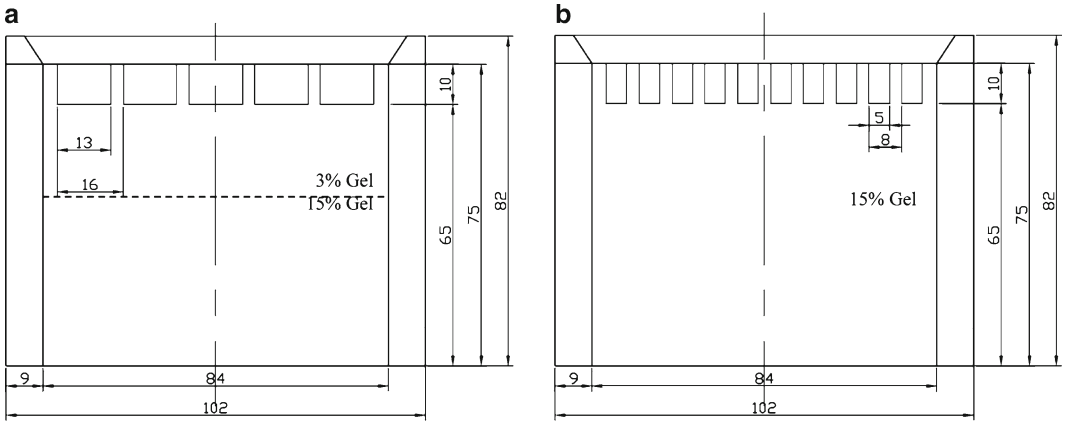


Fig. 4. The sizes of Bio-RAD® mini-protean tetra cell casting modules. The left module is for the first PAGE (1.5 mm), and the right module is for the second and third PAGE (0.75 mm).

3.2.1. Preparation of 15% Gel 10 mL with 7 M Urea

Assemble the gel caster before handling liquid in a 15-mL cone plastic tube!

4.2 g	Urea
1.0 mL	10× TBE
2.0 mL	ddH2O
3.75 mL	40% Gel solution

1. Heat the tube with the above chemicals in hot water (<55°C) for 3 min to dissolve the urea; cool the tube to room temperature. Add ddH_2O to 10 mL.
2. Degas the mixture for 10 min in vacuum.

3.2.2. Preparation of 0% Gel Stock 100 mL with 7 M Urea

42 g	Urea
10 mL	10× TBE
50 mL	ddH_2O

Heat (<55°C) for 3 min to dissolve the urea, cool to room temperature, and add ddH_2O to 100 mL. Degas the mixture for 10 min in vacuum. This stock could be kept at room temperature for 2 months.

3.2.3. Preparation of 3% PAGE Gel

Take 1.0 mL 15% gel to mix with 4 mL 0% gel in a 15-mL cone tube.

1. Add 40 µL APS to a tube with 9.0 mL 15% gel and 7 M urea.

3.2.4. Gel Cast

2. Add 5 µL TEMED and mix well in less than 60 s.
3. Use 6.5–7.0 mL of the 9.0 mL of the 15% gel to cast into the 1.5 mm-thick spacer between the two pieces of glass in a Bio-RAD® mini-protean tetra cell, gently cover the gel surface with water and wait for 1 h.
4. A clear interface between the gel and water could be seen. Pour the water out and use filter paper to suck remaining water on the gel surface.
5. Add 20 µL of APS and 2.5 µL of TEMED to the 5 mL of 3% gel with 7 M urea, cast it on the top of the 15% gel, and insert a 5-teeth comb. Wait for 1 h.

3.2.5. Run the Gel

The fraction of miRNA-size RNA in total RNA is so small that even if a large amount (400 µg) of total RNA is used in the first PAGE, it is still hard to see miRNA bands stained with a sensitive stain: GelStar® nucleic acid. An internal control RNA from miRCat® kit had to be used to mix with the total RNA sample to show the bands of miRNA size RNA. To reduce the amount of 21 nt miSPIKE™ 21-mer internal RNA control, we added only the marker to the sample in one lane of a gel. For higher PAGE

resolution, the sample volume in a well should be as small as possible. Usually, a sample volume in a well is less than 30% of the well volume.

1. Take the comb away and assemble the gel between the two glasses into the Bio-RAD® mini-protean tetra cell system.
2. Fill the gel trough with 1 × TBE (~600 mL) and prerun the gel for 30 min at 100 V.
3. Rinse each well immediately before loading.
4. Mix 20 μL 6 × Loading dye with 100 μL × 3 μg/μL total RNA and load the mixture 24 μL/well to four wells.
5. The rest of 24 μL mixture is mixed with 2 μL × 1.7 μg/μL 21-nt marker and loaded in the center well.
6. Run the gel at 100.0 V constant voltage until the dye moves to the bottom of the lanes (~2 h).
7. Move the gel in a tray with 100 mL 1 × TAE buffer and add 2.5 μL GetStar™ dye to the tray for staining RNA in the gel. Gently shake the tray for about 1 h.
8. Take a picture of the gel shined with UV (Fig. 5).
9. Cut out a gel slice 2 mm above and below the 21 nt RNA marker band in each lane and put each piece of gel in a 2-mL tube.

3.2.6. Short RNA Recovery

1. Add 200 μL water to each 2-mL plastic tube with a piece of gel cut from a lane at the location of a 21-nt band shown in Fig. 5.
2. Use a sterile, RNase-free plastic rod to break the gel slice into small pieces in the tube.
3. Heat the gel suspension in the tube to 70°C for 10 min.
4. Centrifuge a DTR Performa Gel Filtration Cartridge (EDGE Biosystems) for 3 min at 850 × *g*.
5. Vortex the suspension and transfer the gel slurry into the DTR cartridge, centrifuge it for 3 min at 850 × *g*.
6. Discard the DTR column.
7. Add 3 μL 10 mg/mL glycogen, 25 μL of 3 M NaOAc (pH 5.2), and 900 μL ice-cold 100% EtOH to the flow-throw. Mix by inversion and hold at −80°C for 30 min.
8. Centrifuge the tube at full speed (>10,000 × *g*) for 10 min; discard the supernatant, and dry the pellet for the next procedure.

3.3. Short RNA Ligation to 3′ Linker and the Second PAGE

Once the short RNA fraction has been separated by the first PAGE, the short RNAs could be ligated with 3′ linkers. Because of the presence of an activated 5′ adenylation (12), the 3′ linker can be ligated to short RNA chains by T4 RNA ligase in the absence of ATP without circularization of small RNAs. The 3′ end of the linker is blocked with a dideoxycitidine (ddC) to prevent

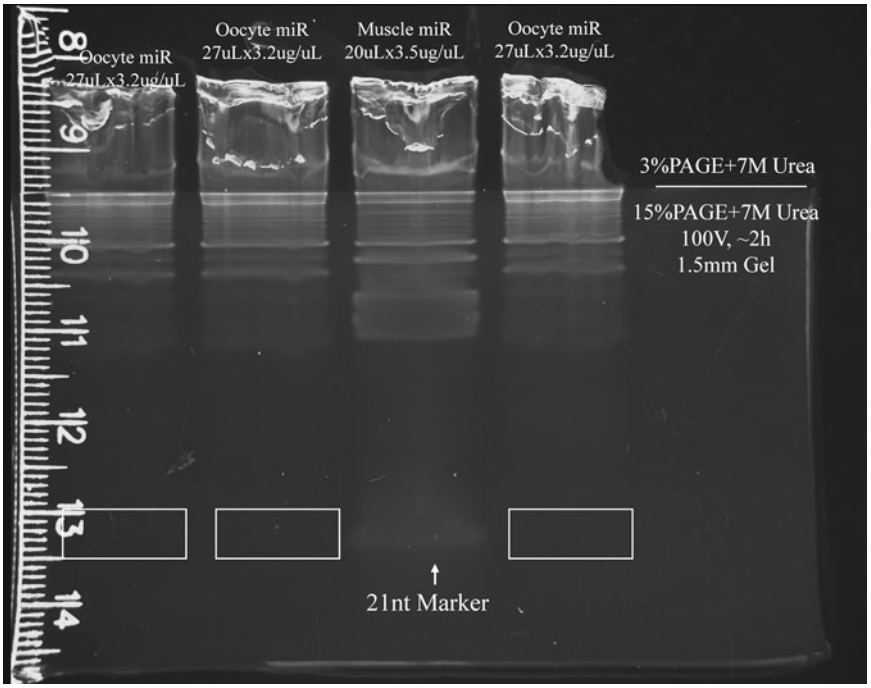


Fig. 5. UV photos of the first PAGE: using (3% + 15%) PAGE with 7 M urea to separate miRNA size RNA from total RNA. The 3% and 15% gel interface was located at 9.5 cm on the rulers shown in the photo. Only the rainbow trout muscle total RNA was mixed with the 21-nt marker whose band was shown between 13.0 and 13.5 cm. The 21-nt band indicated miroRNA locations on the gel as shown by three rectangles.

circularization of the ligated short RNAs. For example, the mature sequence of let-7 found on chromosome X in *C. elegans* (12) is 5′-ugagguaguagguuguauaguu-3′. The miRNA let-7 was also found in fetal bovine ovaries in our experiment. The sequence of the ligation product of a let-7 with a 3′ linker should be 5′-ugagguaguagguuguauaguu-CTGTAGGCACCATCAAT/3ddC-3′.

Recovered small RNA fraction	*10 µL*	*TUBE 6* (IDT water)
3′ RNA linker (50 µM)	1 µL	*TUBE 1*
10× ligation buffer	2 µL	*TUBE 8*
Ligation enhancer	6 µL	*TUBE 9*
T4 RNA ligase (1 U/µL)	1 µL	*TUBE 13 (diluted)*

The ligation took place in an RNase-free 0.2-ml PCR plastic tube. Add the above reagents into the tube and incubate at room temperature for 2 h. The total liquid volume in the tube should be 20 µL. The T4 RNA ligase from Tube 13 should be diluted from 5 to 1 unit/µL for ligation by the 1× ligation buffer, which

is obtained by a dilution of an aliquot of tube 8 by water from tube 6. Follow the following 12 steps to separate the 3′ linker ligated small RNA from the reaction mixture.

1. Add 80 µL IDTE (pH 7.5) to the reaction tube	*TUBE 7*
2. Transfer entire volume in the tube to an RNase-free 1.5-ml tube	
3. Add 3 µL glycogen (10 mg/mL) to the 1.5-ml tube	*TUBE 11*
4. Add 10 µL 3.0 M NaOAc	*TUBE 12*
5. Add 250 µL –20°C 100% EtOH	
6. Mix by inversion or vortex briefly.	
7. Place the tube at –80°C for 30 min.	
8. Centrifuge at full speed (~16,000 × *g*) for 10 min.	
9. Pour off the supernatant	
10. Dry the pellet completely	
11. Resuspend in 10-µL IDT DNase/RNase/pyrogen-free water.	
12. Mix the 10-µL short RNAs with 2-µL 6× loading dye and load it in one lane or two lanes (Fig. 6) in a 0.75-mm-15% PAGE with 7-M urea. Also, add a 50-bp DNA ladder with steps in another lane of the gel. Drive the samples with 100 V in the second PAGE (0.75-mm 15% gel with 7-M urea) for about 1 h. The gel is cut and the reaction product is extracted by a gel filtration cartridge as described above.	

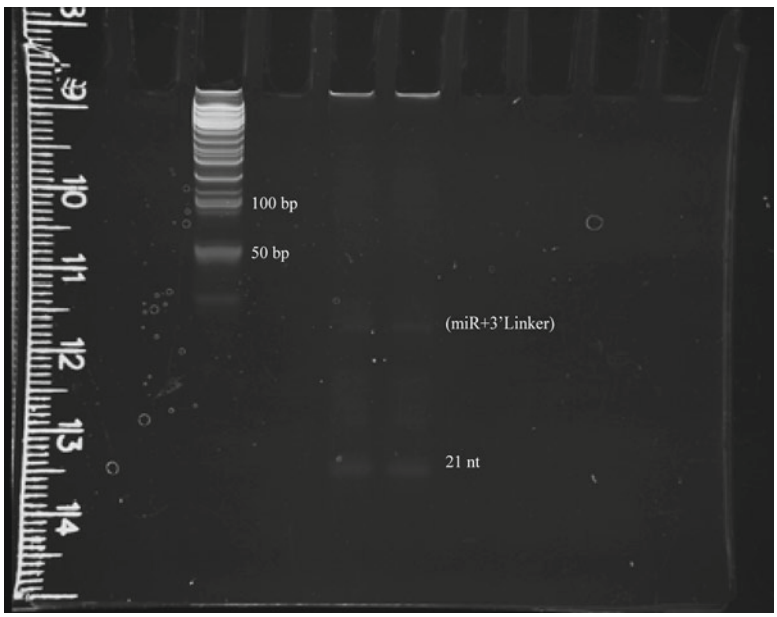


Fig. 6. A UV photo of the second PAGE: Rainbow Trout small RNA and 3′ linker ligation reaction. According to the left ruler on the photo, the 21-nt marker bands were located at 13.5 cm, and the bands of ligation products were located at around 11.7 cm.

3.4. The Ligation of 5′ Linkers to Small RNA Ligated with 3′ Linkers

The next ligation is to link a 5′ linker to a short RNA ligated with a 3′ linker in the presence of 1.0 mM ATP. Spin the tube with purified short RNAs ligated with 3′ linkers in a centrifuge for 10 min at 16,000 × *g*. Discard the supernatant and dry the pellet. Add 8 µL IDT water. Put the vial on ice.

Small RNAs ligated with 3′ linkers	8 µL	IDT water
5′ Linker (50 µM)	1 µL	TUBE 2
10× Ligation buffer	2 µL	TUBE 8
Ligation enhancer	6 µL	TUBE 9
10 mM ATP	2 µL	TUBE 10
T4 RNA ligase (5 U/µL)	1 µL	TUBE 13

In an RNase-free 0.2-mL plastic tube, add the above reagents. The total liquid volume in the tube should be 20 µL. Incubate the tube for the reaction at room temperature for 2 h. Follow the 11 steps given below to finish the 5′ linker ligation reactions so that short RNAs of miRNA size have a same 5′ cap and a same 3′ tail. For example, the sequence of a let-7 miRNA with a 5′ and a 3′ linker should be: 5′-TGGAATucucgggcaccaaggU-ugagguagu-agguuguauaguu-CTGTAG GCACCATCAAT/3ddC-3′.

1. Add 80 µL IDTE (pH 7.5) to the reaction tube.	TUBE 7
2. Transfer entire volume in the tube to an RNase-free 1.5-ml plastic tube.	
3. Add 3 µL glycogen (10 mg/ml) to the 1.5-ml tube.	TUBE 11
4. Add 10 µL 3.0 M NaOAc	TUBE 12
5. Add 250 µL –20°C 100% EtOH	
6. Mix by vortexing briefly	
7. Place tube at –80°C for 30 min.	
8. Centrifuge at full speed (~16,000 × *g*) for 10 min.	
9. Pour off the supernatant	
10. Dry the pellet	
11. Resuspend in 11 µL IDT DNase/RNase/ pyrogen-free water.	

3.5. Reverse Transcription

A short RNA ligated with a 5′ and a 3′ linkers contains both RNA and DNA regions, as shown by small and capital letters in the

sequence of let-7 with a 5′ linker cap and a 3′ linker tail. With the RT/REV primer in tube 5, a reverse transcriptase, SuperScript™ III RT from Invitrogen could synthesize a complementary DNA (cDNA) strand from a single-stranded RNA–DNA hybrid. For example, the cDNA sequence of the let-7 with a 5′ linker cap and a 3′ linker tail should be: 3′-ACCTTAAGAGCCCGTGGTTCCA-ACTCCATCATCCAACATATCAA-GACATCCGTGGT AGTTAG-5′. In this sequence, at the 5′ end between the last two "–" was the sequence of the RT/REV primer in tube 5.

In the recovered sample tube with 11 μL IDT DNase/RNase/pyrogen-free water, add the following reagents:

dNTPs (10 mM)	1.0 μL	
RT primer (10 μM)	1.0 μL	TUBE 5

Incubate tube with 13.0 μL liquid at 65°C for 5 min and then place the tube on ice, add the following reagents into the tube:

5× First strand buffer	4 μL	
0.1 M DTT	1 μL	
IDT water	1 μL	Tube 7
SuperScript III RT (200 U/μL)	1 μL	

The total liquid volume in the tube should be 20.0 μL. In a PCR machine, incubate the tube at 50°C for 1 h, and then change temperature to 70°C for 15 min. This reverse transcription products could be stored at –20°C until needed.

3.6. PCR Amplification and the Third PAGE

Add and mix PCR reagents in a 2-mL plastic tube and then aliquot the mixture in the tube into six 0.2-mL PCR tubes so that each of the PCR tube have the following reagents. The total liquid volume in each of the PCR tube should be 50.0 μL.

1. Reverse transcription product	3.0 μL	
2. IDT water	35.5 μL	TUBE 6
3. 10× PCR buffer	5.0 μL	
4. $MgCl_2$ (1.5 mM)	3.0 μL	
5. dNTPs (10 mM)	1.0 μL	
6. Forward primer (10 μM)	1.0 μL	TUBE 4
7. Reverse primer (10 μM)	1.0 μL	TUBE 5
8. Taq polymerase (5 U/μL)	0.5 μL	

Put the six PCR tubes into a PCR machine, set the PCR conditions as: 10 min at 95.0°C and then 35 cycles of 30 s at 95°C, 30 s at 52°C, and 30 s at 72°C, at last 5 min at 72°C. After the PCR amplification process is finished, use 0.75-mm 15% PAGE with 7 M urea gel to check the 5′ linker ligation results. The size of the sample should be around 65 bp. Figure 7 shows a UV picture of the third PAGE for 10 μL PCR products from one of the six PCR tubes. The products that ran in the lane six and a 50-bp DNA ladder that ran in the lane seven on the gel are shown in the picture. Two clear bands can be seen in the sample lane. From the 50-bp DNA marker band located at 10.5 cm on the ruler shown in the picture, the up sample band located at 10.3 cm should come from short pieces of DNA amplified through cDNAs of short RNAs ligated with 5′ and 3′ linkers. The gel around the up bands is cut to extract PCR products.

3.7. T–A Cloning with Promega® pGEM T Easy Kit

Since a Taq PCR product had a 3′ overhang of adenosine (A), the PCR products could be ligated into a linearized plasmid vector with a complementary single 5′ thymidine (T) overhangs at both ends as shown in Fig. 8. The plasmid vector with two 5′ T overhangs is called a T-vector, and the ligation of the vector with a Taq PCR product containing two 3′ A overhangs is called T–A cloning. Add and mix the following reagents by pipetting into a 200-μL PCR tube. Incubate the reactions overnight at 4°C. It is

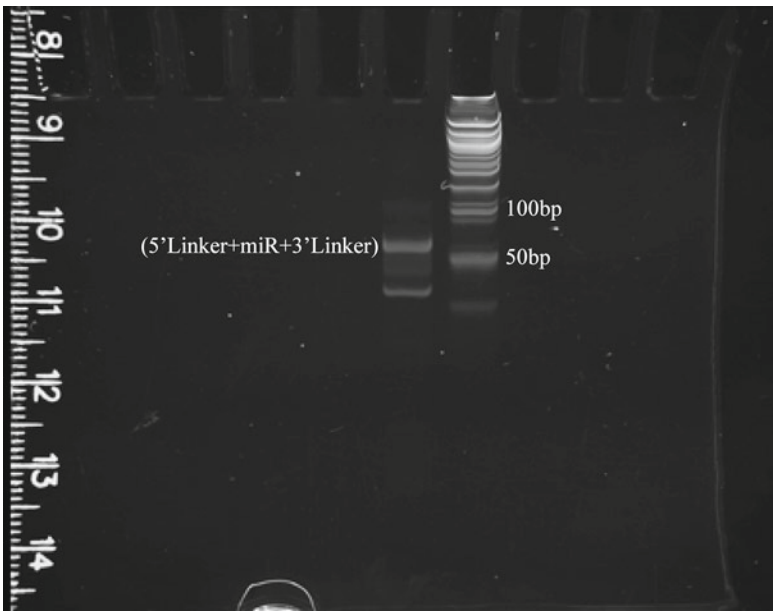


Fig. 7. A UV photo of the third PAGE: The lane 6 and lane 7 were the PCR product of cDNA and a 50 bp DNA ladder, respectively. The band of micro RNAs ligated with a 5′ linker and a 3′ linker appeared at 10.3 cm on the ruler in the photo.

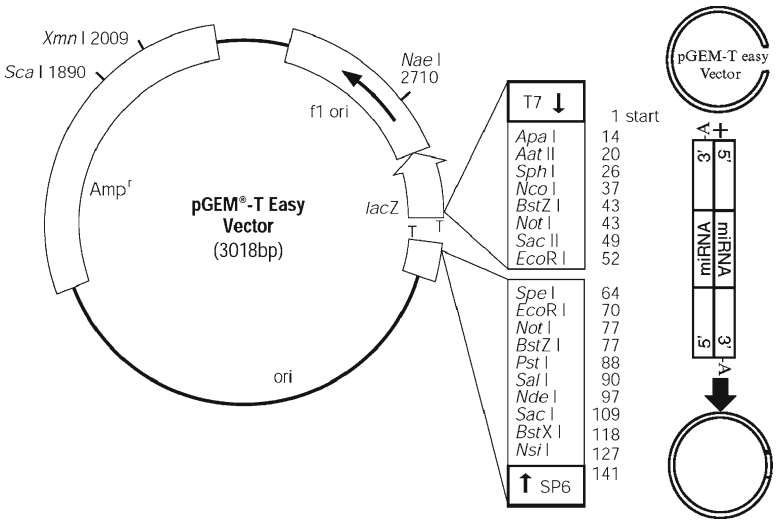


Fig. 8. A pGEM®-T easy plasmid vector circle map and the place for T–A cloning.

important to vortex the 2× rapid ligation buffer before each use. A double-stranded oligonucleotide that came from a short RNA with a 5′ linker cap and a 3′ linker tail would be ligated into the gap in the pGEM-T easy vector as shown in Fig. 8.

2× Rapid ligation buffer	5 μL
pGEM-T easy vector (50 ng)	1 μL
PCR product	3 μL
T4 DNA ligase (3 Weiss units/μL)	1 μL

3.8. Using Heat Shock to Transfect E. coli by Plasmids

1. Remove a tube with frozen JM109 high-efficiency competent *E. coli* cells (Promega® pGEM T easy kit) from −80°C and place it on ice for about 5 min. Mix the just-thawed cells by gently flicking the tube.
2. Take 48 μL of cells into a 200-μL PCR tube with 2 μL TA ligation product. Gently mix the bacteria with the TA ligation product and then place the tube on ice for 20 min.
3. Heat-shock the cells for 50 s in a PCR machine at 42°C, and immediately put the tube on ice for 2 min.
4. Transfer the 50-μL liquid with transfected *E. coli* in the 200-μL PCR tube to a 15-mL snap plastic tube with 950-μL SOC medium (room temperature) and then incubate the tube for 1.5 h at 37°C with shaking (~150 rpm).
5. Plate 100 μL of the transformation culture onto duplicate LB/ampicillin plates.

6. Incubate the plates overnight (16–24 h) at 37°C.
7. Pick 16 colonies with 10-μL plastic tips and drop each of the tips into a 17×100 mm polypropylene tube with 4-mL LB and 100-μg/mL ampicillin. Culture the transformed *E. coli* overnight.
8. Transfer 200 μL liquid from the 4-mL LB bacterial solution to a 1.5-mL tube for each of the 16 tubes, centrifuge the sixteen 1.5-mL tubes for 1 min at 14,000×*g*, and discard the supernatant.
9. Add 20 μL water to each tube, vortex them.
10. Heat at 95°C for 10 min and centrifuge the tubes for 1 min at 14,000×*g*.
11. Transfer 2 μL supernatant from each tube to a 200μL tube for PCR to select *E. coli* colonies transfected with plasmid containing small RNA.
12. PCR preparation: PCR mixture solution 3 μL, M13F/R 2 μL, Taq 0.2 μL and H2O 12.8 μL.
13. Insert the PCR tube with 20-μL liquid into a PCR machine with the setting: 5 min at 94°C, (30 s at 94°C, 30 s at 60°C, 30 s at 72°C)×25cycles, 10 min at 72°C and 4°C.
14. Use a 1.5% agarose gel to check the TA ligation results. The picture in Fig. 9 shows the agarose gel electrophoresis results.

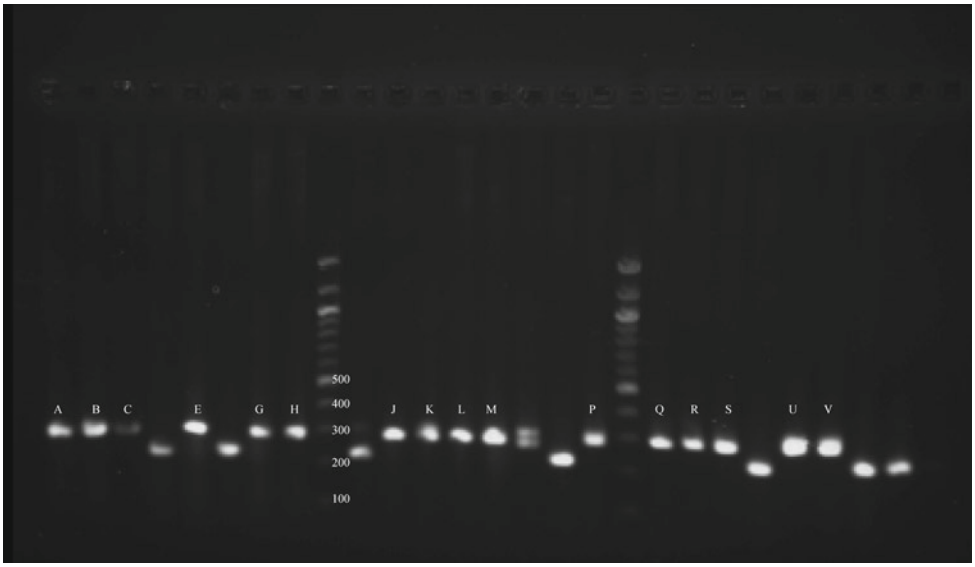


Fig. 9. An UV photo of a 1.5% agarose gel electrophoresis for a fetal bovine ovary microRNA library. The two marker lanes had 100 bp DNA ladders. The 16 up bands close to 300 bp had plasmids ligated with samples, and the seven lower bands close to 200 bp were empty plasmids.

3.9. Purify Plasmid for Sequencing by QIAprep Spin Miniprep Kit

1. To the colonies with inserted plasmids, resuspend the pelleted bacterial cells in 250 μL of buffer P1 and vortex.
2. Add 250 μL of buffer P2 and invert tube gently 4–6 times to mix.
3. Add 350 μL of buffer N3 and invert tube immediately but gently 4–6 times. The solution should become cloudy.
4. Centrifuge for 10 min at maximum speed in a tabletop microcentrifuge. During centrifugation, prepare QIAprep columns.
5. Apply the supernatant (about 850 μL) from step 4 to the QIAprep column by decanting or pipetting.
6. Centrifuge 30–60 s at maximum g-force. Discard the flow-through.
7. Wash QIAprep spin column by adding 0.75 ml of buffer PB and centrifuge 30–60 s. Discard the flow-through.
8. Centrifuge for an additional 1 min to remove residual wash buffer.
9. Place each QIAprep column in a clean 1.5-mL microfuge tube. To elute DNA, add 50 μL buffer EB or sterile Milli-Q H_2O to the center of each QIAprep column, let stand for 1 min, and centrifuge for 1 min.
10. Mark the 1.5-mL microfuge tubes with eluted DNA, and send them out for sequencing.

3.10. Automatically Search Small RNA Between 5′ and 3′ Linkers

ATCACGCGTGGGAGCTCTCCCATATGGTCGACCTGCA
GGCGGCCGCGAATTCACTAGTGATTTGGAATTCT
CGGGCACCAAGGTTTCAGTCATTGTTTCTGG
TAGTCTGTAGGCACCATCAATCAATCGAATTCCCG
CGGCCGCCATGGCGGCCGGGAGCATGCGACGTCG
GGCCCAATTCGCCCTATAGTGAGTCGTATTACAATT
CACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAA
ACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCA
CATCCCCCTTTCGCCAGCTGGCGTAATAGCGAA
GAGGCCCGCACCGATCGCCCTTCCCAACAGTTG
CGCAGCCTGAATGGCGAATGGACGCGCCCT
GTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG
TTACGCGCAGCGTGACCGCTACACTTGCCAGC
GCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTT
CCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAG
CTCTAAATCGGGGGCTCCCTTTAGGGTTCCGAT
TTAGTGCTTTACGGCACCTCGACCCCAAAAAACT
TGATTAGGGTGATGGTTCACGTAGTGGGCCAT
CGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAG
TCCACGTTCTTTAATAGTGGACTCTTGTTCCAAA
CTGGAACAACACTCAACCCTATCTCGGTCTATTC

TTTTGATTTATAAGGGATTTTGCCGATTTCGGC
CTATTGGTTAAAAAATGAGCTGATTTAACAA
AAATTTAACGCGAATTTTAACAAAATATTAACGC
TTACAATTTCCTGATGCGGTATTTTCTCCTT
ACGCATCTGTGCGGTATTTCACACCGCATCAGGT
GGCACTTTTCGGGGAAATGTGCGCGGAANCCC
TATTTGTTTATTTTTCTAAATACATTCAAATATG
TATCCGCTCATGANACAATAACCCTGATAAATG
CTTCAATAATATTGAAAAANGAAGAGTATGAG
TATTCAACATTTTCCNTGTCNCCCNTANTNCCNTT
TTTTGCGGCANTTTGGCNTTCCNGTTTTGGCTCC
CCNAAAANNCTGGGGAAANTAAAAAANNNTAAAANAA
ATTGGGGGNCNNANNGGGTTANANCAANGGAAT
TCAAANNNGNAAANCTTNNAAATTTTNCCCCNAA
AAAAATTTTCAANNAAAAAANNTTTAAANTNTN
TTNNGGGNNGNATTNTCCCNNTNNNCCNGGN
AAAAAAANNNNNNCCCNAANANTTTNAAAAA
ANNGNNNAN

The above is a DNA sequencing text file content from a DNA sequence company of one of our rainbow trout micro RNA samples. We could use Microsoft WORD or another software (bioEditor) to find the insert sequence between the 5′ linker (underline) and the 3′ linker (double underline), but it is not an easy task because sometimes sequence text files are in 3′–5′ order. Even if a sequence text file is in 5′–3′ order, the sequencing text file content is not a single long string. For example, the text shown above was separated into 17 segments, and the 5′ linker was in two segments. Also, because of measurement noises, sometimes a sequencing machine could not determine a nucleotide in a sequence; many "N" symbols were used to indicate the errors.

The combination of all of the problems made finding the insert sequences from thousands of text files by WORD process software a tedious task. For example, for our rainbow trout and fetal ovary micro RNA libraries, we had 1,399 and 685 such text files, respectively. I wrote a MathCAD program that could be used to automatically search microRNA insert sequences from thousands of sequence text files. The program just needed to read a Microsoft Excel file with one column of sequence text file names. A screen print of the program is shown in Fig. 10. There are three parts in the program: The first part is for a Microsoft Excel file input. In the Excel file, one column holds all text file names for each sequence. The second part is the search engine, which is too long to be shown for the chapter. The program is available upon request. The third part is for output. It exported the search result into an Excel file and a table as shown in the figure. If the program could only find 5′ and 3′ linkers, it would report "No-Linker-No-miR". For some special conditions, the program would suggest the user to "search miR by yourself".

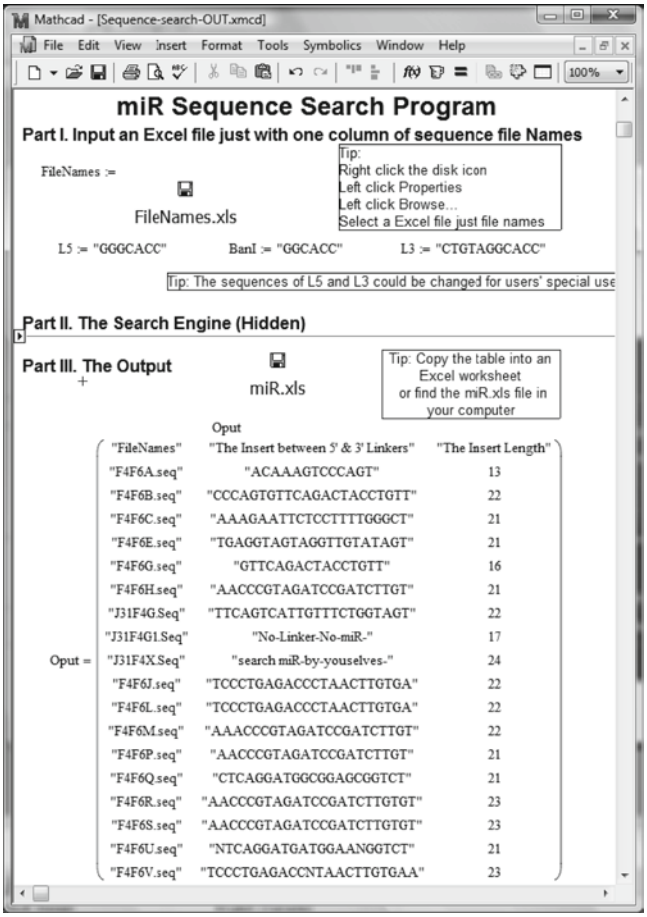


Fig. 10. A screen snap of the text search MathCAD program. The part II of the program was too long to show. In the third part of the program, most of the output sequences came from bacterial colonies shown in Fig. 9.

If higher percentage of bacterial colonies is found to be with microRNA size short RNAs, all the rest PCR products (~240 μL) of the cDNA would be used to run through a 1.5-mM 15% PAGE with 7 M urea gel to select the PCR products. These selected PCR products could be directly sent out for high-throughput DNA sequencing (454 Life Sciences, a Roche). But for traditional DNA sequencing techniques, the selected PCR products must be ligated into plasmids and amplified by bacteria *E. coli*, as in the same steps shown above. After all the insert sequences between the 5′ and 3′ linkers are found by searching thousands of sequencing text files, microRNAs can be found among these inserts by searching a published microRNA database (13).

4. Notes

1. To extract microRNA-size RNA from total RNA, almost all available published protocols suggested using a single concentration (~15%) PAGE. Practically, a single concentration PAGE could handle a very small amount total RNA. To extract enough microRNA-size RNA for a library cloning, people had to run many single concentration PAGEs. It was a tedious job and wasted a lot of 21-nt marker. With a double concentration PAGE (see Subheading 3.2), we could handle hundred micrograms total RNA in a single piece of gel, and only one lane out of five lanes was needed to use the 21-nt marker to locate the microRNA-size RNAs in the gel.
2. The third PAGE is very important (see Subheading 3.6). Without this step, many shorter RNA fragments (something between the two clear bands shown in Fig. 7) would be amplified by bacteria, and a lot of money and time had to be wasted for the sequencing of these no microRNA fragments.

Acknowledgment

All the experimental data were obtained from Dr. Jianbo Yao's Laboratory of Animal Biotechnology and Genomics, Division of Animal and Nutritional Sciences, West Virginia University, Morgantown, WV 26506-6108, USA. Dr. Yao had read the manuscript and gave a permit to publish the paper without his name as a coauthor.

References

1. Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14, *Cell 75*, 843–854.
2. Wienholds, E., and Plasterk, R. (2005) MicroRNA function in animal development, *FEBS Lett 579*, 5911–5922.
3. Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets, *Cell 120*, 15–20.
4. Zhang, B., Wang, Q., and Pan, X. (2007) MicroRNAs and their regulatory roles in animals and plants, *J Cell Physiol 210*, 279–289.
5. Yoon, S., and De Micheli, G. (2005) Prediction of regulatory modules comprising microRNAs and target genes, *Bioinformatics 21(Suppl 2)*, ii93–ii100.
6. Xie, X., Lu, J., Kulbokas, E. J., Golub, T. R., Mootha, V., Lindblad-Toh, K., Lander, E. S., and Kellis, M. (2005) Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals, *Nature 434*, 338–345.
7. Lu, S., Sun, Y. H., Shi, R., Clark, C., Li, L., and Chiang, V. L. (2005) Novel and mechanical stress-responsive MicroRNAs in *Populus trichocarpa* that are absent from Arabidopsis, *Plant Cell 17*, 2186–2203.

8. Yao, Y., Guo, G., Ni, Z., Sunkar, R., Du, J., Zhu, J. K., and Sun, Q. (2007) Cloning and characterization of microRNAs from wheat (*Triticum aestivum* L.), *Genome Biol 8*, R96.
9. http://www.invitrogen.com
10. http://www.idtdna.com/
11. http://www.bio-rad.com/
12. Lau, N., Lim, L., Wienstein, E., and Bartel, D. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*, *Science 294*, 858–862.
13. http://microrna.sanger.ac.uk/

Chapter 11

MicroRNA Regulation of Growth Factor Receptor Signaling in Human Cancer Cells

Keith M. Giles, Andrew Barker, Priscilla M. Zhang, Michael R. Epis, and Peter J. Leedman

Abstract

Aberrant expression of the epidermal growth factor receptor (EGFR) and/or human epidermal growth factor receptor 2 (HER2) is a feature of many human tumors and is associated with disease progression, treatment resistance, and poor prognosis. Protein kinase B/Akt, an important downstream effector of these receptor tyrosine kinases, induces signaling pathways that control cancer cell proliferation, invasion, angiogenesis, and apoptosis resistance. MicroRNAs (miRNAs), small noncoding RNAs that bind to the 3′-untranslated region of target mRNAs, are now recognized to play key roles in the regulation of gene expression, particularly in tumor development and metastasis. We have shown that miRNA-7 (miR-7) and miRNA-331-3p (miR-331-3p) directly regulate expression of EGFR and HER2, respectively, in glioblastoma and prostate cancer cell lines. As a consequence, miR-7 and miR-331-3p reduce Akt activity and thus have the capacity to regulate a signaling pathway critical to the development and progression of glioblastoma and prostate cancer. This chapter provides a detailed approach outlining how to confirm that a putative target of a miRNA is a direct target, and subsequent assessment of downstream signaling mediators.

Key words: MicroRNA, EGFR, HER2, miR-7, miR-331-3p, Glioblastoma, Prostate cancer, Transfection, Immunoblotting, Luciferase reporter gene assay

1. Introduction

Overexpression of the epidermal growth factor receptor (EGFR) and/or human epidermal growth factor receptor 2 (HER2) receptor tyrosine kinases occurs in many human tumors and is associated with disease progression, treatment resistance, and poor prognosis (1, 2). Protein kinase B/Akt is an important downstream effector of EGFR and HER2 signaling which is frequently activated in human cancers and regulates pathways that

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_11, © Springer Science+Business Media, LLC 2011

control cancer cell proliferation, invasion, angiogenesis, and apoptosis resistance (3). Thus, EGFR and HER2 represent important targets for the development of new therapeutics, including small molecule tyrosine kinase inhibitors such as erlotinib (4) and monoclonal antibodies such as trastuzumab (5).

MicroRNAs (miRNAs) comprise a large family of small, ~22 nucleotide, noncoding RNAs that bind with imperfect complementarity to the 3′-untranslated regions (3′-UTRs) of target mRNAs, causing translational repression or message degradation (6, 7). miRNAs have important roles in normal cellular development and function (8, 9), and altered expression of miRNAs is associated with cancer (10).

Bioinformatic analyses (TargetScan; (11)) predicted EGFR and HER2 as putative targets of miR-7 and miR-331-3p, respectively. Using glioblastoma and prostate cancer cell lines, we have demonstrated that these miRNAs directly block EGFR and HER2 expression and that this is associated with reduced activity of the downstream Akt signaling pathway (12, 13).

Akt is activated by phosphatidylinositol 3-kinase (PI3K)-dependent phosphorylation at serine 473 (14). Antibodies recognizing phosphorylated and total Akt are commercially available and allow direct assessment of the impact of a specific miRNA on this signaling pathway by immunoblotting. Similarly, EGFR and HER2 expression can be assessed following transfection of cancer cell lines with miRNA, by using commercially available antibodies in immunoblotting assays. Luciferase reporter assays enable confirmation that miR-7 and miR-331-3p directly target specific binding sites within the 3′-UTRs of EGFR and HER2 mRNA. Together, these techniques are essential for confirming miRNA regulation of a putative mRNA target.

2. Materials

2.1. Cell Culture and Transfections with miRNA Precursors and Luciferase Reporter Plasmids

1. Dulbecco's modified Eagle's medium (DMEM) (Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (FBS; Gibco/BRL, Bethesda, MD) (see Note 1), Opti-MEM reduced serum media (from Gibco/BRL), phosphate buffered saline (PBS; Invitrogen).
2. TrypLE express (Invitrogen), a temperature-stable alternative to trypsin.
3. Synthetic precursor miRNA molecules (Ambion/Applied Biosystems, Austin, TX) are resuspended at 1 or 30 μM in nuclease-free water and stored in aliquots at –20°C (see Note 2).
4. Lipofectamine 2000 transfection reagent (Invitrogen) (see Note 3).

5. Firefly luciferase reporter plasmid (pGL3-control; Promega, Madison, WI; or pMIR-REPORT-luciferase, Ambion) containing an insert from EGFR or HER2 mRNA 3′-UTR immediately downstream of the firefly luciferase open reading frame (ORF) at a concentration of 100 ng/μl, and *Renilla* luciferase reporter plasmid (pRL-SV40, Promega) for normalization at a concentration of 1 ng/μl (see Note 4).

2.2. Preparation of Protein Extracts for SDS-PAGE and Immunoblotting

1. Cell scrapers (Corning Life Sciences, Lowell, MA).
2. Cytoplasmic extraction buffer (CEB) for cell lysis: 10 mM HEPES pH 7.5, 3 mM $MgCl_2$, 40 mM KCl, 5% glycerol, 0.2% Nonidet P-40. Sterilize by autoclaving and store at 4°C. Add one complete Mini EDTA-free Protease Inhibitor Cocktail tablet (Roche, Indianapolis, IN) and one PhosSTOP Phosphatase Inhibitor Cocktail tablet (from Roche) per 10 mL of CEB prior to use, to inhibit protease and phosphatase activity in cell lysates.
3. Bio-Rad protein assay reagent (Bio-Rad, Hercules, CA) for determination of total protein concentrations from cell lysates.

2.3. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

1. NuPAGE Novex 4–12% Bis-Tris Precast Gels (from Invitrogen). Store at 4°C.
2. XCell SureLock Mini-Cell SDS-PAGE gel electrophoresis system (Invitrogen).
3. NuPAGE MOPS SDS running buffer (20×; for Bis-Tris Gels) (Invitrogen). Dilute to 1× with deionized water.
4. NuPAGE sample reducing agent (10×; Invitrogen) and NuPAGE antioxidant (Invitrogen). Store at 4°C.
5. Precision plus protein kaleidoscope prestained SDS-PAGE markers (Bio-Rad). Store at –20°C.

2.4. Immunoblotting for EGFR, HER2, β-Actin, Active Akt and Total Akt

1. Transfer buffer: 25 mM Tris, 190 mM glycine, 10% (v/v) methanol, 0.1% (v/v) NuPAGE antioxidant (Invitrogen).
2. Polyvinylidene difluoride (PVDF) membrane (Roche).
3. Precut filter paper (90 × 70 mm) (Bio-Rad).
4. Tris-buffered saline with Tween (TBS-T): 20 mM Tris–HCl, pH 7.4, 150 mM NaCl, 0.1% Tween 20.
5. Blocking buffer: 5% (w/v) nonfat dried milk in TBS-T.
6. Primary antibodies (see Note 5): anti-EGFR (Abcam, UK, ab52894; 1:5,000), anti-HER2 (Abcam, ab8054-1; 1:1,000), anti-phospho-HER2 (Abcam, UK, ab47755-100; 1:1,000), anti-Akt (Cell Signaling Technology, Danvers, MA, #9272; 1:1,000), anti-phospho-Akt (Ser-473; Cell Signaling Technology, #4060; 1:500), anti-β-actin (Abcam,

ab6276; 1:15,000). Dilute antibodies in 1% (w/v) nonfat dried milk in TBS-T prior to use.

7. Secondary antibodies: antimouse/rabbit IgG conjugated to horse radish peroxidase (HRP; General Electric Healthcare, Uppsala, Sweden). Dilute antibodies to 1:10,000 (v/v) in 5% (w/v) nonfat dried milk in TBS-T prior to use. The appropriate secondary antibodies are used according to the species in which the primary antibody to be detected was raised. For instance, an antimouse IgG secondary antibody is used to detect a mouse primary antibody raised against the protein of interest.
8. Enhanced chemiluminescence (ECL) plus reagents and ECL-Hyperfilm (from GE Healthcare).

2.5. Luciferase Reporter Assay

1. Dual-luciferase reporter assay system (from Promega), consisting of passive lysis buffer (5×), luciferase assay substrate, luciferase assay buffer II, Stop and Glo substrate (50×), and Stop and Glo buffer. Dilute passive lysis buffer to 1× working stock in distilled water immediately before use. Resuspend lyophilized luciferase assay substrate in luciferase assay buffer II; this can be stored at –20°C for up to 1 month or at –80°C for up to 1 year. Dilute Stop and Glo substrate to 1× final concentration in Stop and Glo buffer before use. If necessary, this reagent may be stored at –20°C for up to 2 weeks.
2. White-walled 96-well plates (Costar; from Corning Life Sciences) for measurement of luciferase activities in cell lysates.

3. Methods

The confirmation of EGFR and HER2 mRNA as direct targets of miR-7 and miR-331-3p, respectively, requires evidence (1) that transfection of EGFR/HER2-expressing human cancer cells with miR-7/miR-331-3p represses EGFR or HER2 protein expression and (2) that transfection of cells with miR-7/miR-331-3p directly blocks expression of a chimeric reporter gene consisting of the EGFR/HER2 mRNA 3′-UTR fused to a firefly luciferase ORF. Importantly, mutation of the putative miRNA binding site within the 3′-UTR predicted to abrogate binding by the miRNA should be associated with reversal of the effect, providing critical validation that the site is directly bound by the miRNA of interest. SDS-PAGE and immunoblotting is used to demonstrate specific regulation of EGFR and HER2 protein expression in cancer cell lines by miR-7 and miR-331-3p. Further insight into the mechanism of this regulation can be obtained by analyzing

the effect of the miRNA upon target mRNA expression. If the transfected miRNA represses expression of its target protein and its target mRNA, one might conclude that the miRNA induces decay of that mRNA; on the other hand, if the miRNA downregulates expression of its target protein, yet does not induce significant downregulation of its target mRNA, then it is more likely that this regulation occurs at the translational level. Using quantitative RT-PCR, we have shown that miR-7 and miR-331-3p regulate degradation of EGFR and HER2 mRNA, respectively (12, 13).

Ultimately, it is important to determine whether miR-7 and miR-331-3p regulate signaling pathways that are critical to tumor formation and progression. This can be accomplished by SDS-PAGE and immunoblotting with lysates harvested from cancer cell lines transfected with miR-7 or miR-331-3p. It is important that cellular proteins are harvested with lysis buffer containing inhibitors to prevent protease and phosphatase activity. Activation of Akt requires phosphorylation at serine 473 in a PI3K kinase-dependent manner (14), and thus, the use of a phosphospecific antibody against serine 473 of Akt allows assessment of the degree of Akt activation in a protein sample by immunoblotting. miR-7 and miR-331-3p directly target the EGFR and HER2 mRNA 3′-UTRs, blocking expression of these target molecules and reducing activity of the Akt signaling pathway in glioblastoma and prostate cancer cells (see data figures).

3.1. Reverse Transfection of U87MG and LNCaP Human Cancer Cell Lines with miRNA Precursors and Luciferase Reporter Plasmids

1. U87MG human glioblastoma, LNCaP human prostate cancer, and A549 human non-small cell lung cancer cells are passaged when approaching confluence with TrypLE express to provide new maintenance cultures in 175 cm^2 tissue culture flasks and experimental cultures in six-well plates for transfection with miRNA precursor molecules to obtain cell lysates for immunoblotting, or 24-well plates for luciferase reporter gene assays.
2. Materials required for the reverse transfection protocol are prepared: miRNA precursor molecules are diluted in nuclease-free water at appropriate stock concentration for dilution at 1:1,000 into cultures (see Note 6); plasmid DNA is prepared at 100 ng/μL (pGL3-based firefly luciferase vector) or 10 ng/μL (pMIR-REPORT-based firefly luciferase vector), and 1 ng/μL (*Renilla* luciferase vector) in nuclease-free water; Optimem and Lipofectamine 2000 are brought to room temperature; cancer cells are harvested by trypsinization, counted using a hemocytometer, and resuspended at $1.0–2.0 \times 10^5$ cells/mL (see Note 7).
3. The experiment should include replicate wells for the miRNA of interest, for a negative control (NC) miRNA and transfection reagent (Lipofectamine 2000) only. For luciferase reporter

gene assays, plasmids carrying mutated miRNA binding sites and/or empty reporter DNA (no miRNA binding sites) should also be included.

4. Transfection format for six-well plates for cell lysis and immunoblotting: prepare master mixes for transfection, assuming a final miRNA precursor molecule concentration of 30 nM, as follows (per 1× well):

	Part A	Part B
Optimem	241.67 µL	248.00 µL
Lipofectamine 2000	8.33 µL (see Note 8)	
miRNA precursor (30 µM stock)		2.00 µL
Total	250 µL	250 µL

5. Transfection format for 24-well plates for luciferase reporter gene assays: prepare master mixes for transfection, assuming a final miRNA precursor molecule concentration of 1 nM, as follows (per 1× well) (see Note 9):

	Part A	Part B
Optimem	48.5 µL	43.4 µL
Lipofectamine 2000	1.5 µL (see Note 8)	
miRNA precursor (1 µM stock)		0.6 µL
Firefly luciferase plasmid (100 ng/µL)		1.0 µl
Renilla luciferase plasmid (1 ng/µL)		5.0 µl
Total	50 µL	50 µL

6. Prepare Part A, mix gently, and incubate at room temperature for 5 min. Meanwhile, prepare Part B. Add Part B to Part A in a dropwise fashion, mix gently, and incubate at room temperature for 20 min to allow formation of transfection complexes.
7. Aliquot 500 µL of the combined transfection mixture to each well of a six-well plate (or 100 µL to each well of a 24-well plate). Carefully overlay the transfection mixture with 1.5 mL of cell suspension (for a six-well plate; $1.5–3.0 \times 10^5$ cells per well; total volume of cells plus transfection mixture is 2 mL per well) or 0.5 mL of cell suspension (for a 24-well plate; 0.5–1.0×105 cells per well; total volume of cells plus transfection

mixture is 0.6 mL per well). Swirl the plates gently to ensure even coverage and place in a 37°C cell culture incubator at 5% CO_2 for 72 h (prior to protein extraction and SDS-PAGE and immunoblotting), or 24 h (for luciferase reporter gene assays) (see Note 10).

3.2. Preparation of Cell Lysates for Immunoblotting Analysis of miRNA Target Gene Expression

1. Cells are removed from the incubator and placed on ice. Media is removed and cells are rinsed once gently with 2 mL of cold PBS per well of a six-well plate. After removal of PBS, add 80–200 µL of CEB lysis buffer (containing fresh protease and phosphatase inhibitors; see Subheading 2.2) to each well and return plates to ice.
2. Scrape and collect cells to 1.5-mL Eppendorf tubes and leave on ice for 20–30 min to permit cell lysis (see Note 11). Clear lysates by centrifugation at 13,000 × *g* at 4°C for 5 min, transfer supernatants to new Eppendorf tubes and store at –80°C.
3. Prior to SDS-PAGE, determine total protein concentrations of lysates using the Bio-Rad protein assay (based on the Bradford method (15)). The inclusion of standards (bovine serum albumin) ranging from 0.5 to 5.0 mg/mL will allow comparison to a standard curve, providing a relative measurement of total protein concentration.

3.3. SDS-PAGE

1. SDS-PAGE and immunoblotting is used for the analysis of relative levels of specific proteins in cell lysates (prepared in Subheading 3.2). The technique, derived from Laemmli (16) involves electrophoretic separation of total cellular proteins based on their molecular weight on a polyacrylamide gel under denaturing conditions, and their transfer onto a PVDF membrane for detection of proteins of interest (Subheading 3.4).
2. Sample preparation: in a 1.5 mL tube, place an appropriate volume of protein lysate corresponding to 15 µg of protein, based on the Bio-Rad protein assay protein quantitation results (see Subheading 3.2 in step 3). To each of these tubes, add 2 µL of 10× reducing agent, 5 µL of 4× LDS sample buffer, and distilled water to a final volume of 20 µL.
3. Centrifuge sample tubes briefly to ensure all components are collected at the base of the tube, and incubate at 70°C for 10 min to denature proteins. After heating, mix samples and centrifuge tubes briefly to collect samples at the bottom of the tubes, and cool on ice.
4. Prepare the apparatus for SDS-PAGE: assemble the XCell SureLock Mini-Cell SDS-PAGE gel electrophoresis tank according to the manufacturer's instructions (Invitrogen). After removing well combs and sticker lining, rinse NuPAGE

Novex 4–12% Bis-Tris Precast Gels with distilled water and secure into place in the electrophoresis tank.

5. Fill the inner chamber of the tank (chamber between the gels) with 1× NuPAGE MOPS SDS running buffer, supplemented with 500 μL of NuPAGE antioxidant, until the buffer level covers the wells. Fill the outer chamber with 200 mL of the same running buffer without NuPAGE antioxidant.
6. Load the protein samples into wells of the gel. In a separate well on each gel, load a suitable volume of MW markers (e.g., Bio-Rad Precision Plus Protein Kaleidoscope Standards).
7. Run gels at 150 V (constant voltage) and 4°C until the dye front has almost reached the base of the tank, or until the MW region where the proteins of interest lie is well separated. During this time, prepare the transfer buffer and prechill at 4°C.

3.4. Immunoblotting for EGFR, HER2, β-Actin, Active Akt, and Total Akt

1. This protocol is derived from the method of Towbin and coworkers (17). Transfer: following SDS-PAGE, disassemble the gel apparatus and casting plates, and transfer gels into the cold transfer buffer.
2. Assemble the Bio-Rad transfer cassettes by placing, in order, on the "black" side of the cassette, a transfer sponge, a piece of precut filter paper, the polyacrylamide gel, a piece of PVDF Western blotting membrane that has been prewetted with 100% methanol, a piece of precut filter paper, and a transfer sponge. Presoak all items with transfer buffer. A 15-mL centrifuge tube may be used to smooth out air bubbles while setting up the transfer cassettes. This ensures that all surfaces are in contact, allowing for efficient transfer. Close the cassette and place into the Bio-Rad transfer tank (containing a magnetic stirrer bar) such that the PVDF membrane is between the gel and the anode.
3. Perform the transfer overnight at 25 V (constant voltage) at 4°C, with gentle stirring.
4. After transfer, the protein-bound PVDF membranes may either be rinsed in TBS-T and stored at 4°C for use at a later date, or transferred directly into blocking buffer for immunoblotting.
5. Blocking: cover the PVDF membranes with an adequate volume of blocking buffer (see Note 12) and gently rock at RT for 1 h.
6. Using the MW markers as a guide, trim and divide the PVDF membranes into smaller blots to fit trays, in order to immunoblot for the desired proteins. Place the diluted primary antibodies in the trays containing the appropriate blots (based

on the predicted MW for each protein of interest), and incubate for 1 h at RT, with gentle rocking.

7. Remove the primary antibodies and wash the blots for 3 × 10 min in blocking buffer.
8. Incubate blots with the diluted secondary antibodies for 1 h at RT, with gentle rocking.
9. Discard the diluted secondary antibody solutions and wash the blots for 3 × 10 min in TBS-T.
10. Freshly prepare an ECL Plus (see Note 13) reagent mixture (at a ratio of 1 mL of solution A to 25 µL of solution B), and incubate on the PVDF blots for 5 min. Drain off excess ECL reagent and place blots between plastic sheets. Place the sheets containing the blots into a cassette and expose to ECL-Hyperfilm in a dark room for 1 s–15 min as required, after which film is processed in a film developer. Examples of the signals for EGFR, HER2, β-actin, Akt, and phospho-Akt are shown in Fig. 1.

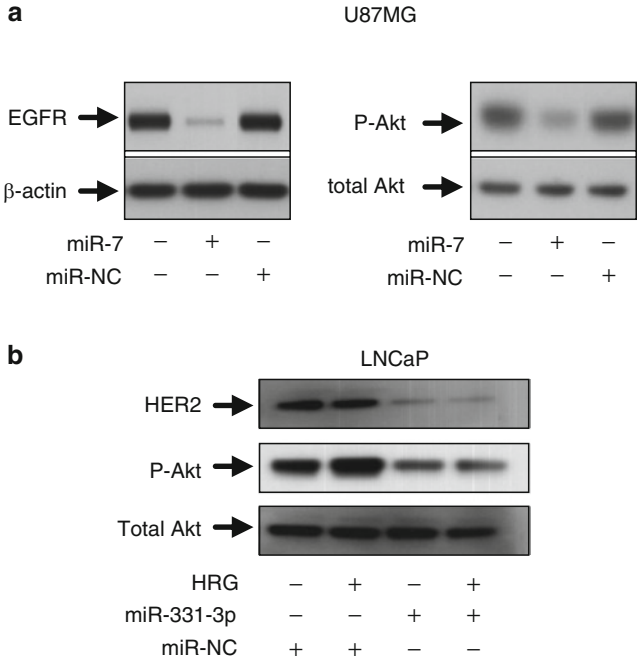


Fig. 1. Regulation of EGFR and HER2 expression and Akt activity by miR-7 and miR-331-3p in U87MG and LNCaP cells. (**a**) U87MG and (**b**) LNCaP cells were reverse transfected with miR-7 or miR-331-3p for 3d, and processed for immunoblotting analysis of EGFR, HER2, total Akt, phosphorylated Akt (P-Akt), and β-actin expression. In some wells, LNCaP cells were stimulated with heregulin (HRG; 50 ng/mL for 20 min) to activate P-Akt via HER2 signaling; this effect was blocked by miR-331-3p.

3.5. Luciferase Reporter Assay

1. This procedure involves the use of Promega's dual luciferase reporter assay system for quantitation of firefly and *Renilla* luciferase activities from a single lysate of cells cotransfected with firefly and *Renilla* luciferase plasmids and miRNA precursors. The 3′-UTR sequence containing a putative miRNA binding region is present within the firefly luciferase vector, whereas the *Renilla* luciferase vector is used for normalization of transfection efficiency and to correct for nonspecific effects of miRNA precursor transfection.
2. At 24-h post transfection with miRNA precursors and luciferase plasmids, remove cell culture media, rinse wells carefully with 1 mL of PBS, and remove PBS.
3. Add 75 µL of 1× passive lysis buffer to each well of the 24-well plate that contains cells. Since scraping is difficult in wells of this size, passive lysis is recommended, with plates being placed on a rocking platform at room temperature for 15 min. Alternatively, place the plates at –80°C; they can be stored here for up to several weeks for analysis of luciferase activities, otherwise, a single freeze-thaw cycle will improve cell lysis (see Note 14).
4. Clear the cell lysates by transferring them to 1.5-mL Eppendorf tubes and centrifuging at top speed at 4°C for 3 min.
5. Transfer 20 µL of cleared supernatant to a white-walled 96-well plate prior to reading.
6. Prime a luminometer fitted with reagent injectors with luciferase assay reagent II and Stop and Glo reagent. Measure firefly luciferase activities by injecting 50 µL of luciferase assay reagent II into each well, and measure *Renilla* luciferase activities by injecting 50 µL of Stop and Glo reagent into each well. A 2-s premeasurement delay should be followed by a 10-s measurement period for each reporter assay.
7. Process raw data by dividing each firefly luciferase measurement by its corresponding *Renilla* luciferase measurement, giving corrected/normalized luciferase activity. Statistical analysis of the data should also be performed using Microsoft Excel or other statistical software. Examples of normalized luciferase data are shown in Figs. 2e, f and 3.

Fig. 2. (continued) (**f**) Luciferase reporter assay with A549 cells that were transfected with WT or MT EGFR target site (A, B, C, or D) 3′-UTR firefly-luciferase reporter, control *Renilla* luciferase, and miR-7 or miR-NC precursor. Relative luciferase expression values are the ratio of miR-7-treated reporter vector compared with the same miR-NC-treated reporter (±SD), and demonstrate the reduced expression of reporters bearing putative target sites B and C, but not A, by miR-7. Furthermore, the presence of sites B and C in the same reporter construct (EGFR D) conferred additive repression with miR-7. In all cases, mutation of the miR-7 seed-binding regions (B-MT, C-MT, D-MT) abrogated the repression of reporter gene expression by miR-7. *Asterisk* indicates a significant difference from miR-NC-transfected controls ($p < 0.05$) (Reproduced from ref. (12) with permission from The American Society for Biochemistry and Molecular Biology.).

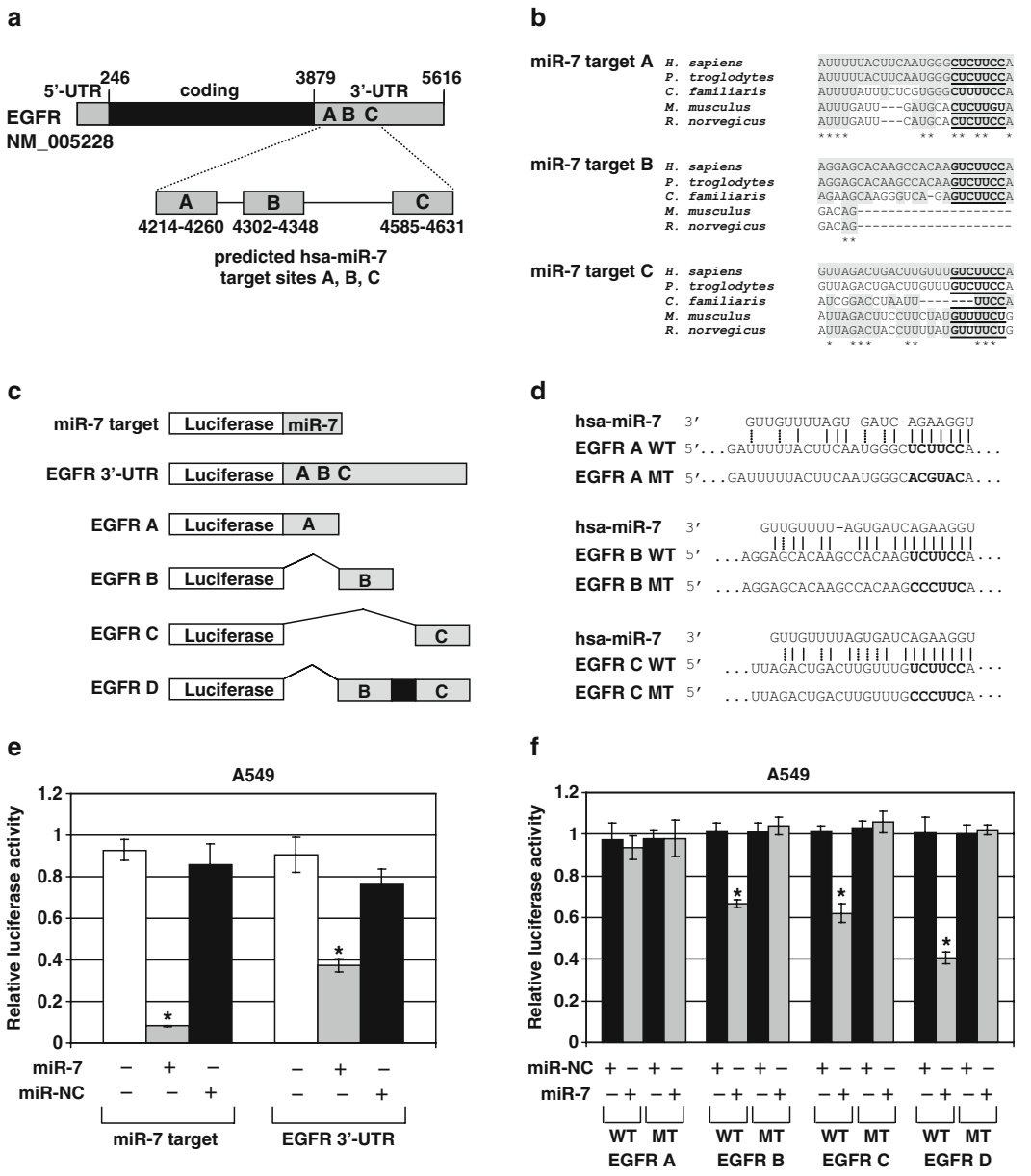


Fig. 2. Identification of two specific miR-7 target sites within the non-conserved EGFR mRNA 3′-UTR. (**a**) Schematic representation of the EGFR mRNA with three 3′-UTR miR-7 binding sites (A, B, C). (**b**) Sequence alignment of predicted EGFR mRNA 3′-UTR miR-7 target sites showing lack of conservation between species. The miR-7 seed target sequence (UCUUCC) is in *bold* and *underlined*, conserved nucleotides are *shaded*, and *stars* indicate nucleotides conserved across all species. (**c**) Schematic representation of firefly luciferase reporter constructs for miR-7 perfect target, full-length wild-type EGFR 3′-UTR, and truncated EGFR 3′-UTR (A, B, C, D) target sites. (**d**) Sequence of wild-type (WT) and mutant (MT) EGFR mRNA 3′-UTR miR-7 target sites. Mutations predicted to disrupt miRNA:mRNA binding were made in the seed target region. (**e**) Luciferase reporter assay to verify activity of miR-7 upon a miR-7 target site and full-length wild-type EGFR 3′-UTR. A549 lung cancer cells were transfected with perfect miR-7 target or full length wild-type EGFR 3′-UTR firefly luciferase plasmid and miR-7 or miR-NC precursor. Relative luciferase expression (firefly normalized to *Renilla*) values are expressed as a ratio of reporter vector alone (±SD). Data are representative of at least three independent experiments. *Asterisk* indicates a significant difference from vehicle (Lipofectamine 2000) – treated reporter vector ($p < 0.001$).

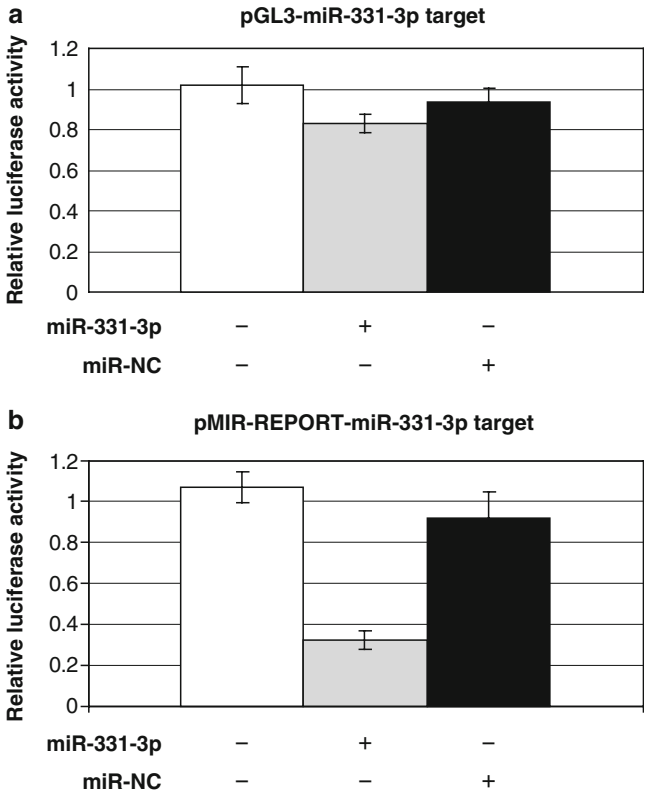


Fig. 3. Repression of pMIR-REPORT-miR-331-3p perfect target reporter, but not pGL3-miR-331-3p perfect target reporter, by miR-331-3p. LNCaP prostate cancer cells were cotransfected with miR-331-3p precursor, a control *Renilla* luciferase plasmid (pRL-SV40), and either pGL3-miR-331-p perfect target (**a**) or pMIR-REPORT-miR-331-3p perfect target (**b**) firefly luciferase reporter plasmid. Relative luciferase expression values were calculated by normalizing firefly luciferase values to their corresponding *Renilla* luciferase measurements (±SD), and the values demonstrate reduced expression of a luciferase reporter gene carrying a consensus miR-331-3p target site within pMIR-REPORT, but not pGL3-control reporter vector.

4. Notes

1. Cell culture media for transfection experiments may contain serum, but should not contain antibiotics, as this will cause toxicity when cells are exposed to the transfection reagent.
2. Despite their name, Ambion's pre-miR molecules should not be confused with precursor miRNA (pre-miRNA) hairpin structures; rather, they mimic the mature miRNA. It is important that a negative control (nontargeting) miRNA precursor molecule is purchased together with the miRNA of interest to determine whether effects observed on target gene expression are specific. Other companies (e.g., QIAGEN, Dharmacon)

provide miRNA "mimics" that can be used as alternatives to the miRNA precursor molecules sold by Ambion/Applied Biosystems. Although we do not detail such experiments here, miRNA inhibitors can also be used with cell lines that express appreciable amounts of a specific miRNA, with the "read out" being an upregulation of the target gene or reporter upon treatment with the specific miRNA inhibitor. miRNA inhibitors incorporating the locked nucleic acid (LNA) technology, developed by Exiqon (Vedbaek, Denmark) can be used interchangeably with the Ambion miRNA precursors in the protocols described here.

3. Lipofectamine 2000 is an effective cationic lipid-based transfection reagent, designed to give high transfection efficiency with minimal toxicity to cells. It can be used for simultaneous transfection of miRNA/siRNA and plasmid DNA. Invitrogen provides information about the use of specific cell lines with Lipofectamine 2000 at http://www.invitrogen.com/celllines. In our hands, Lipofectamine 2000 has proved the only transfection reagent that reliably transfects plasmid DNA and small RNA simultaneously. We do, however, note some variability in the transfection efficiency achieved with different aliquots of Lipofectamine 2000 obtained from the manufacturer, even with aliquots bearing the same lot number.
4. Plasmid DNA for transfection should be prepared using Maxiprep or Midiprep protocols and should be of good quality and purity: predominantly supercoiled, with 260/280 nm absorbance ratios of 1.8–2.0. The insert orientation is confirmed by restriction digestion, and the correct insert sequence is verified by DNA sequencing. Plasmid DNA can be stored at –20°C, but care should be taken to avoid repeated freeze–thawing. Predicted target sequences can be inserted at restriction enzyme sites downstream of the firefly luciferase ORF using standard molecular cloning techniques. Alternatively, several companies offer reporter vectors containing synthesized 3′-UTR target sequences (e.g., GenScript, Piscataway, NJ; Genecopoeia, Rockville, MD; SwitchGear Genomics, Menlo Park, CA). We would suggest initial assessment using the full-length 3′-UTR of the putative target mRNA in luciferase reporter assays. If the results suggest that the miRNA of interest regulates reporter activity of the 3′-UTR, then vectors containing shorter focused regions spanning the putative miRNA target sequence can be generated and tested (as described in (12, 13)). It is recommended that "perfect target" reporter vectors are used as a positive control for cotransfection with miRNA precursor molecules. These vectors contain a target sequence with perfect complementarity to the miRNA of interest; thus, expression of the reporter

gene should be significantly repressed with cotransfection of the miRNA. It is important to use this control to establish that the miRNA precursor was effectively transfected and can block gene expression via binding to its target sequence. In some instances, we have cloned a perfect target sequence into two different reporter vectors and observed repression of one vector but not the other with miRNA precursor. An example of this is shown in Fig. 3, where miR-331-3p blocked luciferase expression from pMIR-REPORT-miR-331-3p target but not from pGL3-miR-331-3p target. We presume this is due to differences in the sequence context of the 3′-UTRs of the two reporter constructs, as differing 3′-UTR RNA secondary structures can affect the interaction between a miRNA and its target (18). Finally, to determine the specificity of a miRNA–mRNA interaction, mutant miRNA target sites should be incorporated in reporter gene assays, in which nucleotide substitutions are introduced within the seed-binding region of the miRNA target site (see Fig. 2d, f).

5. Primary antibodies can be stored at −20°C after use and thawed several times later for reuse. We have tested antibodies to cancer cell signaling molecules from a number of suppliers and recommend the ones listed for immunoblotting experiments.
6. It is suggested that 1 and 30 μM working stocks of miRNA precursor molecules are prepared in nuclease-free water to be used 1:1,000 in transfection reactions, yielding a final miRNA precursor concentration of 1 or 30 nM. This will generally permit changes in target expression (endogenous protein or reporter) to be observed in luciferase assays or by immunoblotting. However, the optimal final concentration of miRNA precursor used should be determined by titration and will typically be between 1 and 50 nM in transfections for immunoblotting, and between 0.1 and 10 nM in transfection for reporter gene assays. Higher concentrations may increase the possibility of nonspecific/off-target effects.
7. Cells should be resuspended at $1.0–2.0 \times 10^5$ cells/mL. For transfections in six-well plates, we routinely use 1.5 mL of this cell suspension per well (i.e., $1.5–3.0 \times 10^5$ cells per well). For transfections in 24-well plates we use 0.5 mL of this suspension per well (i.e., $0.5–1.0 \times 10^5$ cells per well). The precise cell number to be used in the transfection will depend on the cell line being used as well as the time point at which samples are to be harvested. Typically, we harvest protein samples for immunoblotting at 72-h posttransfection although the time needed to observe changes in miRNA target protein expression may vary depending on the protein and its endogenous half-life. The rate of cell growth and the sensitivity

of the cell line to transfection reagent will also have a bearing on the number of cells to be used. We routinely use reverse transfection (i.e., transfection of cells in suspension prior to their attachment to the plate) as this favors higher transfection efficiency and saves time by eliminating the need to preplate cells before transfection (forward transfection). Certain cell lines (e.g., MDA-MB-468) are particularly sensitive to reverse transfection, and better cell viability will be achieved with forward transfection. In such cases, the basic transfection method described above can be used, except transfection mixtures are added to cells 24 h after plating instead of together with cells in suspension.

8. The volume of Lipofectamine 2000 used per well should be determined empirically for each cell line and will depend on the sensitivity of the cell line to the transfection reagent, the transfection efficiency, and results obtained. Volumes between 5 and 10 μL per well can be tested for six-well plates, and between 1.5 and 3 μL per well for 24-well plates.
9. Cotransfection of plasmid DNA and miRNA precursors permits greater sensitivity in the detection of miRNA-dependent regulation of reporter gene expression. Once the reporter has been translated, it cannot be regulated directly by the miRNA of interest.
10. Due to the fairly low toxicity of Lipofectamine 2000, it is not necessary to remove the transfection complexes after transfection; however, the media can be changed after 4–6 h if desired.
11. Freezing cells in lysis buffer at –80°C and thawing prior to clearing by centrifugation may improve lysis and protein recovery from some cell lines.
12. The PVDF membrane used in immunoblotting has a high affinity for proteins. Following transfer of proteins from the gel, it is essential to block the remaining membrane surface to prevent nonspecific binding of the detection antibodies during later steps. A variety of blocking buffers (e.g., skim milk, BSA) have been used to block free sites on a membrane, improving the specificity of the assay by reducing background and increasing the signal to noise ratio. In our hands, 5% skim milk in 1× TBS-T is a suitable blocking buffer for most applications.
13. Chemiluminescent blotting substrates produce a transient signal that persists only as long as the enzyme–substrate reaction is occurring. If the ECL substrate is consumed or the enzyme (HRP-conjugated to the secondary antibody) loses activity, then the reaction will end and the signal will be lost. However, in our experience, the signal should be stable for at least 30–45 min, allowing detection with X-ray film (ECL

Hyperfilm) or digital imaging equipment. Film exposure times are typically shorter than those needed with digital imaging equipment, as film is densely coated with photoreactive molecules and is placed in direct contact with the blot, allowing more light photons to collide with the film emulsion than is the case with a digital sensor located behind focusing lenses. As a result, film is more sensitive than digital imaging equipment.

14. As certain cell lines are more resistant to lysis, we generally recommend freezing cells in 1× passive lysis buffer at –80°C and thawing prior to clearing lysates by centrifugation.

Acknowledgments

This work was supported by the National Health and Medical Research Council of Australia and the Cancer Council of Western Australia.

References

1. Ritter, C. A., and Arteaga, C. L. (2003) The epidermal growth factor receptor-tyrosine kinase: a promising therapeutic target in solid tumors. *Semin Oncol* **30**, 3–11.
2. Menard, S., Pupa, S. M., Campiglio, M., and Tagliabue, E. (2003) Biologic and therapeutic role of HER2 in cancer. *Oncogene* **22**, 6570–8.
3. Tokunaga, E., Oki, E., Egashira, A., Sadanaga, N., Morita, M., Kakeji, Y., and Maehara, Y. (2008) Deregulation of the Akt pathway in human cancer. *Curr Cancer Drug Targets* **8**, 27–36.
4. Bulgaru, A. M., Mani, S., Goel, S., and Perez-Soler, R. (2003) Erlotinib (Tarceva): a promising drug targeting epidermal growth factor receptor tyrosine kinase. *Expert Rev Anticancer Ther* **3**, 269–79.
5. Ranson, M., and Sliwkowski, M. X. (2002) Perspectives on anti-HER monoclonal antibodies. *Oncology* **63, *Suppl*** 1, 17–24.
6. Bartel, D. P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–97.
7. Humphreys, D. T., Westman, B. J., Martin, D. I., and Preiss, T. (2005) MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc Natl Acad Sci U S A* **102**, 16961–6.
8. Chen, J. F., Mandel, E. M., Thomson, J. M., Wu, Q., Callis, T. E., Hammond, S. M., Conlon, F. L., and Wang, D. Z. (2006) The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat Genet* **38**, 228–33.
9. Cheng, A. M., Byrom, M. W., Shelton, J., and Ford, L. P. (2005) Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. *Nucleic Acids Res* **33**, 1290–7.
10. Zhang, B., Pan, X., Cobb, G. P., and Anderson, T. A. (2007) microRNAs as oncogenes and tumor suppressors. *Dev Biol* **302**, 1–12.
11. Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20.
12. Webster, R. J., Giles, K. M., Price, K. J., Zhang, P. M., Mattick, J. S., and Leedman, P. J. (2009) Regulation of epidermal growth factor receptor signaling in human cancer cells by microRNA-7. *J Biol Chem* **284**, 5731–41.
13. Epis, M. R., Giles, K. M., Barker, A., Kendrick, T. S., and Leedman, P. J. (2009) miR-331-3p regulates ERBB-2 expression and androgen receptor signaling in prostate cancer. *J Biol Chem* **284**, 24696–704.

14. Persad, S., Attwell, S., Gray, V., Mawji, N., Deng, J. T., Leung, D., Yan, J., Sanghera, J., Walsh, M. P., and Dedhar, S. (2001) Regulation of protein kinase B/Akt-serine 473 phosphorylation by integrin-linked kinase: critical roles for kinase activity and amino acids arginine 211 and serine 343. *J Biol Chem* **276**, 27462–9.
15. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–54.
16. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–5.
17. Towbin, H., Staehelin, T., and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350–4.
18. Grimson, A., Farh, K. K., Johnston, W. K., Garrett-Engele, P., Lim, L. P., and Bartel, D. P. (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* **27**, 91–105.

Chapter 12

Epigenetic Regulation of MicroRNA Expression in Cancer

Hani Choudhry and James W.F. Catto

Abstract

Epigenetic gene regulation is important in human cancer. Both functional and observational data implicate alterations of histone modifications, DNA promoter methylation, and non-coding RNA expression in carcinogenic roles. We sought to explore the role of aberrant DNA hypermethylation in the regulation of microRNA (miR) expression in human cancer. From human genome databases we calculated that 13 and 28% of human miR genes are located within 3 and 10 kb of a CpG island, respectively. To identify miRs that are regulated by epigenetic mechanisms in cancer, we performed expression profiling prior to and following treatment of cell lines with 5-azacytidine. We used oligonucleotide microarrays to determine miR expression. For miRs whose expression changed following 5-azacytidine treatment, we sequenced the adjacent CpG island and promoter using bisulphite-treated DNA. Here, we describe these methods to enable other researchers to use this approach.

Key words: Cancer, MicroRNA, Epigenetic, DNA methylation, Gene expression, Microarray

1. Introduction

MicroRNAs (miRs) are small non-coding RNA species that post-transcriptionally regulate gene expression. Altered miR expression has been observed in many cancers and is the focus of this book. MiRs with carcinogenic gain of functions and tumour suppressor loss of function have been observed. For example, we recently characterised loss of miRs-99a/100 expression in low-grade bladder cancer (1). Down-regulation of these two miRs leads to upregulation of FGFR3, one of their most conserved target genes. Cancer associated loss of miR expression may occur either from chromosomal deletion, gene mutations, translocation or by epigenetic mechanisms (2, 3).

Epigenetic regulation of miR expression has been reported in various cancers (4, 5). We analysed the latest draft of the human

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_12, © Springer Science+Business Media, LLC 2011

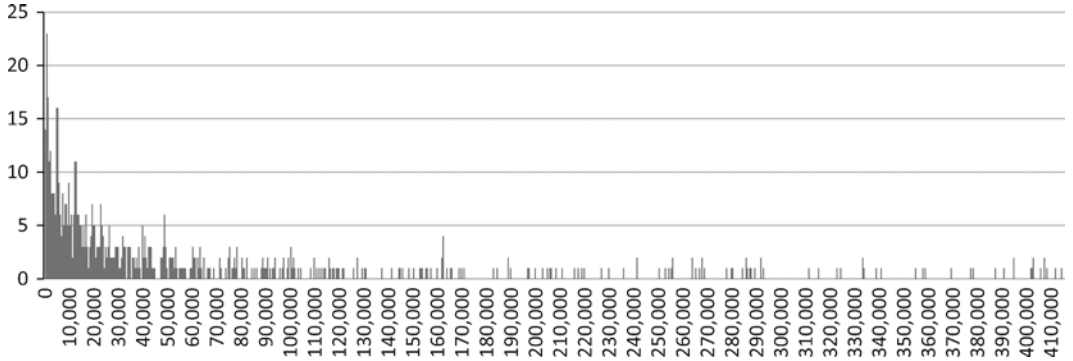


Fig. 1. Distance of human miRNA genes to the nearest CpG island in bp. Data obtained from UCSC genome browser and miRBase (Choudhry and Catto, unpublished data).

genome and miRBase and found that 13 and 28% of human miR genes are located within 3 and 10 kb of a CpG island, respectively (Fig. 1) (Choudry and Catto, unpublished data). Thus, 199 miRs may be susceptible to silencing by aberrant DNA hypermethylation. As each miR targets many genes, this could have a profound effect upon cell behaviour. To establish epigenetic miR regulation various methods can be used. The simplest method is probably to sequence the adjacent CpG islands for miRs that are of interest (6). This is accurate but laborious for high throughput. Here, we describe the chemical re-expression of epigenetically regulated miRs by inhibition of DNA methyltransferase. Thus, high throughput is possible using oligonucleotide expression microarrays. For confirmation, the adjacent CpG islands from miRs with altered expression are sequenced. Functional studies would then be required to confirm the importance of these observational findings (7).

2. Materials

2.1. Cell Culture

1. Human bladder cancer cell lines EJ, RT112, RT4 were obtained from the American Type Culture Collection Manassas, Virginia.
2. Cancer cell lines were cultured in Dulbecco's modified Eagle medium (DMEM) with 10% foetal calf serum (FCS). For long-term use, medium is stored at 4°C and warmed prior to every use.
3. Normal non-immortalised human urothelial cells were expanded from fresh urothelial biopsies using standard methods (8).

4. Complete keratinocyte serum-free medium (KSFMc) (GIBCO invitrogen) was used for normal human urothelial (NHU) cell culture. The medium is available in 500 ml bottle and supplied with 25 mg bovine pituitary extract (BPE), 2.5 μg recombinant epidermal growth factor (rEGF) and 1 mg of cholera toxin.
5. KSFMc medium is prepared by adding rEGF (5 mg/ml) and BPE (50 ng/ml) to 500 ml bottle of KSFMc. Cholera toxin (500 μl, 30 μg/ml) was also added to the medium and was kept at 4°C.
6. Trypsin inhibitor (GIBCO invitrogen) was prepared by adding 205 mg (1 vial) to 5 ml of PBS buffer (pH 7.4). The solution was mixed and filtered through a 0.2 μm sterile filter. Aliquots of 50 μl were stored at −20°C. For each use, one aliquot of trypsin inhibitor was thawed and added to 5 ml of KSFMc to inhibit 1 ml of trypsin–EDTA solution.

2.2. miRNA and Total RNA Extraction

1. The *mir*Vana™ RNA isolation kit (Ambion, Texas) is used according to the manufacturer's instructions for miRNA and total RNA extraction.
2. The kit contains miRNA washing solution 1 (WS1), washing solution 2/3, lysis, binding buffers, miRNA homogenate additive, acid-phenol:chloroform, gel loading buffer II, collection tubes, and elution solution.
3. The kit is stored at 4°C and warmed to room temperature before use.
4. Washing solutions are diluted with absolute ethanol; 21 ml of ethanol is added in WS1, and 40 ml of ethanol is added in washing solution 2/3.

2.3. miRNA Microarray

1. The human miRNA microarray (Agilent Technologies, California) manufactured using Agilent 60-mer SurePrint technology.
2. Each glass slide contained eight individual microarrays consisting of probes to 726 human and 76 human viral miRNAs annotated from the Sanger miRBase (version 10.1).
3. Microarrays were stored in the dark at room temperature.

2.3.1. Complementary Equipments and Kit

1. miRNA microarray hybridisation gasket slides (Agilent Technologies, California) are used to load labelled samples for hybridisation, which is performed in a stainless hybridisation chamber.
2. Glass slide-staining dishes are used for microarray slide washing. The glass dish is washed thoroughly with dH_2O and rinsed twice with gene expression washing buffer 1 before use.

Table 1
Microarray sample labelling and hybridisation kit's components and their storing temperature

Items	Stored
10× Calf intestinal phophatase buffer	–20°C
10× GE blocking agent	Solution at –20°C lyophilised 25°C
10× T4 RNA ligase buffer	–20°C
2× Hi-RPM hybridisation buffer	25°C
Calf intestinal alkaline phosphatase (CIP)	–20°C
Cyanine 3-pCp	–20°C
Dimethyl sulphoxide (DMSO)	–20°C
Nuclease-free water	4°C
T4 RNA ligase	–20°C

3. Agilent sample labelling and hybridisation kit (Agilent) contains all the required components for sample labelling and hybridisation preparation (Table 1).
4. 10× GE blocking agent is supplied lyophilised with the sample labelling and hybridisation kit. Nuclease-free water (125 µl) is added to 10× GE blocking agent vial. The solution is mixed gently by vortexing and heated for 5 min at 37°C.

2.3.2. Gene Expression Wash Buffers

1. Gene expression wash buffers 1 and 2 (Agilent) are prepared by adding 2 ml of 10% of TritonX-102 is added into both gene expression wash buffers. Buffers are mixed thoroughly by inverting the container six times.
2. Gene expression washing buffer 2 (1,000 ml) is dispensed into 1.5 l sterile glass bottle. The bottle is placed in a 37°C water bath and stored overnight before washing arrays. The gene expression washing buffer 2 is kept warmed at 37°C during microarray washing.

2.3.3. TaqMan miR Assays for Real Time-QPCR

1. TaqMan miR assays (Applied Biosystems, California) consist of short hairpin primers for reverse transcription and a minor groove binder (MGB) probe for PCR quantification. TaqMan miR assays are stored at –20°C.
2. For reference genes, we use the median expression value of two endogenous small nuclear RNAs: RNU44 and 48.

3. High-capacity cDNA reverse transcription kit (Applied Biosystems) stored at −20°C, contains 10× RT buffer, 25× dNTP mix (100 mM), reverse transcriptase (50 U/μl), and RNase inhibitor.
4. TaqMan 2× Universal PCR Master Mix (Applied Biosystems) contains AmpliTaq Gold DNA Polymerase, uracil *N*-glycosylase (UNG), dNTPs, dUTP, passive reference (ROX), and optimized buffer components. The mix is stored at 4°C.

2.4. DNA Extraction

1. The gold-standard method to evaluate the methyl status of Cytosine residues is bisulphite sequencing. For this, DNA is treated with sodium bisulphite to induce a bisulphite and methyl-specific sequence change and then sequenced using primers specific to the bisulphite sequence, not the methyl sequence.
2. DNeasy Blood and Tissue kit (Qiagen, Texas) is used according to the manufacturer's instructions for cell lines. The kit contains DNeasy mini spin columns in 2-ml collection tubes, buffer ATL, buffer AL, buffer AE, proteinase K, buffer AW1, and buffer AW2. The kit was stored at room temperature.
3. Buffer AW1 and AW2 are diluted with ethanol (100%); 25 ml and 30 ml of ethanol are added to buffer AW1 and AW2, respectively.

2.4.1. Bisulphite Modification of DNA

1. CpGenome DNA modification kit is used for bisulphite conversion of DNA to evaluate the Cytosine methylation status (Chemicon International, California).
2. The kit contains DNA modification reagents I, II, III, and IV. The entire set of reagents are kept at −20°C and warmed to room temperature before use.
3. Sodium hydroxide (NaOH) (3 M) is freshly prepared by dissolving 1 g in 8.3 ml dH_2O.
4. In a 1.5 ml Eppendorf tube, 900 μl ethanol (100%) is added to 6.6 μl of 3 M NaOH and 93.4 μl of dH_2O is added. This achieves a final concentration of 20 mM NaOH/90% ethanol. The solution is freshly prepared prior to each use.
5. In a universal tube, 1 μl of β-mercaptoethanol is dissolved in 20 ml of dH_2O.
6. Reagent I is prepared by dissolving 0.227 g of DNA Modification Reagent I in 0.571 ml of nuclease-free water with 60 μl of 3 M NaOH (to adjust the pH to 5.0) per sample (see Note 1).
7. Reagent II is prepared by adding 750 μl of β-mercaptoethanol solution (prepared above) to 1.35 g of DNA Modification reagent II for each sample to be modified. The solution is mixed well and wrapped with foil (see Note 1).

2.4.2. Bisulphite-Specific Polymerase Chain Reaction

1. AccuPrime *Taq* DNA polymerase system (Invitrogen, Paisley, UK) contains AccuPrime *Taq* DNA polymerase, 10× AccuPrime PCR buffer I, II and 50 mM magnesium chloride ($MgCl_2$). The kit is stored at –20°C.
2. The 10× AccuPrime PCR buffers contain 20 mM Tris–hydrochloric acid (Tris–HCl) (pH 8.0), 500 mM potassium chloride (KCl), 15 mM $MgCl_2$, 2mM deoxythymidine triphosphate (dTTP), 2 mM deoxyadenosine triphosphate (dATP), 2 mM deoxythymidine triphosphate (dTTP), 2 mM deoxycytidine triphosphate (dCTP), 10% glycerol and thermostable AccuPrime protein.
3. Bisulphite specific primers are purchased (Sigma) and reconstituted to a concentration of 100 μM.
4. For PCR reactions, primers are diluted to a working stock at 10 μM.

2.4.3. Gel Electrophoresis

1. Concentrated (50×) Tris–acetate–EDTA (TAE) buffer is made by adding 121 g Tris base, 9.3 g EDTA and 28.55 ml of glacial acetic acid. Distilled H_2O (dH_2O) is added to make the volume up to 500 ml, and pH is adjusted to 8.3.
2. Ethidium bromide (10 mg/ml) is prepared by dissolving 1 g of ethidium bromide in 100 ml of dH_2O. The solution is wrapped in aluminium foil and stored at 4°C.
3. DNA hyperladder I, II, V (Invitrogen, Paisley, UK) is used as DNA marker (1 μg/μl).

2.4.4. PCR Purification Prior to Sequencing

1. MinElute PCR Purification Kit (Qiagen, Texas) contains MinElute Spin Columns, collection tubes, pH indicator, and buffers (PB, PE, and EB). The kit was stored at room temperature.
2. The PE buffer is diluted by adding 24 ml of 100% ethanol.

3. Methods

Epigenetic gene silencing is associated with aberrant hypermethylation of CpG dinucleotides within the gene promoters (9). This hypermethylation is an active process performed by one of several DNA methyltransferases. Thus, inhibition of these DNA methyltransferases leads to the reversal of aberrant hypermethylation including 5-azacytidine (8). Experiments that detect this re-expression will, therefore, be able to identify genes whose methyl status is reversed by this experiment (Fig. 2). Various experimental designs are possible within this context, but the easiest, most available, and highest throughput is probably the expression microarray. As with typical microarray experiments, one must pay

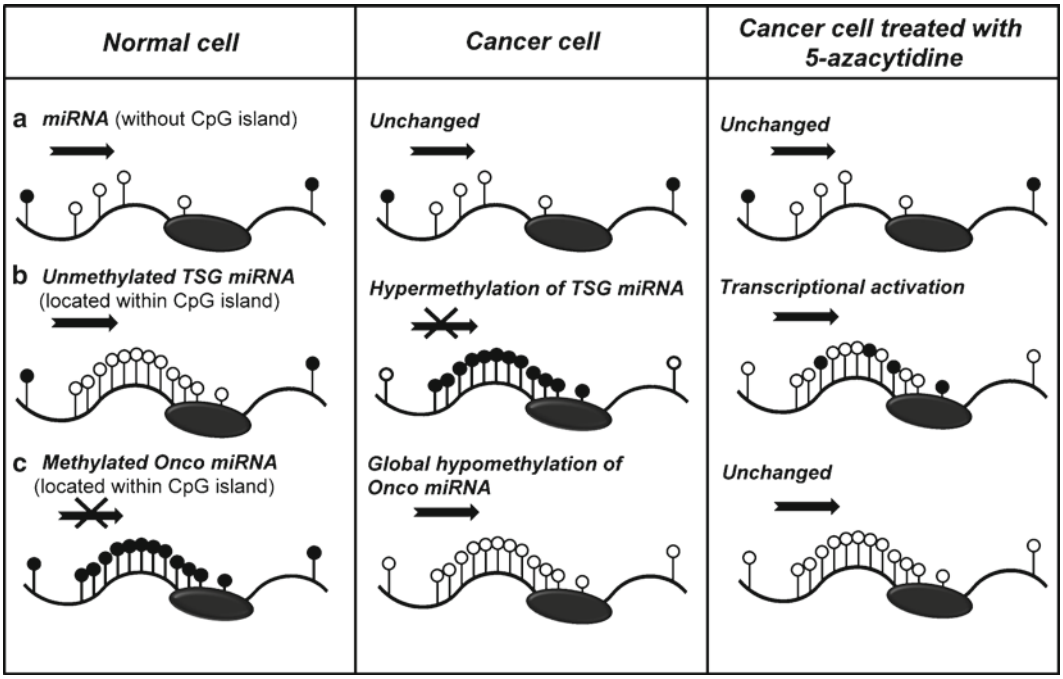


Fig. 2. *DNA methylation patterns of miRNA genes in normal and cancer cells.* (**a**) miRNA genes which are not located close to a CpG islands are unaffected by CpG methylation. (**b**) Tumour suppressor miRNA (TSG miRNA) genes located in a CpG island are unmethylated in normal cells but undergo hypermethylation in cancer cells (silenced). Demethylating agents can reactivate their expression. (**c**) Onco-miRNA genes located in CpG islands are hypermethylated in normal cells (silenced) but hypomethylated in cancer cells (active). They are less affected by demethylating agents.

attention to experimental reproducibility (and therefore use biological and experimental replicates) and validate the results by another method. Here, we use bisulphite sequencing of individual gene promoter regions to validate our microarray findings.

3.1. Cell Culture and 5-Azacytidine Treatment

3.1.1. Control and Treatment Cells Culture

1. Divide cell lines into: control and treatment groups.
2. Each contains triplicate repeats.
3. Cancer cell lines and NHU cells are cultured in sterile conditions in a microbiological grade-2 hood (see Note 2).
4. Warm all media and cell culture reagents to 37°C prior to use.
5. Cancer cell lines are cultured in Nunc T75 culture flasks with DMEM (see Table 2). NHU cells are cultured in Primaria culture flasks using KSFMc.

3.1.2. Preparation of DNA Methyltransferase Inhibition "5-Azacytidine"

1. 5-Azacytidine (Sigma) is supplied as 1.2 mg lyophilised powder at 50-fold concentration. Add 10 ml sterile DMEM to reconstitute and reach a final concentration of 500 μM (see Note 3).
2. The solution is stored at −20°C in 500 μl aliquots.

Table 2
Description of bladder cancer cell line and their culture conditions

Cell line	Differentiation	Growth	Description	Culture medium	Cell passaging
RT4	Well (grade 1)	Grow as monolayer	Well-differentiated transitional papillary urothelial carcinoma and slow-growing tumour	DMEM medium with 10% FCS	1:4 Dilution using 2 ml of (1:1) trypsin/versene mixture
RT112	Moderate (grade 2)	Epithelial-like cells	Fast-growing tumour	DMEM medium with 10% FCS	1:5 Dilution using 2 ml of (1:1) trypsin/versene mixture
EJ	Poor (grade 3)	Mesenchymal-like cells growing monolayer	Fast-growing tumour	DMEM medium with 10% FCS	1:8 Dilution using 2 ml of (1:1) trypsin/versene mixture

3.1.3. Cell Line Treatment

1. The cancer cell lines are treated for 5 days.
2. Day 0: seed cells at a low density. This density depends upon the cell growth rate. Aims for a 70% confluence by day 5 with treatment.
3. Days 1, 3 and 5:
   - *Control group*: old media changed for fresh DMEM (12.55 ml).
   - *Treatment group*: old media is changed for fresh DMEM (12.5 ml) with 50 μl of 5-azacytidine added to each flask (final concentration of 2 μM).

3.1.4. Cell Harvest

1. Cells are grown in monolayers and harvested on day 6.
2. Discard media and wash cells twice with 10 ml of PBS. To release adherent cells, add 2 ml of trypsin–EDTA (Sigma) and incubate for 10 min at 37°C.
3. Add 4 ml media to inactivate trypsin and transfer the cells to universal tubes.
4. Centrifuge cells at 400×*g* for 5 min, carefully aspirate and discard the media, before re-suspending in 1 ml of PBS.
5. Split the cells/PBS into two 1.5 ml Eppendorf tubes (each will contain 0.5 ml of cell solution) and obtain approximate cells counts using a haemocytometer chamber (Sigma).
6. Centrifuge Eppendorf tubes at 400×*g* for 5 min and discard the PBS to obtain a cell pellet for immediate use or storage at –80°C in liquid nitrogen.
7. To confirm successful 5-azacytdine treatment, we perform rtPCR using extracted RNA (from Subheading 3.2) to look for re-expression of a gene known to be methylated in your cell line of choice. For example, we use *EDNRB* (10) in our bladder cancer cell lines (Fig. 3).

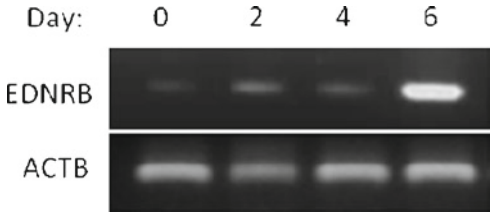


Fig. 3. *Expression of EDNRB RNA in the EJ cell line with 5-azacytidine treatment.* Re-expression of this tumour suppressor gene occurs with inhibition of DNA methyltransferase. β-*Actin* RNA shown as control of the RT-PCR reaction.

3.2. RNA Extraction and miR Quantification

3.2.1. RNA Extraction

1. Several protocols for RNA extraction are described. For miR extraction, it is important to use a method without a column (such as Trizol method) or to use a high ethanol concentration to bind small RNA species to the silica membrane within a column. Otherwise, small species will be washed through the column without retrieval. We use the mirVana RNA isolation kit (Ambion, Texas) according to the manufacturer's instructions.
2. Clean the bench and equipment with an RNase decontamination solution.
3. Use fresh or thawed frozen cell pellets and re-suspended in 200 µl PBS to an appropriate number of cells (e.g. total 1×10^7).
4. Add lysis/binding solution (600 µl) and pulse-vortex to homogenise lysate.
5. Add 60 µl of homogenate additive to the lysate and leave on ice for 10 min.
6. Add 600 µl acid-phenol:chloroform and pulse-vortex for 60 s. Centrifuge the sample for 5 min at 10,000 × *g*.
7. Remove the upper aqueous phase and add 100% ethanol (375 µl). Load this mixture into the filter column provided in the kit. Up to 700 µl can be applied at a time.
8. Centrifuged at 10,000 × *g* for 15 s to bind the RNA to the membrane.
9. The flow-through contains small RNA and is used for miR extraction (use for Subheading 3.2.2).
10. Place the filter column into a new collection tube and apply 700 µl of WS1. Centrifuge for 10 s at 10,000 × *g*. Discard the flow-through and change the collection tube.
11. Apply 500 µl of washing solution 2/3 and centrifuge at 10,000 × *g* for 10 s.
12. Repeat this step twice, before centrifuging the column without solution for 1 min, to dry the membrane (see Note 4).
13. Place the column into a new collection tube and add 100 µl of pre-heated (95°C) elution solution to the centre of the filter.
14. Centrifuge the column and collection tube at 13,000 × *g* for 30 s to recover the RNA. Store at –80°C until used.

3.2.2. miR Extraction

1. The flow-through from the first RNA extraction step (see step 9 in Subheading 3.2.1) is used for miR extraction. Determine the approximate volume of this flow-through and add two-thirds volume of 100% ethanol, before mixing.

2. Pipette this mixture in 700 µl aliquots into a new filter column and centrifuge for 15 s at 10,000×*g*. Pass all the mixture through the column.
3. Apply WS1 (700 µl) to the filter column and centrifuge for 10 s at 10,000×*g*. Discard the flow-through and change the collection tube.
4. Apply washing solution 2/3 (500 µl) to the filter column and centrifuge for 10 s at 10,000×*g*. Discard the flow-through and change the collection tube.
5. Repeat this step twice, before centrifuging the column without solution for 1 min, to dry the membrane (see Note 4). Then, the cartridge is transferred to a fresh collection tube (provided in kit).
6. Apply 100 µl pre-heated (95°C) elution solution to the centre of the filter.
7. Tubes are centrifuged for 30 s at maximum speed to recover the miRNA and stored at –20°C until use.

3.3. miR Expression Microarray

1. Single-colour Agilent human miRNA microarrays are used to investigate miRNA expression.
2. Microarray processing consists of five main steps: (a) *sample preparation*, (b) *preparation of labelling reaction*, (c) *sample hybridisation*, (d) *microarray slide washing*, *and* (e) *microarray scanning*. All the steps were performed according to miRNA microarray system's protocol with slight changes to achieve best array outcome (Agilent Technologies, California).

3.3.1. Sample Preparation

1. Total RNA concentration and quality should be first determined (e.g. using a nano-drop spectrophotometer).
2. Total RNA from each cell lines is diluted to 50 ng/µl in nuclease-free water for each individual microarray (see Note 5).

3.3.2. Preparation of Labelling Reaction

1. Diluted total RNA 2 µl (100 ng) is added to a 1.5 ml Eppendorf tube and kept on ice (see Note 5).
2. Immediately prior to use, the calf intestinal alkaline phosphatase (CIP) master mix is prepared by combining 10× CIP buffer (3.6 µl), nuclease-free water (9.9 µl), and CIP (4.5 µl). These volumes are per 9 µl reaction (see Note 6).
3. Add 2 µl CIP master mix to each RNA sample for a total reaction volume of 4 µl, and mix by gently pipetting.
4. De-phosphorylate the samples by incubating at 37°C for 30 min (see Note 7).
5. Add 2.8 µl of 100% dimethyl sulphoxide (DMSO) to each sample and incubate at 100°C for 7 min. *Caution*: samples

should not be incubated less than 5 min and no more than 10 min, as this leads to inefficient sample labelling.

6. Transfer samples immediately to ice-water (see Note 8).
7. Immediately prior to use, the ligation master mix is prepared by gently mixing of 10× T4 RNA ligase buffer (9 µl), Cyanine1-pCP (27 µl) and T4 RNA ligase (4.5 µl). These volumes are per 9 µl reaction (see Note 9). *Caution*: the ligation master mix should be used within 15 min of mixing all components. Delay to do so may affect the labelling efficiency.
8. Immediately, 4.5 µl of the ligation master mix is added to each sample tube for a total reaction volume of 11.3 µl
9. Tubes are gently mixed by pipetting and gently spin down.
10. The reactions are then transferred to PCR tubes and incubated at 16°C in the thermal cycler (Applied Biosystems 2700) for 2 h.
11. Samples are then transferred to 1.5-ml Eppendorf tubes and dried up using a vacuum concentrator (left for 4 h in vacuum concentrator).
12. To confirm dehydration, tubes are flicked hard on the bottom and check that the pellets do not move or spread (see Note 10).

3.3.3. Sample Hybridisation

1. Nuclease-free water (18 µl) is added to resuspend the dried samples.
2. GE blocking agent (10×) and Hi-RPM hybridisation buffer (2×) are added to the samples, 4.5 and 22.5 µl, respectively. The mixture is gently mixed by vortexing.
3. Incubated at 100°C for 5 min and then immediately transfer to ice water for 5 min. *Caution*: the mixture must not be left in the ice water bath for more than 15 min. Longer incubations may adversely affect the hybridisation results.
4. To prepare the hybridisation assembly, a clean gasket slide is loaded on into Agilent SureHyb chamber with the label is facing up and aligned with the chamber base.
5. Slowly entire volume of the hybridisation sample is loaded into the gasket well (see Note 11).
6. Carefully, the "active side" of the array is placed onto the gasket slide so that the "Agilent-labelled barcode" is facing down and "numeric barcode" is facing up.
7. The SureHyb chamber cover is placed on top as this tightened the clamp onto the chamber.
8. The assembled chamber is vertically rotated to wet the gasket and assess the mobility of the bubbles.

9. The assembled slide chamber is placed in a hybridisation oven at 55°C, rotating at 20 revolutions per minute (RPM) for 20 h (see Notes 12).

3.3.4. Microarray Slides Washing and Scanning

1. The microarray wash procedure for Agilent's miRNA platform is done in environments where ozone levels are 50 ppb or less.
2. The gene expression wash buffer 2 is warmed overnight (as explained above)
3. Slide-staining dish #1 is filled with gene expression wash buffer 1 at room temperature.
4. Place a slide rack into slide-staining dish #2 and filled with enough gene expression wash buffer 1 at room temperature to cover the slide rack. Stir with a magnetic bar stirrer.
5. Place a slide rack into slide-staining dish #3 and fill to approximately three-fourths full with warmed gene expression wash buffer 2 (at 37°C). This dish is stirred and warmed to 37°C by a heated magnetic bar stirrer.
6. The hybridisation chamber assembly is removed from the hybridisation oven after incubation and placed on a flat surface.
7. The thumbscrew of chamber is loosened and the chamber cover removed.
8. With gloved fingers, the array-gasket sandwich is removed from the chamber base by grabbing the slides from their ends and quickly submerges into slide-staining dish #1 containing Gene Expression Wash Buffer 1.
9. Once array-gasket sandwich is completely submerged in Gene Expression Wash Buffer 1, gently open from the barcode end only.
10. The gasket slide is dropped to the bottom of the staining dish, and the microarray slide is transferred into the slide rack in slide-staining dish #2 containing Gene Expression Wash Buffer 1 at room temperature. *Caution*: the microarray slide should be transferred by touching only the barcode portion of the microarray slide or its edges.
11. Place the microarray into the slide rack in slide-staining dish #2 and wash at moderate stir speed for 5 min.
12. Transfer the microarray into the slide rack in slide-staining dish #3 and wash in the warmed Gene Expression Wash Buffer 2 (at 37°C) for 5 min (see Note 13).
13. The slide rack containing the microarray is removed slowly to minimising droplets on the slides. It takes about 5–10 s to remove the slide rack.

14. After washing, slides are immediately scanned to minimise the impact of environmental oxidants on signal intensities.
15. For scanning, slides are placed into the slide holder such that the numeric barcode side is visible.
16. The assembled slide holder is placed into a scanner carousel.
17. An Agilent DNA microarray scanner with high resolution technology© from Agilent Technologies Microarray core facility, University of Sheffield) was used to scan the slides.
18. The scanner is set according to the miRNA microarray recommended scanning protocol (*scan sitting*: values for 8×15 k formats, *scan region*: scan area (61×21.6 mm), *sac resolution* (μm): 5, *5 µm scanning mode*: single press, *extended dynamic range* (*XDR*): selected, *dye channel*: green, *green photo-multiplier tube* (*PMT*): XDR high 100%, XDR low 5%).
19. After scanning, slides are stored in orange slide boxes given with the kit, covered with aluminium foil and stored at room temperature.

3.3.5. Data Extraction and Data Analysis

1. A grid template (provided with kit in CD) is uploaded on the computer. The Grid templates contain all the information about probes on the microarray.
2. The Agilent Feature Extraction software 10.5.1.1 is used to extract miRNA microarray data. Feature extraction software processes array information from probe features and provide measurement of miRNA expression in experiments and quality control.
3. Feature extraction is performed on all microarrays, and the extracted data is further analysed using Microsoft excel and Genespring GX.

3.3.6. Real Time Quantitative PCR (RT-qPCR) Validation

1. To confirm altered miR expression (detected by microarray), we recommend individual real time RT-PCR analysis of RNA from the control and treatment cell lines.
2. Currently, individual miR assays are available from Applied Biosystems and Ambion. Each includes a hairpin RT-PCR primer and a MGB probe for quantitative.
3. TaqMan miR assay consists of two steps:
   - *First*: RNA sample is reverse transcripted (RT-PCR) to form a complementary DNA (cDNA) using specific miRNA stem-looped primers (provided in kit).
   - *Second*: cDNA is amplified using specific miRNA real-time PCR primers together with probe and TaqMan Universal PCR master mix.
4. For miRNA reverse transcription (RT-PCR), master mix is prepared by adding 100 mM dNTPs (0.15 µl), reverse transcriptase

50 U/μl (1 μl), 10× reverse transcription buffer (1.5 μl), RNase inhibitor 20 U/μl (0.19 μl), nuclease-free water (4.16 μl), and miRNA stem-looped RT-primers (3 μl).

5. The RT-PCR protocol is 30 min at 16°C, followed by 30 min 42°C and 5 min 82°C.
6. For Real Time, the PCR master mix reaction is prepared by mixing TaqMan 2× Universal PCR master mix (10 μl), nuclease-free water (7.67 μl), miRNA real time primers (1 μl), and cDNA (1.33 μl).
7. The protocol for qPCR: initial temperature for enzyme activation (10 min at 95°C), denature (15 s at 95°C), annealing and extension (60 s at 60°C), PCR for 50 cycles.
8. The PCR reaction mixture (20 μl) was loaded onto ABI 384-well clear optical reaction plate. The plate was sealed with optical adhesive cover and briefly centrifuged to spin down the contents and eliminate air bubbles.
9. The Applied Biosystems 79000HT system is used for real time-PCR.
10. The data was analysed using the Sequence Detection System (SDS) software.2.21.
11. We compare expression of the miR of interest with two endogenous snRNA (RNU44 and RNU48).

3.4. DNA Extraction

1. There are many reliable methods for DNA extraction. We currently use the DNeasy Blood and Tissue kit (Qiagen, Texas).
2. Use fresh or thawed frozen cell pellets (around 5×10^6 cells) and re-suspend in 200 ml PBS. Add proteinase K (20 ml) to lyse cells.
3. Add Buffer AL (200 ml) to aid nucleic acid binding to the silica membrane within the column. Mix immediately by pulse-vortexing.
4. Incubate tubes at 56°C for 10 min and then add 200 ml 100% ethanol.
5. Transfer the solution to a DNeasy mini spin column, within a 2-ml collection tube, and centrifuge for 1 min at 6,500×g. DNA is selectively bound to silica-based membrane. Discard the flow-through and replace the 2-ml column collection tube.
6. Add buffer AW1 (500 ml) to the column and centrifuged for 3 min at 6,500×g.
7. Discard the flow-through and collection tubes. Place the DNeasy mini columns within a new 2 ml collection tube.
8. Add Buffer AW2 (500 ml) to the DNeasy mini columns and centrifuge for 3 min at 20,000×g to dry the DNeasy membrane. Discard the collection tube and flow-through.

9. For DNA elution, 200 ml of buffer AE is added directly onto the DNeasy membrane and incubated at room temperature for 1 min. The tubes are centrifuged for 1 min at 6,500 × g to retrieve the eluted DNA.
10. Store DNA at –20 to –80°C for subsequent use (depending upon storage time).

3.4.1. Bisulphite Modification of DNA

1. There are numerous bisulphite treatment protocols and kits. We currently use the CpGenome DNA modification kit (Millipore, California).
2. In a 1.5 ml Eppendorf tube, the genomic DNA (1 μg) is added to 7 μl of 3 M NaOH in 98 μl of nuclease-free water.
3. The DNA Modification reagent IV (2 μl) is added to the solution and mixed well.
4. Incubate the tubes for 10 min in heat block at 50°C.
5. Freshly prepared DNA modification reagent I (550 μl) is added to tubes and vortexed.
6. The tubes are incubated in heat block at 50°C for 16 h "overnight" in the dark.
7. Resuspend DNA modification III agent by vortexing vigorously and add 5 μl to the DNA solution.
8. Freshly prepared DNA modification Reagent II (750 μl) is added and mixed briefly (the mixture should turn into yellow).
9. The tubes are incubated at room temperature for 10 min and centrifuged for 1 min at 5,000 × *g* (A small white pellet should be present).
10. The supernatant is discarded from the pellet.
11. To the pellet, 1 ml of 70% ethanol is added, vortexed, spin at 5,000 × *g* for 1 min, and supernatant discarded. Perform this step three times.
12. After the supernatant from the third wash has been removed, the tubes are centrifuged at high speed for 2 min. The remaining supernatant is removed with a plastic pipette tip.
13. Freshly prepared 20 mM NaOH/90% ethanol solution (50 μl) is added the tubes and vortexed.
14. The tubes are incubated at room temperature for 5 min.
15. 90% of ethanol (1 ml) is added to tube, pulse-vortexed and then centrifuge at 5,000 × *g* for 1 min. Carefully discard the supernatant. This step is performed twice.
16. After the supernatant from the second wash has been removed, the samples are centrifuged at high speed for 3 min. The remaining supernatant is removed with a plastic pipette tip.
17. Tubes are allowed to dry up for 30 min at room temperature (alcohol odour should be diminished).

18. The pellet is dissolved by adding 30 µl of Tris–EDTA buffer (pH.8) and vortexed well.
19. The tubes are incubated in heat block at 60°C for 15 min and then centrifuged at 13,000 × *g* for 3 min.
20. The supernatant (contains modified DNA) is transferred to a clean tube. For short term use the modified DNA is stored at –25°C for up to 2 months, while for long-term use, the modified DNA is stored at –80°C for up to 6 months (see Note 14). The DNA should be stored in aluminium foil to protect it from the light.
21. The modified DNA can now be sequenced to look at individual cell lines or used for Methyl specific PCR (MSP) or Quantified MSP (9, 11) if a cohort of samples needs analysing.
22. Bisulphite specific primers can be made from genomic sequences using a variety of software tools such as MethPrimer (see Subheading 3.4.2, step 3).

3.4.2. Genomic Sequencing of Bisulphite Modified DNA

1. Bisulphite sequencing is used to investigate the individual methylation status of CpG islands in cells or samples. In this case one can look at miRs that have changed expression following 5-azacytdine treatment.
2. Obtain the genomic sequence of interest (e.g. CpG Island adjacent to miR of note) from a genome repository such as Ensembl or The University of California Santa Cruz (UCSC).
3. Export the genomic region (as a FASTA file) into software that can generate primers for bisulphite-treated DNA such as MethPrimer (http://www.urogene.org/methprimer/index1.html).
4. Optimise the generated primers using universally methylated DNA (Invitrogen, Paisley, UK). For Bisulphite-specific PCR, we use the Accuprime Taq polymerase (Invitrogen, Paisley, UK).
5. The bisulphite-specific PCR master mix (20 µl) is prepared in clean sterile environment in a PCR setup room by adding Accuprime Taq DNA polymerase (0.5 µl), 10× AccuPrime PCR buffer I (2 µl), bisulphite-specific primers (10 pm/µl) (4 µl), nuclease-free water (9.5 µl), and bisulphite-modified DNA (50 ng/µl) (4 µl).
6. The bisulphite specific PCR protocol we use is: denaturisation (3 min at 95°C), denature (30 min at 95°C), annealing (40 s at Tm for ACTB is 56°C and for EDNRB is 57°C), extension (30 s at 68°C), and final extension (5 min at 68°C). Forty cycles are required.
7. Run 4 µl of the PCR product on an agarose gel using electrophoresis to confirm the success of the amplification and the absence of non-specific products.

8. The remaining PCR product is purified using QIAquick PCR purification kit (Qiagen, Texas). This step is performed to purify and clean the PCR reaction from contaminants (enzymes, unincorporated nucleotides) before sequencing.
9. Five volumes of Buffer PB are added to one volume of PCR sample and mixed well.
10. QIAquick spin column is placed in a provided 2-ml collection tube.
11. The sample is applied into the QIAquick column and centrifuge for 1 min at 17,900 × *g*. Flow-through is discarded, and the QIAquick column is placed back into the same tube.
12. For washing, 0.75 ml of buffer PE is added to the QIAquick column and centrifuge for 1 min at 17,900 × *g*.
13. Flow-through is discarded, and the QIAquick column is placed back in the same tube. An additional centrifuge step is performed for 1 min.
14. QIAquick column is transferred into a clean 1.5-ml microcentrifuge tube.
15. To elute DNA, 15 µl of Buffer EB (10 mM Tris–Cl, pH 8.5) is applied to the centre of the QIAquick membrane and centrifuge the column for 1 min at 17,900 × *g*.
16. To confirm the presence of clean DNA product, 5 µl of purified DNA sample is run on agarose gel electrophoresis before pursuing the sequencing step.
17. The purified DNA samples (10 µl) are sequenced using appropriate technology. For example we use BigDye terminator sequencing (Applied Biosystems 3730 DNA analyser).
18. The data is analysed using a sequence analyser such as Sequencher software (Gene Codes Corporation, http://www.genecodes.com/) (Fig. 4).

4. Notes

1. Reagents I and II are freshly prepared on day of experiment.
2. Equipment required for cell culture are wiped down with 70% ethanol before transfer into the hood (including media bottles, reagents, flasks, pipettes, and waste container)
3. The 5-azacytine is light sensitive, thus, it should be wrapped in aluminium foil or kept in dark tube.
4. It is important to dry mini spin column membrane, since residual ethanol may interfere with subsequent reaction. To

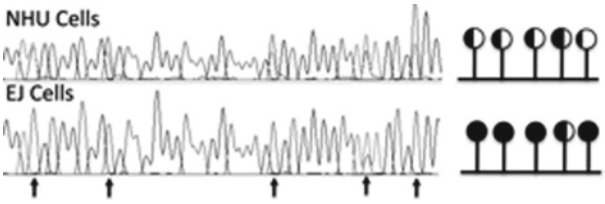


Fig. 4. *Aberrant methylation of CpG dinucleotides in EJ and NHU cells*. Bisulphite sequencing reveals the conversion unmethylated cytosine to thymidine bases and the maintenance of methylated cytosine residues. For EJ there is almost complete protection of the 5 methylated cytosine residues. In the NHU cells the 5 residues are hemimethylated, as shown by the presence of two alleles. The lollipops demonstrate the proportion of methylation (black) for each of the five cytosines.

ensure that no residual ethanol is carried over, and addition centrifugation is carried out.

5. The step should be performed on a nuclease free and 70% ethanol clean bench.
6. All the components of the CIP master mix should be added in the same order as indicated above and should be maintained on ice during preparation.
7. The samples can be stored at –80°C after incubation.
8. An ice-water bath should be used, instead of just crushed ice, to ensure that the samples remain properly denatured.
9. Prior use of 10× T4 RNA ligase buffer is warmed at 37°C and vortexed until all precipitate is dissolved. Before adding it into the master mix, the 10× T4 RNA ligase buffer should be cooled to room temperature. Failure to do so will affect the T4 RNA ligase activity and thus the labelling efficiency.
10. The sample must be completely dried after labelling. Residual DMSO will adversely affect the hybridisation results.
11. The pipette tip or the hybridisation solution must not touch the gasket walls. Allowing liquid to touch the gasket wall greatly increases the likelihood of gasket leakage.
12. The recommended hybridised time is 20 h and should be consistent. However, hybridisation can occur for longer than 20 h, but different hybridisation times may adversely affect the data outcome.
13. Gene expression washing buffer 2 should be maintained at 37°C during the wash step. Failure to do so may result in alterations of experimental results.
14. The modified DNA should be stored in aluminium foil to protect it from the light.

Acknowledgments

The authors would like to thank Mark Meuth, Freddie Hamdy, Helen Owen, Saiful Miah, Helen Bryant, Katie Myers, Maggie Glover, Ewa Dudziec, Louise Goodwin, Leila Ayandi, and Sheila Blizard for their help, support, and encouragement with this work. This work was funded by a GSK Clinician Scientist fellowship to JWFC and project grants to JWFC from Yorkshire Cancer Research, Sheffield Hospitals Charitable trust, and the European Union (Framework 7).

References

1. Catto, J. W., Miah, S., Owen, H. C., Bryant, H., Dudziec, E., Larre, S., Milo, M., Rehman, I., Rosario, D. J., DiMartino, E., Knowles, M. A., Meuth, M., Harris, A. L., and Hamdy, F. C. (2009) Distinct microRNA alterations characterize high and low grade bladder cancer, *Cancer Research 69*, 8472–8481.
2. Lamy, P., Andersen, C. L., Dyrskjot, L., Torring, N., Orntoft, T., and Wiuf, C. (2006) Are microRNAs located in genomic regions associated with cancer? *British Journal of Cancer 95*, 1415–1418.
3. Croce, C. M. (2009) Causes and consequences of microRNA dysregulation in cancer, *Nature Reviews 10*, 704–714.
4. Saito, Y., Liang, G., Egger, G., Friedman, J. M., Chuang, J. C., Coetzee, G. A., and Jones, P. A. (2006) Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells, *Cancer Cell 9*, 435–443.
5. Lujambio, A., Calin, G. A., Villanueva, A., Ropero, S., Sanchez-Cespedes, M., Blanco, D., Montuenga, L. M., Rossi, S., Nicoloso, M. S., Faller, W. J., Gallagher, W. M., Eccles, S. A., Croce, C. M., and Esteller, M. (2008) A microRNA DNA methylation signature for human cancer metastasis, *Proceedings of the National Academy of Sciences of the United States of America 105*, 13556–13561.
6. Lehmann, U., Hasemeier, B., Christgen, M., Muller, M., Romermann, D., Langer, F., and Kreipe, H. (2008) Epigenetic inactivation of microRNA gene hsa-mir-9-1 in human breast cancer, *Journal of Pathology 214*, 17–24.
7. Saito, Y., and Jones, P. A. (2006) Epigenetic activation of tumor suppressor microRNAs in human cancer cells, *Cell Cycle 5*, 2220–2222.
8. Southgate, J., Hutton, K. A., Thomas, D. F., and Trejdosiewicz, L. K. (1994) Normal human urothelial cells in vitro: proliferation and induction of stratification, *Laboratory Investigation 71*, 583–594.
9. Catto, J. W., Azzouzi, A. R., Rehman, I., Feeley, K. M., Cross, S. S., Amira, N., Fromont, G., Sibony, M., Cussenot, O., Meuth, M., and Hamdy, F. C. (2005) Promoter hypermethylation is associated with tumor location, stage, and subsequent progression in transitional cell carcinoma, *Journal of Clinical Oncology 23*, 2903–2910.
10. Yates, D. R., Rehman, I., Abbod, M. F., Meuth, M., Cross, S. S., Linkens, D. A., Hamdy, F. C., and Catto, J. W. (2007) Promoter hypermethylation identifies progression risk in bladder cancer, *Clinical Cancer Research 13*, 2046–2053.
11. Dhawan, D., Hamdy, F. C., Rehman, I., Patterson, J., Cross, S. S., Feeley, K. M., Stephenson, Y., Meuth, M., and Catto, J. W. (2006) Evidence for the early onset of aberrant promoter methylation in urothelial carcinoma, *Journal of Pathology 209*, 336–343.

Chapter 13

In Vitro Functional Study of miR-126 in Leukemia

Zejuan Li and Jianjun Chen

Abstract

MicroRNAs (miRNAs, miRs) are postulated to be important regulators in various cancers, including leukemia. In a large-scale miRNA expression profiling analysis of 435 human miRNAs in 52 acute myeloid leukemia (AML) samples, we found that miR-126 and its minor counterpart in biogenesis, namely, miR-126*, were specifically aberrantly overexpressed in core binding factor (CBF) AMLs including both t(8;21)/*AML1-ETO* and inv(16)/*CBFB-MYH11* samples. Our in vitro gain- and loss-of-function experiments showed that forced expression of miR-126 inhibited apoptosis and increased the viability of AML cells, whereas the opposite effect was observed when endogenous expression of miR-126 was knocked down. In addition, through in vitro colony-forming/replating assays, we demonstrated that forced expression of miR-126 enhanced proliferation and colony-forming/replating capacity of mouse normal bone marrow progenitor cells alone and particularly, in cooperation with *AML1-ETO*, a fusion gene resulting from t(8;21). Thus, our data shows that miR-126 may play a critical role in the development of CBF leukemias. In the present chapter, the materials and protocols for the study of miR-126 in leukemia are described.

Key words: MicroRNA, miR-126, qPCR, Apoptosis, Cell viability, In vitro colony-forming/replating assay, CBF leukemia

1. Introduction

MicroRNAs (miRNAs) are endogenous small non-protein coding RNAs that can play important regulatory roles in diverse processes (1, 2). Emerging data show that miRNAs may function as oncogenes or tumor suppressors (2–15). In a large-scale analysis of miRNA expression profiling in acute myeloid leukemias (AMLs) using a bead-based miRNA expression profiling method (16), we observed that miR-126 and miR-126* were specifically highly expressed in core binding factor (CBF) AMLs, including both t(8;21)/*AML1*(*RUNX1*)-*ETO*(*RUNX1T1*) and

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_13, © Springer Science+Business Media, LLC 2011

inv(16)/*CBF*β-*MYH11* samples (17, 18). CBF AMLs represent 10–20% of primary AMLs; their differential expression patterns were confirmed by quantitative real-time PCR (i.e., qPCR) (19, 20). Gain- and loss-of-function experiments showed that miR-126/126* inhibits apoptosis and increase the viability of AML cells (Fig. 1).

Acute leukemias, like other human cancers, occur as the consequence of more than one mutation (21). Primary oncogenic events, such as those triggered by chromosomal rearrangements, are generally insufficient by themselves to cause leukemia and require secondary cooperating mutations to generate a fully transformed cell (22). For example, conditional expression of knock-in

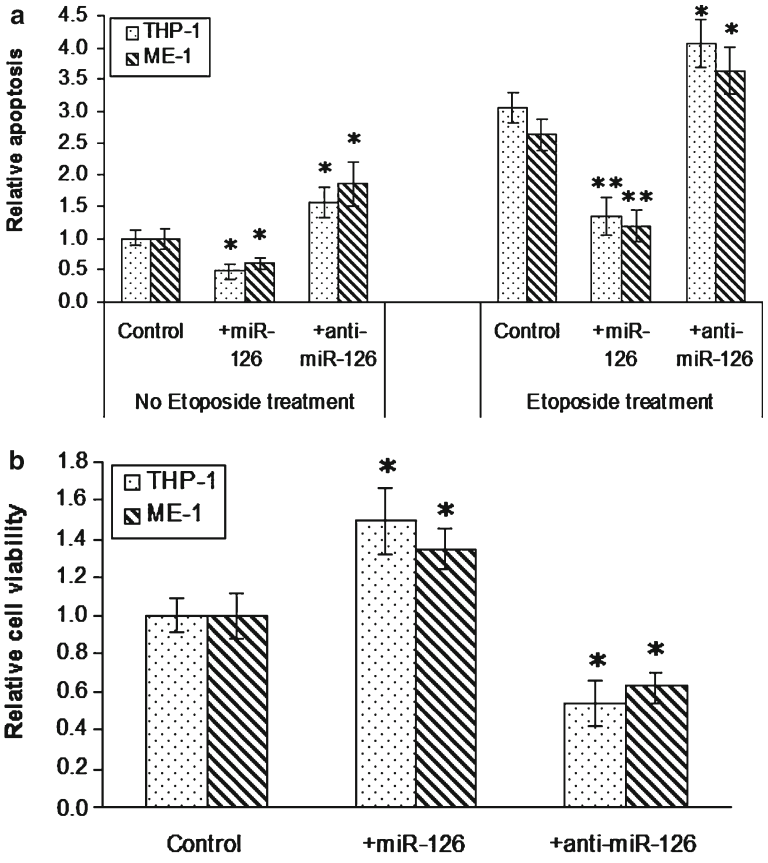


Fig. 1. miR-126 inhibits cell apoptosis and increases cell viability. (**a**) Forced expression of miR-126 by transduced MSCVpuro-miR-126 plasmid significantly inhibited, whereas downregulation of miR-126 by anti-miR-126 inhibitor significantly increased apoptosis in THP-1 and ME-1 cells with or without Etoposide treatment. (**b**) Forced expression of miR-126 significantly increased, whereas downregulation of miR-126 significantly decreased cell viability in THP-1 and ME-1 cells. Normalized mean values of three independent experiments and standard error (mean ± SE) are shown. *$p < 0.05$; **$p < 0.001$ (paired *t*-test).

CBFβ-MYH11 or *AML1-ETO* (AE) fusion gene, or expression of *AML1-ETO* following retroviral transduction/transplantation does not result in leukemia without a "second hit," such as those induced by *N*-ethyl-*N*-nitrosourea (ENU) (23–26). Although several secondary mutagenic events including FLT3, RAS, KIT, PU.1, and AML1 mutations, as well as truncated AE fusion (e.g., *AML1-ETO9a* (AE9a)) have been proposed (27), other previously unidentified secondary mutagenic events may also play an important role in the development of CBF leukemias. Due to the specific overexpression of miR-126 in CBF AMLs and its potential oncogenicity, we hypothesized that its aberrant overexpression may contribute to the development of CBF AMLs as a "second hit" in cooperation with the primary oncogenic events such as t(8;21)/inv(16) translocations. Indeed, our in vitro colony-forming/replating assays showed that miR-126 has a very significant synergistic effect with CBF leukemia fusions such as *AML1-ETO* in transforming mouse normal bone marrow progenitor cells in vitro (Fig. 2). Thus, our data suggest that miR-126 may play a critical role in the development of CBF leukemia. Here, we report the detailed materials and protocols for the above functional assays of miR-126, which would be very helpful for investigators to study other miRNAs in leukemogenesis.

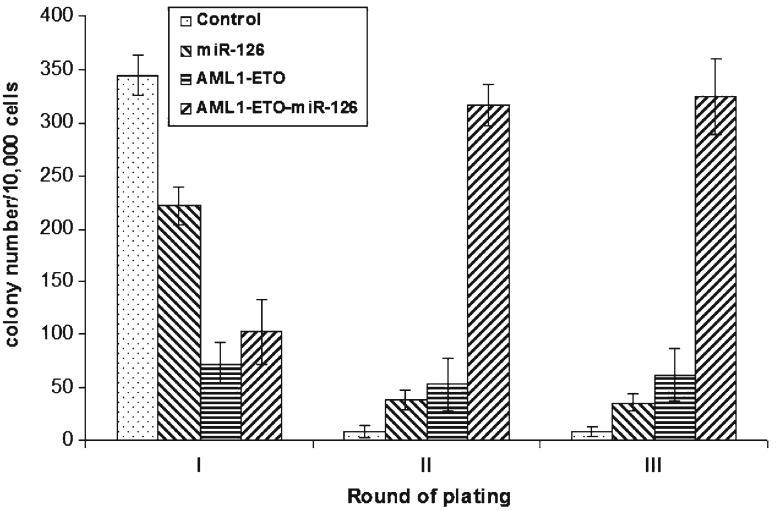


Fig. 2. miR-126 enhanced colony forming capacity of mouse normal bone marrow progenitor cells alone and particularly, in cooperation with *AML1-ETO* (*AE*). Only the colonies that contained at least 50 cells each were counted. Note that the colony numbers of first round of plating largely reflected the transduction efficiency, which was related to the size of the plasmids with a higher transduction efficiency for a smaller sized plasmid. Thus, the low colony number of the control in the second and third rounds of plating is not due to a low efficiency of transduction.

2. Materials

2.1. TaqMan qPCR Assay of the miRNAs

1. QIAGEN miRNeasy Mini Kit (QIAGEN, Valencia, CA) for total RNA (including miRNA) preparation.
2. TaqMan® MicroRNA Reverse Transcription Kit, TaqMan® Universal PCR Master Mix, TaqMan® MicroRNA Assays for miR-126 and for U6 endogenous control RNA are purchased from Applied Biosystems (Foster City, CA).
3. The 7900HT real-time PCR system (Applied Biosystems, Foster City, CA) is used to perform qPCR.

2.2. Construction of Retroviral Vectors

1. Primers to amplify miR-126 precursor sequence (236 bp): miR126-precursor-F: 5′-GATCTCTCGAGCGGAGCCTCA TATCAGCCAA-3′; miR126-precursor-R: 5′-TCCCAGAATT CTCTGCGCAGAGGCTCAGGGT-3′.
2. Genomic DNA template isolated from human normal bone marrow mononuclear cells purchased from AllCells (Emeryville, CA).
3. MSCVpuro (Clontech, Mountain View, CA) is a retroviral vector for cloning and gene expression.
4. *AML1-ETO* fusion gene in a vector MIGR1.

2.3. Apoptosis Assay

1. THP-1 and ME-1 cells.
2. Anti-miR126 inhibitor and inhibitor control (i.e., scrambled oligonucleotides) are purchased from Dharmacon (Chicago, IL).
3. Etoposide (Sigma, St. Louis, MO).
4. ApoONE Homogenous Caspase 3/7 Assay (Promega, Madison, WI).

2.4. Cell Viability Assay

1. THP-1 and ME-1 cells.
2. Anti-miR126 inhibitor and inhibitor control (i.e., scrambled oligonucleotides) are purchased from Dharmacon (Chicago, IL).
3. CellTiter-Blue Reagent (Promega, Madison, WI).

2.5. Retrovirus Preparation and Viral Titer Determination

1. HEK293T (a human embryonic kidney cell line; ATCC, Manassas, VA) for generation of retrovirus.
2. Rat1a cells (an embryonic rat fibroblast cell line; ATCC, Manassas, VA) for monitoring the viral titer.
3. PCL-Eco vector (IMGENEX, San Diego, CA) is a retroviral packaging vector.
4. QIAGEN Effectene transfection reagent (QIAGEN, Valencia, CA) for transfection of DNA (i.e., retroviral constructs) into 293T cells.

2.6. Colony-Forming/ Replating Assays

1. A cohort of 4- to 6-week-old normal C57BL/6J (known as B6) mice of both sexes as bone marrow donor mice.
2. Lyse red blood cells with ammonium chloride solution (0.8% NH_4Cl with 0.1 mM EDTA) (Stemcell Technologies, Vancouver, Canada).
3. Lineage Cell Depletion Kit (Miltenyi Biotec Inc., Auburn, CA), a cocktail of immunomagnetically labeled antibodies directed against CD5, CD45R (B220), CD11b (Mac-1), Gr-1 (Ly-6G/C), 7-4 (or 7/4; neutrophils), and Ter-119, is used to isolate Lin(-) cells.
4. Polybrene, 35 mm Petri dishes (Stemcell Technologies, Vancouver, Canada), β-mercaptoethanol, and Puromycin (Sigma, St. Louis, MO).
5. Murine stem cell factor (SCF), IL-3, and IL-6, and GM-CSF (R&D Systems, Minneapolis, MN).
6. Methocult M3231 methylcellulose medium (Stemcell Technologies, Vancouver, Canada).

2.7. Cell Culture and DNA Transfection

1. Dulbecco modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum and 1% penicillin–streptomycin (Invitrogen) for HEK293T cell culture.
2. Effectene Transfection Reagent is purchased from Qiagen (Valencia, CA) for DNA transfection of HEK293T cells.
3. ATCC-formulated RPMI-1640 (ATCC) containing 1% penicillin–streptomycin, 0.05 mM 2-mercaptoethanol, and 10% FBS (Invitrogen) for THP-1 cell culture.
4. RPMI-1640 (Invitrogen) supplemented with 20% FBS, 1% penicillin–streptomycin, and 1% HEPES (Invitrogen) for ME-1 cell culture.
5. Cell Line Nucleofector Kit V following program V01 and X01, respectively, using a nucleofector device (Amaxa Biosystems, Berlin, Germany) for DNA transfection of THP-1 and ME-1 cells.

3. Methods

3.1. TaqMan qPCR Assay of the miRNAs

1. Prepare total RNA (containing all the miRNA molecules) with QIAGEN miRNeasy Mini Kit (see Note 1). Dilute RNA to 1.6 ng/μl with DEPC water.
2. Use 8 ng of total RNA per the reverse transcription (RT; see Note 2). Prepare RT reaction with 0.15 μl of 100 mM dNTPs (with dTTP), 1 μl of MultiScribe™ Reverse Transcriptase (50 U/μl), 1.5 μl of 10× Reverse Transcription

Buffer, 0.19 μl of RNase Inhibitor (20 U/μl), 5 μl of RNA (1.6 ng/μl), and 4.16 μl of Nuclease-free water (see Note 3). The conditions are: 16°C for 30 min, 42°C for 30 min, and 85°C for 5 min.

3. Prepare PCR reaction with 1 μl of TaqMan MicroRNA Assay (20×), 1.33 μl of Product from RT reaction, 10 μl of TaqMan 2× Universal PCR Master Mix (No AmpErase UNGa), and 7.67 μl of Nuclease-free water (see Note 4). qPCR is performed on the Applied Biosystems 7900HT real-time PCR system. The PCR conditions are: 95°C for 10 min and then 15 s at 95°C and 1 min at 60°C for 40 cycles. U6 RNA is used as endogenous control.
4. Each candidate gene is assayed in triplicate.
5. ΔC_t values are used for the analysis. Briefly, in each sample, a $\Delta Ct_{(miRNA\text{-}U6)}$, which is equal to the difference in threshold cycles for miRNA (i.e., target) and U6 RNA (i.e., reference) [i.e., $\Delta Ct_{(miRNA\text{-}U6)} = Ct_{miRNA} - Ct_{U6}$] is calculated.
6. Hierarchical clustering (28) is performed on the PCR data. Values for each candidate gene are mean-centered before clustering. TIGR TMeV (version 4.0) (29) is used for the analysis.

3.2. Construction of Retroviral Vectors

1. The precursor sequence of miR-126 (236 bp) is PCR amplified from human normal bone marrow mononuclear cells.
2. miR-126 precursor fragment is then cloned into MSCVpuro through *XhoI* (CTCGAG) and *EcoRI* (GAATTC) sites as "MSCV-LTR-*BglII*-*XhoI*-miR-126 (precursor)-*EcoRI*-Pgk-puromycin," and named MSCVpuro-miR126.
3. AE fusion gene is isolated from the original construct, namely, MIGR1-AE, and then cloned into the MSCVpuro and MSCVpuro-miR-126 vectors, respectively. In MSCVpuro-AE-miR-126, *AE* was placed in front of miR-126 and close to the promoter region.
4. The inserts are confirmed by DNA sequencing.
5. A total of four retroviral constructs including MSCVpuro (used as control), MSCVpuro-miR-126, MSCVpuro-AE, and MSCVpuro-AE-miR-126 are ready for further functional studies.

3.3. Apoptosis Assay

1. THP-1 or ME-1 cells are plated at a concentration of 10,000 cells per well in triplicate for each condition in a 96-well plate 24 h after transfection with MSCVpuro-miR-126, the control plasmid MSCVpuro, anti-miR126 inhibitor, or inhibitor control.

2. Etoposide (5 μM) or the same volume of DMSO (mock treatment) is added to each well.
3. Caspase-3 and caspase-7 activation are detected using ApoONE Homogenous Caspase 3/7 Assay 24 h later following the manufacturer's manual.

3.4. Cell Viability Assay

1. THP-1 or ME-1 cells are plated at a concentration of 10,000 cells per well in triplicate for each condition in a 96-well plate 24 h after transfection with MSCVpuro-miR-126, MSCVpuro, anti-miR126 inhibitor, or inhibitor control.
2. Metabolic activity of the cells is determined using CellTiter-Blue Reagent (see Note 5) 24 h later following the manufacturer's manual.

3.5. Retrovirus Preparation and Viral Titer Determination

1. To produce the ecotropic retrovirus, 0.5×10^6 293T cells are plated in a 60 mm dish the day before transfection.
2. 1.8 μg of retroviral vector DNA and 1.2 μg PCL-Eco vector (IMGENEX, San Diego, CA) are transfected using the QIAGEN Effectene transfection reagent.
3. Then, medium is changed with 1 ml of 10% FBS/DMEM after 20 h of transfection.
4. After 48 h of transfection, the virus-containing medium is collected and filtered with a 0.45 μm cellulose acetate (low protein binding) filter (see Note 6). For each construct, two dishes of cells are transfected and the virus-containing medium is mixed together.
5. Rat1a cells (an embryonic rat fibroblast cell line) are infected with virus-containing medium to monitor the viral titer.

3.6. Colony-Forming and Replating Assays

1. Bone marrow cells are obtained from a cohort of 4- to 6-week-old normal C57BL/6J (known as B6) mice of both sexes. Inject mice with 150 mg/Kg of 5-Fluorouracil (40–50 μl) 5 days before sacrificing them.
2. Red blood cells are lysed in ammonium chloride solution (0.8% NH_4Cl with 0.1 mM EDTA) and then passed through a 40 μM nylon mesh to create a single cell suspension.
3. To enrich primitive hematopoietic progenitors from the bone marrow cells, mature erythroid, lymphoid, and myeloid cells are depleted by incubation with a cocktail of immunomagnetically labeled antibodies directed against CD5, CD45R (B220), CD11b (Mac-1), Gr-1 (Ly-6G/C), 7-4 (or 7/4; neutrophils), and Ter-119 using Lineage Cell Depletion Kit.
4. The antibody-labeled cells are then incubated with magnetic colloid and pass through a magnetic column. Spin down and

resuspend the lin(-) cells (hematopoietic progenitor cells) in RPMI media containing 10% FBS and supplemented with 100 ng/ml murine SCF, 10 ng/ml IL-3, and 10 ng/ml IL-6, and 55 μM β-mercaptoethanol (see Note 7). Culture the cells overnight.

5. The next day, an aliquot of 2×10^4 hematopoietic progenitor cells (will be plated into two dishes after transduction) in a volume of 0.4 ml in medium are added to 1.6 ml retroviral supernatant of a single construct (with an adjusted titer of 1×10^6 CFU/ml) and 4 μg/ml polybrene in a conical tube.
6. Centrifuge the tubes at $2{,}000 \times g$ for 2.5 h at 30°C (i.e., "spinoculation" (30–34)) and then replace with fresh media and incubate for 20 h at 37°C.
7. Next day, the same procedure is repeated once.
8. On the day following the second spinoculation, an equivalent of 1×10^4 of the initial lineage-depleted cells are plated into duplicate 35 mm Petri dishes in 1.1 ml of Methocult M3230 methylcellulose medium (Stem Cell Technologies Inc.) containing 10 ng each of murine recombinant IL-3, IL-6, and granulocyte-macrophage colony-stimulating factor per ml and 100 ng of murine recombinant SCF per ml with 1.5 μg of puromycin per ml. For each sample, there are two dishes.
9. Cultures are incubated at 37°C in a humidified atmosphere of 5% CO_2 in air.
10. The colonies are replated every 7 days under the same conditions for 3–6 times (see Note 8).
11. Averages and error margins are determined from two independent transduction experiments. Only the colonies that contain at least 50 cells each are counted.

4. Notes

1. Purify the RNA with QIAGEN miRNeasy Mini Kit using the protocol: Purification of Total RNA, Including Small RNAs, from Animal Cells.
2. Before performing RT-PCR, do not denature the RNA. Denaturation of the RNA may reduce the yield of cDNA.
3. Prepare RT master mix. Mix gently. Centrifuge to bring the solution to the bottom of the tube. Incubate the tube on ice for 5 min and keep on ice until you are ready to load the thermal cycler.
4. Keep all TaqMan MicroRNA Assays protected from light, in the freezer, until you are ready to use them. Excessive exposure to light may affect the fluorescent probes.

5. CellTiter-Blue cell viability assay can be multiplexed with the ApoONE Homogenous Caspase 3/7 assay. CellTiter-Blue Reagent is added to each well 24 h after transfection, and the cells are incubated 3 h prior to recording fluorescence ($560_{Ex}/590_{Em}$). Then caspase activity is measured in the same wells by adding 120 μl of the ApoONE Homogenous Caspase 3/7 assay Reagent. Cells are incubated for three additional hours at ambient temperature prior to recording fluorescence ($485_{Ex}/527_{Em}$).
6. The viral supernatant can be collected from 36 to 72 h after transfection. Viral titer will drop dramatically after 72 h of transfection.
7. When spin down the mouse progenitor cells, stop the spinning without brake.
8. In our previous in vitro colony-forming/replating assays, we conducted only three rounds of plating (35). However, mouse bone marrow cells obtained from different strains of mice may have different capacity in forming colonies. If cells transduced with an empty vector still form a relatively large number (e.g., over 50) of tertiary colonies (colonies formed at the end of the third plating in methylcellulose) as reported by Okuda et al. (36), replating the cells for all of the transductions until cells transduced with empty vector(s) form no or very few colonies is necessary.

Acknowledgements

We thank Dr. P. Paul Liu at National Human Genome Research Institute for providing the ME-1 cell line and Dr. Dong-Er Zhang for providing the MIGR1-AE construct. This work was supported in part by National Institutes of Health (NIH) CA127277 (J.C.) and the G. Harold and Leila Y. Mathers Charitable Foundation (J.C.).

References

1. Bartel, D. P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell 116*, 281–297.
2. Wu, W., Sun, M., Zou, G. M., and Chen, J. (2007) MicroRNA and cancer: current status and prospective. *Int J Cancer 120*, 953–960.
3. Croce, C. M., and Calin, G. A. (2005) miRNAs, cancer, and stem cell division. *Cell 122*, 6–7.
4. Esquela-Kerscher, A., and Slack, F. J. (2006) Oncomirs – microRNAs with a role in cancer. *Nat Rev Cancer 6*, 259–269.
5. He, L., Thomson, J. M., Hemann, M. T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S. W., Hannon, G. J., and Hammond, S. M. (2005) A microRNA polycistron as a potential human oncogene. *Nature 435*, 828–833.
6. Tagawa, H., Karube, K., Tsuzuki, S., Ohshima, K., and Seto, M. (2007) Synergistic action of the microRNA-17 polycistron and Myc in aggressive cancer development. *Cancer Sci 98*, 1482–1490.
7. Voorhoeve, P. M., le Sage, C., Schrier, M., Gillis, A. J., Stoop, H., Nagel, R., Liu, Y. P.,

van Duijse, J., Drost, J., Griekspoor, A., Zlotorynski, E., Yabuta, N., De Vita, G., Nojima, H., Looijenga, L. H., and Agami, R. (2006) A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. *Cell 124*, 1169–1181.

8. Chan, J. A., Krichevsky, A. M., and Kosik, K. S. (2005) MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res 65*, 6029–6033.
9. Kluiver, J., Poppema, S., de Jong, D., Blokzijl, T., Harms, G., Jacobs, S., Kroesen, B. J., and van den Berg, A. (2005) BIC and miR-155 are highly expressed in Hodgkin, primary mediastinal and diffuse large B cell lymphomas. *J Pathol 207*, 243–249.
10. Costinean, S., Zanesi, N., Pekarsky, Y., Tili, E., Volinia, S., Heerema, N., and Croce, C. M. (2006) Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice. *Proc Natl Acad Sci U S A 103*, 7024–7029.
11. Ma, L., Teruya-Feldstein, J., and Weinberg, R. A. (2007) Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature 449*, 682–688.
12. Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F., and Croce, C. M. (2002) Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A 99*, 15524–15529.
13. Cimmino, A., Calin, G. A., Fabbri, M., Iorio, M. V., Ferracin, M., Shimizu, M., Wojcik, S. E., Aqeilan, R. I., Zupo, S., Dono, M., Rassenti, L., Alder, H., Volinia, S., Liu, C. G., Kipps, T. J., Negrini, M., and Croce, C. M. (2005) miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc Natl Acad Sci USA 102*, 13944–13949.
14. Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K. L., Brown, D., and Slack, F. J. (2005) RAS is regulated by the let-7 microRNA family. *Cell 120*, 635–647.
15. He, L., He, X., Lim, L. P., de Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., Ridzon, D., Jackson, A. L., Linsley, P. S., Chen, C., Lowe, S. W., Cleary, M. A., and Hannon, G. J. (2007) A microRNA component of the p53 tumour suppressor network. *Nature 447*, 1130–1134.
16. Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R., and Golub, T. R. (2005) MicroRNA expression profiles classify human cancers. *Nature 435*, 834–838.
17. Bloomfield, C. D., Ruppert, A. S., Mrozek, K., Kolitz, J. E., Moore, J. O., Mayer, R. J., Edwards, C. G., Sterling, L. J., Vardiman, J. W., Carroll, A. J., Pettenati, M. J., Stamberg, J., Byrd, J. C., Marcucci, G., and Larson, R. A. (2004) Core binding factor acute myeloid leukemia. Cancer and Leukemia Group B (CALGB) Study 8461. *Ann Hematol 83 Suppl 1*, S84–S85.
18. de Bruijn, M. F., and Speck, N. A. (2004) Core-binding factors in hematopoiesis and immune function. *Oncogene 23*, 4238–4248.
19. Look, A. T. (1997) Oncogenic transcription factors in the human acute leukemias. *Science 278*, 1059–1064.
20. Strout, M. P., Marcucci, G., Caligiuri, M. A., and Bloomfield, C. D. (1999) Core-binding factor (CBF) and MLL-associated primary acute myeloid leukemia: biology and clinical implications. *Ann Hematol 78*, 251–264.
21. Gilliland, D. G. (2001) Hematologic malignancies. *Curr Opin Hematol 8*, 189–191.
22. Pui, C. H., and Jeha, S. (2007) New therapeutic strategies for the treatment of acute lymphoblastic leukaemia. *Nat Rev Drug Discov 6*, 149–165.
23. Castilla, L. H., Garrett, L., Adya, N., Orlic, D., Dutra, A., Anderson, S., Owens, J., Eckhaus, M., Bodine, D., and Liu, P. P. (1999) The fusion gene Cbfb-MYH11 blocks myeloid differentiation and predisposes mice to acute myelomonocytic leukaemia. *Nat Genet 23*, 144–146.
24. Yuan, Y., Zhou, L., Miyamoto, T., Iwasaki, H., Harakawa, N., Hetherington, C. J., Burel, S. A., Lagasse, E., Weissman, I. L., Akashi, K., and Zhang, D. E. (2001) AML1-ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations. *Proc Natl Acad Sci USA 98*, 10398–10403.
25. Higuchi, M., O'Brien, D., Kumaravelu, P., Lenny, N., Yeoh, E. J., and Downing, J. R. (2002) Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia. *Cancer Cell 1*, 63–74.
26. de Guzman, C. G., Warren, A. J., Zhang, Z., Gartland, L., Erickson, P., Drabkin, H., Hiebert, S. W., and Klug, C. A. (2002) Hematopoietic stem cell expansion and distinct myeloid developmental abnormalities in a murine model of the AML1-ETO translocation. *Mol Cell Biol 22*, 5506–5517.

27. Peterson, L. F., Boyapati, A., Ahn, E. Y., Biggs, J. R., Okumura, A. J., Lo, M. C., Yan, M., and Zhang, D. E. (2007) Acute myeloid leukemia with the 8q22;21q22 translocation: secondary mutational events and alternative t(8;21) transcripts. *Blood 110*, 799–805.
28. Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA 95*, 14863–14868.
29. Saeed, A. I., Bhagabati, N. K., Braisted, J. C., Liang, W., Sharov, V., Howe, E. A., Li, J., Thiagarajan, M., White, J. A., and Quackenbush, J. (2006) TM4 microarray software suite. *Methods Enzymol 411*, 134–193.
30. Lavau, C., Du, C., Thirman, M., and Zeleznik-Le, N. (2000) Chromatin-related properties of CBP fused to MLL generate a myelodysplastic-like syndrome that evolves into myeloid leukemia. *Embo J 19*, 4655–4664.
31. Lavau, C., Luo, R. T., Du, C., and Thirman, M. J. (2000) Retrovirus-mediated gene transfer of MLL-ELL transforms primary myeloid progenitors and causes acute myeloid leukemias in mice. *Proc Natl Acad Sci USA 97*, 10984–10989.
32. Lavau, C., Szilvassy, S. J., Slany, R., and Cleary, M. L. (1997) Immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX-ENL. *Embo J 16*, 4226–4237.
33. Slany, R. K., Lavau, C., and Cleary, M. L. (1998) The oncogenic capacity of HRX-ENL requires the transcriptional transactivation activity of ENL and the DNA binding motifs of HRX. *Mol Cell Biol 18*, 122–129.
34. Luo, R. T., Lavau, C., Du, C., Simone, F., Polak, P. E., Kawamata, S., and Thirman, M. J. (2001) The elongation domain of ELL is dispensable but its ELL-associated factor 1 interaction domain is essential for MLL-ELL-induced leukemogenesis. *Mol Cell Biol 21*, 5678–5687.
35. Li, Z., Lu, J., Sun, M., Mi, S., Zhang, H., Luo, R. T., Chen, P., Wang, Y., Yan, M., Qian, Z., Neilly, M. B., Jin, J., Zhang, Y., Bohlander, S. K., Zhang, D. E., Larson, R. A., Le Beau, M. M., Thirman, M. J., Golub, T. R., Rowley, J. D., and Chen, J. (2008) Distinct microRNA expression profiles in acute myeloid leukemia with common translocations. *Proc Natl Acad Sci U S A 105*, 15535–15540.
36. Okuda, T., Cai, Z., Yang, S., Lenny, N., Lyu, C. J., van Deursen, J. M., Harada, H., and Downing, J. R. (1998) Expression of a knocked-in AML1-ETO leukemia gene inhibits the establishment of normal definitive hematopoiesis and directly generates dysplastic hematopoietic progenitors. *Blood 91*, 3134–3143.

Chapter 14

Prediction of the Biological Effect of Polymorphisms Within MicroRNA Binding Sites

Debora Landi, Roberto Barale, Federica Gemignani, and Stefano Landi

Abstract

MicroRNAs (miRNAs) are negative gene regulators acting at the 3′UTR level, modulating the translation of cancer-related genes. Single-nucleotide polymorphisms (SNPs) within the 3′UTRs could impact the miRNA-dependent gene regulation either by weakening or by reinforcing the binding sites. Thus, the alteration of the normal regulation of a given gene could affect the individual's risk of cancer. Therefore, it is helpful to develop a tool enabling the researchers to predict which of the many SNPs could really impact the regulation of a target gene. At present, there are several available databases and algorithms able to predict potential binding sites in the 3′UTR of genes. However, each algorithm gives different predictions and none of them gives, for each polymorphism, a direct measurement of the biological impact. We propose an approach allowing the assignment to each polymorphism a ranking of its biological impact. The method is based on a simple elaboration of predictions from preexisting well-established algorithms. As an example, we show the application of this approach to 140 genes candidate for colorectal cancer (CRC). These genes were identified following a genome-wide sequencing of 20,857 transcripts from 18,191 genes in 11 CRC specimens and were found somatically mutated and thought to be crucial for the development of cancer.

Key words: MicroRNA, Colorectal cancer, Polymorphisms, Algorithms

1. Introduction

MicroRNAs (miRNAs) are noncoding single-stranded RNAs of about 21–25 nucleotides regulating gene translation in both plants and animals, by binding the 3′UTR of target mRNAs. MiRNAs are processed from larger (~80-nt) precursor hairpins by the RNase III enzyme Dicer into miRNA:miRNA* duplexes. One strand of these duplexes associates with the RNA-induced silencing complex (RISC), whereas the other is generally degraded (1).

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_14, © Springer Science+Business Media, LLC 2011

The gene expression is regulated by miRNAs at posttranscriptional level in two manners: when there are mismatches between miRNA and target, the translation is reduced or inhibited, whereas in case of perfect matching, the miRNA acts as silencing RNA (siRNA) and the target is degraded by cytosolic enzymes. In the former case, the binding of miRNA to its target does not affect the mRNA level, while in the latter case it does (1). Although siRNA was discovered first in plants, in animal cells miRNAs play the major role. The rules governing the pairing between miRNA and target have been established: the maximal complementarity is restricted to a region called "seed sequence" on the 5′ end of miRNA, spanning the nucleotides 2–8 (2). Within the seed sequence a maximum of one mismatch is tolerated. Moreover, the G:T pairing is admitted in the miRNA–mRNA hybrid (2).

Recent evidences indicate that miRNAs are involved in most biological and pathological processes, including development timing, cellular differentiation, proliferation, apoptosis, insulin secretion, cholesterol biosynthesis, and tumorigenesis (3). However, the binding between miRNA and mRNA can be affected by single-nucleotide polymorphisms (SNPs) that can reside in the target site: SNPs can either weaken or reinforce the binding sites thereby altering the normal regulation of a given gene (4). Therefore, it is helpful to develop a tool enabling the researchers to predict which of the SNPs could really impact the regulation of a target gene. At present, there are several available databases and algorithms able to predict potential binding sites in the 3′UTR of genes. The scanning algorithms are based on sequence complementarity between the mature miRNA and the target site, the binding free energy of the miRNA–target duplex, the evolutionary conservation of the target site sequence, and the target position in aligned UTRs of homologous genes (5). However, each algorithm gives different predictions, and none of them gives, for each polymorphism, a direct measurement of the biological impact. We propose an approach allowing the ranking of each polymorphism falling within 3′UTRs, from the most biologically neutral to the most likely affecting the miRNA binding site. The method is based on a reelaboration of predictions from preexisting well-established algorithms. The putative miRNA binding sites are identified by means of the following algorithms: *miRBase*, *miRanda*, *PicTar*, *Diana-MicroT*, and *TargetScanS* (for all of them the default parameters were used). The method uses the outputs from all the algorithms and does not discharge any prediction, preventing any loss of information. As an example, we show this approach applied to 140 genes candidate for colorectal cancer (CRC). These genes derived from a genome-wide study in which 20,857 transcripts from 18,191 human genes were sequenced in 11 CRC specimens and were identified a total of 140 somatically mutated genes (called *CAN*-genes and reported in Table 1), thought to be crucial for the cancer development (6).

Table 1
Starting list of CAN-genes relevant for colorectal cancer, evaluated for the presence of polymorphic miRNA target sites

Gene symbol	Gene name	Gene symbol	Gene name
ABCA1	ATP-binding cassette, subfamily A, member 1	*KIAA1409*	Hypothetical protein LOC57578
ABCB11	ATP-binding cassette, subfamily B, member 11	*KIAA2022*	KIAA2022
ACSL5	Acyl-CoA synthetase long-chain family member 5	*KRAS*	v-Ki-ras2 Kirsten rat sarcoma viral onc. homolog
ADAM19	ADAM metallopeptidase domain 19	*LAMA1*	Laminin, alpha 1
ADAM29	ADAM metallopeptidase domain 29	*LCN9*	Lipocalin 9
ADAMTS18	ADAM metallopeptidase with thrombospondin mot. 18	*LGR6*	Leucine-rich repeat-containing G protein-coupled receptor 6
ADAMTS20	ADAM metallopeptidase with thrombospondin mot. 20	*LMO7*	LIM domain 7
ADAMTSL3	ADAMTS-like 3	*MAP1B*	Microtubule-associated protein 1B
ADARB2	Adenosine deaminase, RNA-specific, B2	*MAP2*	Microtubule-associated protein 2
AGC1	Aggrecan	*MAP2K7*	Mitogen-activated protein kinase kinase 7
AKAP12	A kinase (PRKA) anchor protein (gravin) 12	*MAPK8IP2*	Mitogen-activated protein kinase 8 interacting prot. 2
AKAP6	A kinase (PRKA) anchor protein 6	*MCM3AP*	Minichrom. mainten. complex comp. 3 assoc. prot.
ALK	Anaplastic lymphoma receptor tyrosine kinase	*MGC20470*	Ubiquitin-like
APC	Adenomatosis polyposis coli	*MKRN3*	Makorin, ring finger protein, 3
ARHGEF10	Rho guanine nucleotide exchange factor (GEF) 10	*MMP2*	Matrix metallopeptidase 2
ATP11A	ATPase, class VI, type 11A	*MYO18B*	Myosin XVIIIB
ATP13A1	ATPase type 13A1	*MYO5C*	Myosin VC
ATP13A5	ATPase type 13A5	*MYOHD1*	Myosin XIX

(continued)

Table 1 (continued)

Gene symbol	Gene name	Gene symbol	Gene name
BCL9	B-cell CLL/lymphoma 9	*NAV3*	Neuron navigator 3
C10orf137	Chromosome 10 open reading frame 137	*NF1*	Neurofibromin 1
C13orf7	Ring finger protein 219	*NOS3*	Nitric oxide synthase 3 (endothelial cell)
C14orf115	Chromosome 14 open reading frame 115	*NTNG1*	Netrin G1
C15orf2	Chromosome 15 open reading frame 2	*NUP210*	Nucleoporin 210 kDa
C1QR1	CD93 molecule	*OR51E1*	Olfactory receptor, family 51, subfamily E, member 1
CACNA2D3	Calcium channel, voltage-dep., alpha 2/delta sub. 3	*P2RX7*	Purinergic receptor P2X, ligand-gated ion channel, 7
CD109	CD109 molecule	*PCDH11X*	Protocadherin 11 X-linked
CHL1	Cell adhesion molecule with homology to L1CAM	*PCDHA9*	Protocadherin alpha 9
CHR415SYT	Synaptotagmin XIV-like	*PIK3CA*	Phosphoinositide-3-kinase, catalytic, α-polypeptide
CLSTN2	Calsyntenin 2	*PKNOX1*	PBX/knotted 1 homeobox 1
CNTN4	Contactin 4	*PLB1*	Phospholipase B1
COL3A1	Collagen, type III, alpha 1	*PLCG2*	Phospholipase C, gamma 2
CPAMD8	C3 and PZP-like, α-2-macr-oglob. domain containing 8	*PRDM9*	PR domain containing 9
CSMD3	CUB and Sushi multiple domains 3	*PRKD1*	Polycystic kidney and hepatic disease 1
CUTL1	Cut-like homeobox 1	*PTEN*	Phosphatase and tensin homolog
CX40.1	Gap junction protein, delta 4, 40.1 kDa	*PTPRD*	Protein tyrosine phosphatase, receptor type, D
DPP10	Dipeptidyl-peptidase 10	*PTPRS*	Protein tyrosine phosphatase, receptor type, S
DSCAML1	Down syndrome cell adhesion molecule like 1	*PTPRU*	Protein tyrosine phosphatase, receptor type, U
DTNB	Dystrobrevin, beta	*RASGRF2*	Ras protein-specific guanine nucleotide-releasing factor 2
EDD1	UBIQUITIN protein ligase E3 comp. n-recognin 5	*RET*	Ret proto-oncogene

(continued)

Table 1 (continued)

Gene symbol	Gene name	Gene symbol	Gene name
EPHA3	EPH receptor A3	*ROBO1*	Roundabout, axon guidance receptor, homolog 1
EPHB6	EPH receptor B6	*RUNX1T1*	Runt-related transcription factor 1; translocated to 1
ERCC6	Excision repair cross-complementarity, complementarity group 6	*SCN3B*	Sodium channel, voltage-gated, type III, beta
EYA4	Eyes absent homolog 4	*SEC8L1*	Exocyst complex component 4
F8	Coagulation factor VIII, procoagulant component	*SEMA3D*	Sema dom., immunogl. dom., short basic dom., 3D
FBN2	Fibrillin 2	*SFRS6*	Splicing factor, arginine/serine-rich 6
FBXW7	F-box and WD repeat domain containing 7	*SH3TC1*	SH3 domain and tetratricopeptide repeats 1
FLJ10404	Hypothetical protein FLJ10404	*SHANK1*	SH3 and multiple ankyrin repeat domains 1
FLJ13305	Hypothetical protein FLJ13305	*SLC22A15*	Solute carrier family 22, member 15
FLNC	Filamin C, gamma (actin binding protein 280)	*SLC29A1*	Solute carrier family 29, member 1
FN1	Fibronectin 1	*SMAD2*	SMAD family member 2
GALNS	Galactosamine (*N*-acetyl)-6-sulfate sulfatase	*SMAD3*	SMAD family member 3
GLI3	GLI-Kruppel family member GLI3	*SMAD4*	SMAD family member 4
GNAS	GNAS complex locus	*SMTN*	Smoothelin
GPR112	G protein-coupled receptor 112	*SORL1*	Sortilin-rel. receptor, LDLR class A repeats-containing
GPR158	G protein-coupled receptor 158	*STAB1*	Stabilin 1
GRID1	Glutamate receptor, ionotropic, delta 1	*TAF2*	TAF2 RNA pol II, TATA box bind. protein
GRM1	Glutamate receptor, metabotropic 1	*TBX22*	T-box 22
GUCY1A2	Guanylate cyclase 1, soluble, alpha 2	*TCERG1L*	Transcription elongation regulator 1-like
HAPIP	Kalirin, RhoGEF kinase	*TCF7L2*	Transcription factor 7-like 2

(continued)

Table 1 (continued)

Gene symbol	Gene name	Gene symbol	Gene name
HAPLN1	Hyaluronan and proteoglycan link protein 1	*TGFBR2*	Transforming growth factor, beta receptor II
HIST1H1B	Histone cluster 1, H1b	*TGM3*	Transglutaminase 3
IGFBP3	Insulin-like growth factor binding protein 3	*TIAM1*	T-cell lymphoma invasion and metastasis 1
IGSF22	Immunoglobulin superfamily, member 22	*TLR9*	Toll-like receptor 9
IRS4	Insulin receptor substrate 4	*TNN*	Tenascin N
ITGAE	Integrin, alpha E	*TP53*	Tumor protein p53
K6IRS3	Keratin 73	*UHRF2*	Ubiquitin-like, contain. PHD and RING finger dom.2
KCNQ5	Potassium voltage-gated chan., KQT-like subf., 5	*UQCRC2*	Ubiquinol-cytochrome c reductase core protein II
KIAA0182	KIAA0182	*ZNF262*	Zinc finger, MYM-type 4
KIAA0367	KIAA0367	*ZNF442*	Zinc finger protein 442
KIAA0556	KIAA0556	*ZNF521*	Zinc finger protein 521

2. Materials

For this work, we used the followings web sites:

2.1. Prediction of miRNAs Binding Sites

1. *miRBase* (http://www.mirbase.org/).
2. *miRAnda* (http://www.microrna.org/).
3. *PicTar* (http://pictar.mdc-berlin.de/).
4. *Diana-MicroT* (http://www.diana.pcbi.upenn.edu/cgi-bin/micro_t.cgi).
5. *TargetScan Human* (http://www.targetscan.org/).

2.2. Selection of 3′ UTR

1. UCSC genome browser (http://www.genome.ucsc.edu).

2.3. Search of SNPs in the Target Sites

1. SNP database (dbSNP; http://www.ncbi.nlm.nih.gov/SNP/).

2.4. Calculation of the Gibbs Binding Free Energy

1. *RNAcofold* (http://rna.tbi.univie.ac.at/cgi-bin/RNAcofold.cgi).

3. Methods

3.1. Prediction of miRNAs Binding Sites with miRBase

miRBase (http://www.mirbase.org/) is a database shared in three parts: miRBase Registry includes the miRNA gene nomenclature; miRBase Sequence is the primary online repository for miRNA sequences data and annotation; miRBase Targets is a comprehensive new database of predicted miRNA target genes (7).

1. Follow the link: "Targets."
2. Follow the link: "Enter."
3. Follow the link: "Search" on the left.
4. Insert the name of the gene (e.g., *ABCA1*), select the genome (e.g., *Homo sapiens*), and push "Search."
5. Push "View" to have the list of miRNAs and the binding sites.

3.2. Prediction of miRNAs Binding Sites with miRAnda

miRAnda (http://www.microrna.org/) is an algorithm that considers the sequence complementarity between the mature miRNA and the target site, binding free energy of the miRNA–target duplex, and the evolutionary conservation of the target position in aligned UTRs of homologous genes (5).

1. Follow the link: "Target mRNA."
2. Insert the target mRNA (e.g., *ABCA1*), the species (e.g., *H. sapiens*), and push "Go."
3. Follow the link: "View alignment details" to have the list of miRNAs and the binding sites.

3.3. Prediction of miRNAs Binding Sites with PicTar

PicTar (http://pictar.mdc-berlin.de/) computes a maximum likelihood score that a given RNA sequence (3′UTR region) is targeted by a fixed set of miRNAs (8).

1. Push "click here for Pic Tar predictions in vertebrates, flies, and nematodes (Lall et al. 2006)."
2. Insert the name of the gene (e.g., *ABCA1*) and push "Search for all miRNAs predicted to target a gene" to have the list of miRNAs and the binding sites.
3. The output consists of the predicted miRNA binding sites (only the seed match is shown), highlighted in yellow onto the multispecies alignment (*H. sapiens* on the top). All the miRNAs analyzed are reported independently on separate sections of the output.

3.4. Prediction of miRNAs Binding Sites with Diana-MicroT

Diana-MicroT (http://www.diana.pcbi.upenn.edu/cgi-bin/micro_t.cgi) finds microRNA/target duplexes that are conserved in humans and mice with the minimum free energy (10). This

algorithm at the moment does not allow us to insert the gene name directly. A 3′UTR sequence must be provided, and only one sequence is allowed. In this case, for all the *CAN*-genes mutated in CRC, we selected their 3′UTR regions defined as transcribed sequences from the stop codon to the end of the last exon of each gene.

There are several different ways to extract 3′UTRs. We will illustrate one of them based on the UCSC genome browser (http://www.genome.ucsc.edu).

1. Follow the link: "Genome Browser."
2. Insert the name of the gene (e.g., *ABCA1*).
3. Follow the link under "RefSeq Genes."
4. On the graphical interface, click the arrowed line under "RefSeq Genes."
5. Follow the link "Genomic Sequence."
6. Check only the box marked as "3′UTR exons" and push "Submit."
7. Save the FASTA sequences as ".txt" file (these will be used in step 9).
8. In the DIANA MicroTest page, select "human" in the box at the top on the right (in this case, it will use its own internal miRNA database). Alternatively, one can insert the miRNAs sequences in FASTA format in the box marked as "Enter miRNA sequence."
9. Insert the 3′UTR of the gene (e.g., *APC*) in FASTA format in the box marked as "Enter your DNA sequence (fasta format)."
10. Push "Invia query" to have the list of miRNAs and the binding sites.

3.5. Prediction of miRNAs Binding Sites with TargetScan Human5.1

TargetScan Human (http://www.targetscan.org/) searches the 3′UTRs for segments of perfect Watson–Crick complementarity to bases 2–8 of the miRNA (numbered from the 5′ end) and assigns a binding free energy of the miRNA–target duplex (11), given an internal database of miRNAs and UTR sequences.

1. Insert the name of the gene (e.g., *APC*) in the box marked as "Enter a human Entrez Gene symbol (e.g., 'LIN28')."
2. Push "Submit" to have the list of miRNAs and the binding sites.

3.6. Search of SNPs in the Target Sites

The predicted putative miRNA binding sites were screened for the presence of SNPs, by an extensive search in the SNP database (dbSNP; http://www.ncbi.nlm.nih.gov/SNP/). In our case, out of 140 genes, 109 genes did not bear SNPs within the predicted miRNA binding sites. Within the remaining 31 genes, we found 61 SNPs.

1. Within dbSNP insert the name of the gene (e.g., *APC*) in the input box and push "Go."
2. Follow the tab: "Human."
3. Click on any SNP.
4. In the section "GeneView" check the circle marked as "in gene region" and push "Go" to have a complete list of SNPs within the gene. Select only the SNPs in 3′UTR region.

3.7. Selection of SNPs

As criterion for polymorphisms selection, we excluded the SNPs having the minor allele frequency (MAF) lower than 0.24 in Caucasians. The selection was performed based on frequencies reported for Caucasians, as in future, case–control association studies will be carried out on this ethnic group.

1. As second criterion of selection, we excluded all the miRNAs predicted in steps 1–5 not expressed in colorectal cells. The list of miRNAs expressed in colorectal cell has been taken from a study by Cummins et al. (12). With this second selection criterion, only 38 SNPs remained in 38 different targets sites (Tables 2–4).

3.8. Calculation of the Gibbs Binding Free Energy with RNAcofold

For the selected SNPs, the algorithm *RNAcofold* (http://www.rna.tbi.univie.ac.at/cgi-bin/RNAcofold.cgi) was run to assess the Gibbs binding free energy (ΔG, expressed in kJ/mol), both for the common and the variant alleles. The algorithm *RNAcofold* computes the hybridization energy and base-pairing pattern of two RNA sequences (13). The difference of the free energies between the two alleles was computed as "variation of ΔG" (i.e., $\Delta\Delta G$).

1. Insert the sequence (direction 5′–3′) of the polymorphic predicted target site (NOT the whole 3′UTR) bearing the common allele in the input box marked as "Paste or type your first sequence here." Alternatively, one can upload a file containing the sequence of the polymorphic target site in FASTA format.
2. Insert the sequence of the miRNA (direction 5′–3′) in the input box marked as "Paste or type your second sequence here" (or upload a file containing the sequence of the miRNA in FASTA format).
3. Push "Proceed" to get the results.
4. Note the value of the free energy of the thermodynamic ensemble.
5. Come back to the home page.
6. Insert the sequence of the predicted polymorphic target site with the variant allele in the first input box.
7. Insert the sequence of the miRNA in the second input box.
8. Push "Proceed" to get the results.

Table 2
Selected SNPs with MAF≥0.25 in Caucasians with $|\Delta\Delta G|_{tot}$≥3.21 kJ/mol (upper tertile of the distribution of $|\Delta\Delta G|_{tot}$)

Genes	SNPs	Var.	Freq.	miRNA and $|\Delta\Delta G|$ (kJ/mol)	$|\Delta\Delta G|_{tot}$ (kJ/mol)
PKNOX1	rs2839629	G/A	0.41	miR-18a (1.82); miR-22 (3.97); miR-193b (3.33); miR-196b (0.60); miR-302c (0.37); miR-324-3p (3.48)	13.57
GALNS	rs1141390	T/C	0.25	miR-149 (0.10); miR-21 (1.34); miR-125b (0.04); miR-125a-5p (2.42); miR-199a-5p (1.79); miR-199b-5p (2.38); miR-205 (1.18); miR-222 (0.11); miR-330-5p (1.09); miR-345 (0.01); miR-362-5p (0.41); miR-421 (0.51); miR-498 (0.78); miR-502-5p (0.89)	13.05
ABCA1	rs4149338	C/T	0.28	miR-149 (1.19); miR-154 (0.66); miR-210 (1.93); miR-299-5p (0.67); miR-323 (2.29); miR-331-5p (1.25); miR-382 (1.81)	9.80
PKNOX1	rs378528	C/T	0.25	miR-184 (1.95); miR-185 (2.19); miR-210 (1.05); miR-214 (1.75); miR-320 (1.18); miR-423-5p (0.97)	45.43
ADARB2	rs904960	T/C	0.35	miR-25 (1.11); miR-32 (2.97); miR-363 (0.51); miR-375 (1.71)	6.30
KRAS	rs712	T/G	0.49	miR-185 (0.80); miR-193a-5p (1.47); miR-200c (1.65); miR-200b (2.23)	6.15
NUP210	rs1048650	C/T	0.47	miR-27a (0.75); miR-193a-5p (0.14); miR-210 (1.42); miR-330-5p (0.62); miR-342-5p (1.91)	4.84
ABCA1	rs4149339	C/T	0.28	miR-154 (0.01); miR-199a-5p (0.32); miR-199b-5p (0.79); miR-210 (1.20); miR-345 (2.44)	4.76
C1QR1	rs2749813	G/A	0.36	miR-188 (1.55); miR-373 (0.02); miR-518b (3.05)	4.62
ABCB11	rs496550	G/A	0.43	miR-363 (0.27); miR-214 (2.68); miR-24 (1.60)	4.55
ADARB2	rs1046914	A/G	0.35	miR-193b (0.13); miR-345 (3.31)	3.44
MMP2	rs7201	A/C	0.50	miR-361 (3.21); miR-373 (0.17)	3.38
MAP2K7	rs3679	C/T	0.35	miR-197 (3.21)	3.21

Table 3
Selected SNPs with MAF ≥ 0.25 in Caucasians and with $0.46 < |\Delta\Delta G|_{tot} \leq 3.21$ kJ/mol (second tertile of the distribution of $|\Delta\Delta G|_{tot}$)

Genes	SNPs	Var.	Freq.	miRNA and $\Delta\Delta G$ (kJ/mol)	$\lvert\Delta\Delta G\rvert_{tot}$ (kJ/mol)
ABCB11	rs473351	G/A	0.43	miR-188-5p (0.23); miR-205 (2.81)	3.04
CD109	rs3012518	G/A	0.44	miR-28-5p (0.03); miR-34c-5p (0.07); miR-148b (0.19); miR-148a (0.19); miR-184 (0.00); miR-185 (0.20); miR-197 (0.40); miR-214 (1.57); miR-299-3p (0.16)	2.81
ADARB2	rs1129227	T/C	0.32	miR-103 (1.53); miR-107 (1.21)	2.74
EYA4	rs3734279	C/T	0.38	miR-186 (2.53)	2.53
C1QR1	rs6076019	T/C	0.45	miR-326 (1.60); miR-145 (0.51)	2.11
SEMA3D	rs10261725	C/T	0.37	miR-517a (0.50); miR-517c (1.19)	1.69
KIAA0182	rs709805	G/A	0.35	miR-193b (1.43)	1.43
PRKD1	rs11984	A/G	0.48	miR-125b (0.018); miR-375 (0.47); miR-339 (0.53)	1.01
ABCB11	rs495714	A/G	0.40	miR-196a (0.09); miR-7 (0.57); miR-330-5p (0.09); miR-339-5p (0.14)	0.89
SMAD3	rs2278670	C/T	0.28	miR-320 (0.61); miR-423-5p (0.01)	0.63
KRAS	rs12245	T/A	0.44	miR-300-3p (0.08); miR-296 (0.55)	0.63
IGSF22	rs2289965	A/G	0.25	miR-142-3p (0.01); miR-324-5p (0.45)	0.46

Table 4
Selected SNPs with MAF≥0.25 in Caucasians and with 19.05 kJ/mol $|\Delta\Delta G|_{tot}$≤0.46 kJ/mol (third tertile of the distribution of $|\Delta\Delta G|_{tot}$)

Genes	SNPs	Var.	Freq.	miRNA and ΔΔ*G* (kJ/mol)	$\|\Delta\Delta G\|_{tot}$ (kJ/mol)
DSCAML1	rs2925768	G/A	0.38	miR-34a (0.07); miR-210 (0.35)	0.42
TGM3	rs214831	G/A	0.36	miR-373 (0.42)	0.42
TGM3	rs214832	C/T	0.43	miR-299-5p (0.42)	0.42
PKHD1	rs2784198	A/C	0.38	miR-557 (0.35)	0.35
CSMD3	rs5894075	G/T	0.32	let-7c (0.19); let-7b (0.15)	0.34
ATP11A	rs3742232	T/C	0.41	miR-139-3p (0.31)	0.31
NUP210	rs354476	C/T	0.48	miR-193a-5p (0.03); miR-323-5p (0.26)	0.29
ACSL5	rs8624	C/T	0.30	miR-103 (0.21)	0.21
KIAA0182	rs8571	G/A	0.37	miR-150 (0.00); miR-502 (0.10)	0.10
ADAMTSL3	rs2135551	A/G	0.29	miR-10a (0.15); miR-10b (0.16); miR-381 (0.19)	23.47
SMAD3	rs3743343	T/C	0.27	miR-31 (0.04)	0.04
EPHA3	rs7650466	C/T	0.29	miR-497 (0.01)	0.01
PKNOX1	rs2839628	C/G	0.26	miR-330-5p (0.00)	0.00

9. Note the value of the free energy of the thermodynamic ensemble.
10. Calculate the difference of the free energies (variation of ΔG, i.e., $\Delta\Delta$G) between the two alleles. More negative values correspond to higher stability of the RNA:miRNA duplex.

3.9. Calculation of the Sum of all Absolute Values of $\Delta\Delta G$

The typical result is quite heterogeneous for each predicted target. For some genes, the same sequence is predicted to be targeted by several miRNAs, whereas for others only one miRNA may be predicted. Given the fact that the predictions are based on probabilistic calculations, at least theoretically, one polymorphic target binding more miRNAs should weigh more than those targets binding one or few miRNAs. In fact, the more miRNAs are predicted to bind to a given target, the more likely it is that at least one of them truly binds the target. In order to account for these different weights, as parameter for predicting the biological impact of each polymorphism, the sum of the absolute values of $\Delta\Delta G$s should be used for each SNP (i.e., $|\Delta\Delta G|_{tot} = \Sigma|\Delta\Delta G|$).

3.10. Classification of the SNPs

In order to give a priority list of SNPs having an impact on miRNA binding, we ranked the values of $|\Delta\Delta G|_{tot}$ and classified the SNPs in three groups corresponding to three tertiles. The first tertile ($|\Delta\Delta G|_{tot} \geq 3.21$ kJ/mol) is composed of SNPs having a predicted high impact on the biology of the miRNA binding sites (reported in Table 2). The second tertile ($0.46 < |\Delta\Delta G|_{tot} \leq 3.21$ kJ/mol) is composed of SNPs with a predicted mild biological activity (Table 3), whereas within the third tertile ($|\Delta\Delta G_{tot}| \leq 0.46$ kJ/mol) belong SNPs with, likely, a weak activity (Table 4).

4. Notes

1. Here, we provided the basis for a reasoned algorithm-driven selection of SNPs. It is important to stress that all the polymorphisms predicted as having a strong impact on the miRNA biology could be good candidate to be tested for functional assays or in case–control association studies. The proposed approach could help to select only SNPs having a (putative) biological function, minimizing the workflow and the costs.

Acknowledgements

This work was, in part, funded by Associazione Italiana Ricerca Cancro (Italy) investigator grant 2008 and NCI (NIH-USA) Grant code R03CA115062, anno 2006–2008. Small grants in cancer epidemiology.

References

1. Bartel, D. P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function, *Cell 116*, 281–297.
2. Tomari, Y. and Zamore, P. D. (2005) Perspective: machines for RNAi, *Genes Dev. 19*, 517–529.
3. Sood, P., Krek, A., Zavolan, M., Macino, G., and Rajewsky, N. (2006) Cell-type-specific signatures of microRNAs on target mRNA expression, *Proc. Natl. Acad. Sci. U. S. A. 103*, 2746–2751.
4. Chen, K., Song, F., Calin, G., Wei, Q., Hao, X., and Zhang, W. (2008) Polymorphisms in microRNA targets: a gold mine for molecular epidemiology, *Carcinogenesis 29*, 1306–1311.
5. John, B., Enright, A. J., Aravin, A., Tuschl, T., Sander, C., and Marks, D. S. (2004) Human microRNA targets, *PLoS Biol. 2*, e363.
6. Wood, L. D., Parsons, D. W., Jones, S., Lin, J., Sjoblom, T., Leary, R. J., Shen, D., Boca, S. M., Barber, T., Ptak, J., Silliman, N., Szabo, S., Dezso, Z., Ustyanksky, V., Nikolskaya, T., Nikolsky, Y., Karchin, R., Wilson, P. A., Kaminker, J. S., Zhang, Z., Croshaw, R., Willis, J., Dawson, D., Shipitsin, M., Willson, J. K., Sukumar, S., Polyak, K., Park, B. H., Pethiyagoda, C. L., Pant, P. V., Ballinger, D. G., Sparks, A. B., Hartigan, J., Smith, D. R., Suh, E., Papadopoulos, N., Buckhaults, P., Markowitz, S. D., Parmigiani, G., Kinzler, K. W., Velculescu, V. E., and Vogelstein, B. (2007) The genomic landscapes of human breast and colorectal cancers, *Science 318*, 1108–1113.
7. Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A., and Enright, A. J. (2006) miRBase: microRNA sequences, targets and gene nomenclature, *Nucleic Acids Res. 34*, D140–D144.
8. Krek, A., Grun, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., MacMenamin, P., da Piedade, I, Gunsalus, K. C., Stoffel, M., and Rajewsky, N. (2005) Combinatorial microRNA target predictions, *Nat. Genet. 37*, 495–500.
9. Rusinov, V., Baev, V., Minkov, I. N., and Tabler, M. (2005) MicroInspector: a web tool for detection of miRNA binding sites in an RNA sequence, *Nucleic Acids Res. 33*, W696–W700.
10. Kiriakidou, M., Nelson, P. T., Kouranov, A., Fitziev, P., Bouyioukos, C., Mourelatos, Z., and Hatzigeorgiou, A. (2004) A combined computational–experimental approach predicts human microRNA targets, *Genes Dev. 18*, 1165–1178.
11. Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003) Prediction of mammalian microRNA targets, *Cell 115*, 787–798.
12. Cummins, J. M., He, Y., Leary, R. J., Pagliarini, R., Diaz, L. A., Jr., Sjoblom, T., Barad, O., Bentwich, Z., Szafranska, A. E., Labourier, E., Raymond, C. K., Roberts, B. S., Juhl, H., Kinzler, K. W., Vogelstein, B., and Velculescu, V. E. (2006) The colorectal microRNAome, *Proc. Natl. Acad. Sci. U. S. A. 103*, 3687–3692.
13. Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R., and Hofacker, I. L. (2008) The Vienna RNA websuite, *Nucleic Acids Res. 36*, W70–W74.
14. Landi, D., Gemignani, F., Naccarati, A., Pardini, B., Vodicka, P., Vodickova, L., Novotny, J., Forsti, A., Hemminki, K., Canzian, F., and Landi, S. (2008) Polymorphisms within micro-RNA-binding sites and risk of sporadic colorectal cancer, *Carcinogenesis 29*, 579–584.

Chapter 15

The Guideline of the Design and Validation of MiRNA Mimics

Zhiguo Wang

Abstract

The miRNA mimic technology (miR-Mimic) is an innovative approach for gene silencing. This approach is to generate nonnatural double-stranded miRNA-like RNA fragments. Such an RNA fragment is designed to have its 5′-end bearing a partially complementary motif to the selected sequence in the 3′UTR unique to the target gene. Once introduced into cells, this RNA fragment, mimicking an endogenous miRNA, can bind specifically to its target gene and produce posttranscriptional repression, more specifically translational inhibition, of the gene. Unlike endogenous miRNAs, miR-Mimics act in a gene-specific fashion. The miR-Mimic approach belongs to the "miRNA-targeting" and "miRNA-gain-of-function" strategy and is primarily used as an exogenous tool to study gene function by targeting mRNA through miRNA-like actions in mammalian cells. The technology was developed by my research group (Department of Medicine, Montreal Heart Institute, University of Montreal) in 2007 (Xiao, et al. J Cell Physiol 212:285–292, 2007; Xiao et al. Nat Cell Biol, in review).

Key words: miRNAs, Gene expression, miR-Mimic, miRNA interference (miRNAi), RNA interference (RNAi)

1. Introduction

RNA interference (RNAi) is a well-known strategy for gene silencing; this strategy takes the advantages of the capability of small double-stranded RNA molecules (siRNAs) to bind RNA-induced silencing complex (RISC) on one hand and to bind target genes (mRNAs) on the other hand (3–5). Through such dual interactions, siRNAs elicit powerful knockdown of gene expression by degrading their target mRNAs. Two key characteristics of the RNAi strategy are that the only target of RNAi is mRNAs and that the only outcome of RNAi is silencing of mRNAs. In other words, the RNAi strategy uses siRNAs to interfere directly with mRNAs (mostly protein-coding genes) to silence gene expression.

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_15, © Springer Science+Business Media, LLC 2011

Taking the same concept, I proposed a new concept of microRNA interference (miRNAi) (6). miRNAi manipulates the function, stability, biogenesis, or expression of miRNAs, and as such, it indirectly interferes with expression of protein-coding mRNAs. This concept is based on the thoughts outlined below. The fundamental mechanism of miRNA regulation of gene expression is miRNA–mRNA interaction or binding. A key to interfere with miRNA actions is to disrupt the miRNA–mRNA interaction. In order to achieve this aim, one can either manipulate miRNAs or mRNAs to alter the miRNA–mRNA interaction. For miRNAs, one can mimic miRNA actions either to enhance the miRNA–mRNA interaction or to inhibit miRNAs to break the miRNA–mRNA interaction. Additionally, one can also manipulate mRNA to interrupt the miRNA–mRNA interaction.

A general and unique feature of the miRNA–target RNA interaction is imperfect complementarity between the miRNA guide strand and its target mRNA. Hence, miRNA guide strands usually form bulge structures due to mismatches with its target sequence. The sequence specificity for target recognition by the guide miRNA strand is determined by nucleotides 2–8 of its 5′ region, referred to as the "seed site" (7, 8). This short seed site required for miRISC function raises the potential for a single miRNA to target multiple mRNAs. Indeed, it has been confirmed that unlike a nonnatural siRNA, which targets a particular gene, each single endogenous miRNA has the potential of regulating multiple protein-coding genes, as many as 1,000. On the other hand, each individual gene may be regulated by multiple miRNAs. This implies that the actions of miRNAs are not gene-specific, but sequence-specific, for they can act on any genes carrying motifs matching their seed sites. Thus, when aiming to silence a particular gene using a naturally occurring miRNA, one may actually knockdown a group of genes. This property of miRNAs creates a hurdle for exploiting miRNA function and targets.

The RNAi pathway of siRNA-directed mRNA cleavage and the miRNA-mediated translational repression pathway are genetically and biochemically distinct. In addition to different outcomes, the two pathways have differential requirements for Paz-Piwi domain (PPD) proteins in *Caenorhabditis elegans*. Translational repression by lin-4 and let-7 depends on alg-1 and alg-2 for miRNA processing and/or stability, yet these genes are not required for RNAi (9), whereas rde-1 is needed in RNAi but is not necessary for translational repression (10). Intriguingly, miRNAs capable of translational repression can be conferred the ability to cleave targets with 3′UTRs engineered to contain completely complementary sequences (11, 12). Conversely, functional siRNAs can repress translation of reporter genes containing multiple imperfect binding sites in their 3′UTRs (11, 13). The latter mimics the action of endogenous miRNAs. The findings from these elegant

experiments indicate a possibility of generating nonnatural artificial miRNAs simply by converting siRNAs into miRNAs.

To this end, my laboratory has developed a novel technology called microRNAs Mimics or miR-Mimics technology (see Notes 1–4). The miR-Mimic technology utilizes synthetic, nonnatural nucleic acids that can bind to the unique sequence of target genes (mRNAs) in a gene-specific manner and elicit posttranscriptional repressive effects as an endogenous miRNA does. That is, a miR-Mimic can act only on its particular target gene, but a native miRNA can act on any gene that carries its binding sequence. Additionally, the miR-Mimic technology produces artificial miRNAs that act by the miRNA mechanism. Thus, it will not lead to changes of expression levels of any endogenous miRNAs. The possible applications of miR-Mimics technology are summarized below:

1. To achieve gene silencing by miRNA mechanism in a gene-specific manner: As already mentioned in the previous chapter, miRNA action is not gene-specific. Thus, the gene-silencing action through SC-miRNA strategy is not gene-specific either. The miR-Mimic technology was developed to circumvent this limitation. This property of miR-Mimics can be advantageous over SC-miRNAs when specific-genes need to be knocked down, which happens in many situations.
2. To complement the loss of endogenous miRNAs under certain conditions: In some abnormal conditions, some miRNAs are downregulated in their expression, leading to aberrantly enhanced expression of some protein-coding genes causing diseased phenotypes. Replacement of these miRNAs may reverse this process. Alternatively, application of miR-Mimics targeting the disease-causing genes to prevent their upregulation may be an efficient maneuver to tackle the problem. For example, in the development of cancer, expression of oncogenes is overexpression, partially as a result of the downregulation of their regulating miRNAs. Under such a situation, use of miR-Mimics to keep down of the oncogenes may help to slow the tumorigenesis.
3. Examples of applications: We have examined the application of the technique to cardiac pacemaker genes HCN2 and HCN4 (1). Following the protocols described above, we first identified a stretch of sequence in the 3′UTR unique to the HCN2 (or HCN4) gene that is expectedly long enough for miRNA action. Based on the unique sequence, we designed a 22-nt miR-Mimic that at the 5′ end has eight nucleotides (nucleotides 2–8) and at the 3′-end has seven nucleotides, complementary to the HCN2 (or HCN4) sequence. These miR-Mimics produced substantial repression of HCN channel protein expression with concomitant depression of pacemaker

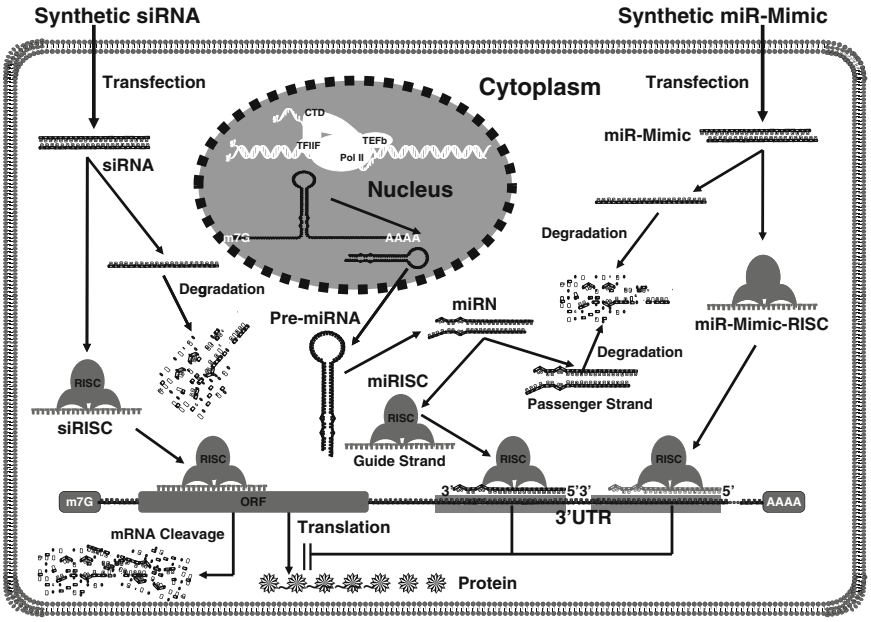


Fig. 1. Schematic presentation of actions of miRNA mimic (miR-Mimic) compared with the miRNA and small interference RNA (siRNA). Synthetic miR-Mimic and siRNA are introduced into the cells, and endogenous miRNA is synthesized by the cell. siRNAs bind to the coding region of target miRNAs and cause mRNA cleavage; miRNAs bind to 3′UTR of multiple target mRNAs and produce nongene-specific posttranscriptional repression to inhibit translation; miR-Mimics bind to 3′UTR of unique target mRNAs and produce gene-specific posttranscriptional repression to inhibit translation.

activities and reduction of beating rate of cultured neonatal myocytes but with minimal effects on their mRNA levels. These cardiac automaticity-targeting miR-Mimics are expected to act like heart rate-reducing agents that have been shown to be beneficial to cardiac function during infarction and to be able to suppress ectopic beats that can be elicited by abnormal automaticity. The results demonstrated a promise of utilizing the technology for gene-specific repression of expression at the protein level based on the principle of miRNA actions.

Additionally, we have also applied the miR-Mimic strategy to repress the oncoprotein Mdm2 and to suppress cancer cell growth in culture (2).

The principle of action of miR-Mimics is illustrated in Fig. 1.

2. Materials

2.1. Luciferase Assay

1. HeLa cells (ATCC, CCL-2.2) or NTera2 cells.
2. psiCheck 2 Vector.
3. Luminometer (Lumat, LB9507) or equivalent.

2.2. Western Blot

1. Setup buffer: 25 mM Tris (do not adjust pH), 190 mM glycine, 20% (v/v) methanol.
2. Transfer buffer: Setup buffer plus 0.05% (w/v) SDS. Store in the transfer apparatus at room temperature (with cooling during use).
3. Supported nitrocellulose membrane from Millipore, Bedford, MA, and 3MM Chr chromatography paper from Whatman, Maidstone, UK.
4. Tris-buffered saline with Tween (TBS-T): prepare 10× stock with 1.37 M NaCl, 27 mM KCl, 250 mM Tris–HCl, pH 7.4, 1% Tween-20. Dilute 100 mL with 900-mL water for use.
5. Blocking buffer: 5% (w/v) nonfat dry milk in TBS-T.
6. Primary antibody dilution buffer: TBS-T supplemented with 2% (w/v) fraction V bovine serum albumen (BSA).
7. Antidually phosphorylated MAPK (Sigma).
8. Secondary antibody: Antimouse IgG conjugated to horse radish peroxidase (Santa Cruz, Santa Cruz, CA).
9. Enhanced chemiluminescent (ECL) reagents from Kirkegaard and Perry (Gaithersburg, MD) and Bio-Max ML film (Kodak, Rochester, NY).

2.3. Real-Time RT-PCR

1. Ambion's mirVana miRNA Isolation Kit (Ambion).
2. The mirVana™ qRT-PCR miRNA Detection Kit (Ambion).
3. Thermocycler (96-well StepOnePlusTM system, A&B Applied Biosystems).
4. mirVana qRT-PCR Primer Sets (Ambion).

3. Methods

A key issue in creating functional miR-Mimics is to ensure the specificity of their interactions with their target mRNAs and to direct each interaction to discrete downstream consequences. Some principles of this interaction have emerged based on several key studies (7, 8, 14). When designing a miR-Mimic, several points should be seriously considered:

1. The key for effective miRNA action is the 5′ portion complementarity of a miRNA to its target mRNA. Complementarity of seven or more bases to residues 2–8 from the 5′-end of a miRNA, the so-called "seed site," is sufficient to confer regulation, even if the target 3′UTR contains only a single site. Sites with weaker "seed site" complementarity require compensatory pairing to the 3′ portion of the miRNA in order to

confer regulation, and extensive pairing to the 3′ portion of the miRNA is not sufficient to confer regulation on its own without a minimal element of 5′ complementarity.

2. The 3′ portion of the miRNA can contribute to efficiency of repression, as a modulator of suppression (8, 15, 16). Thus, in addition to seed-site complementarity, appropriate 3′-end base-pairing can strengthen the effectiveness of miRNAs.
3. If efficient endonuclease cleavage is not required, then base-pairing at the site of cleavage, between bases 10 and 11, should be avoided (17–19).
4. To achieve efficient translational repression, multiple binding sites in the 3′UTR of a target gene may be required, and for direct repression by cleavage of target mRNA, only one binding site is generally sufficient (8, 11, 13, 20).
5. Under certain circumstances, some specific conformations, such as a bulge, in the miRNA:mRNA duplex may help in function as effective repression of the Lin-14 mRNA by the Lin-4 miRNA appears to require a bulge in the miRNA:mRNA duplex (21). The exact mechanisms are unknown, perhaps to allow the recruitment of additional RNA-binding proteins in specific contexts.

The detailed protocols are described below with the miR-Mimics for a cardiac pacemaker channel gene HCN2 and an oncogene Mdm2 as examples.

3.1. Designing miR-Mimics

1. Defining a unique sequence in the target gene: A fundamental requirement to be satisfied is that the 3′UTR of the target gene must contain a unique sequence distinct from other genes to elicit gene-specific action. Similar to designing a siRNA, the first step to design a miRNA mimic is to identify a stretch of sequence in the 3′UTR unique to the gene of interest (target mRNA). But, unlike the full complementarity between a siRNA and its target in any regions of the gene, a miRNA mimic partially base-pairs with the target sequence in the 3′UTR. The length of the sequence should be long enough for miRNA action, which is at least 8 nts and ideally >14 nts.
2. Based on the unique sequence, design 22-nt oligonucleotides that at the 5′ portion have 8 nts (nucleotide positions 1–8 from 5′-end) fully complementary to the target mRNA.
3. To ensure the specificity of binding to the target mRNA, the oligonucleotides should have at least additional 5–6 nts complementary to the target mRNA at the 3′ portion. The base-pairing in this region may not necessarily be continuous and can be grouped into 3-nts clusters.

4. Add an "AU" to the 3′-end of the fragment. As such, a 22-nt miR-Mimic should have at least 14 complementary nucleotides to the target gene plus an "AU" overhang.
5. Design an antisense strand fully complementary to the fragment.
6. Both of the two oligonucleotide fragments need to be artificially synthesized. Many companies provide excellent services for nucleic acids research, such as Integrated DNA Technologies, Ambion, Invitrogen, etc.
7. Anneal the sense and antisense strands of the oligonucleotides to form a duplex miR-Mimic. First mix the two synthetic oligonucleotides at an equal molar concentration in an Eppendorf tube, and then, incubate the sample in a heat block at 95°C for 5 min, and gradually cool the sample to room temperature (22–23°C).
8. Store the miR-Mimic construct at −80°C for future use.
9. Construct a negative control miR-Mimic (NC miR-Mimic) for verifying the effects and specificity of the effects of the miR-Mimic. This NC miR-Mimic should be designed based on the sequence of the miR-Mimic, simply by modifying the miR-Mimic sequence to contain ~5-nt mismatches at the 5′-end "seed site." Such modified or mutated miR-Mimic is expected to lose the ability to bind to the target mRNA with sufficient affinity and is thus rendered incapable of eliciting repressive action. To form an NC miR-Mimic, the annealing procedures described above need to be repeated.
10. Synthesize an anti-miRNA inhibitor (AMO) against the miR-Mimic as an additional negative control. An AMO is designed to be an exact antisense to its target miR-Mimic. An AMO is a single-stranded oligonucleotide (ON) or oligodeoxynucleotide (ODN) fragment.
11. Chemical modification should be done for miR-Mimic to improve its nuclease stability. Many 2′-sugar modifications have been shown to confer this property, including 2′-O-Me, 2′-F, 2′-deoxy, 2′-MOE, and LNA, and have been evaluated and are tolerated to varying degrees in either the sense or antisense strands (22–31). Introduction of phosphorothioate internucleotidic linkages, which can promote plasma protein binding and delay renal clearance of single-stranded oligonucleotides, in addition to conferring nuclease stability (32, 33), is tolerated by siRNAs in cell culture (22, 23, 29, 31), and a beneficial effect on in vivo delivery has been reported (34). The first demonstration of in vivo systemic delivery to the liver was achieved through high-pressure intravenous injection, or hydrodynamic delivery of unmodified siRNAs (35).

Alternative strategies to promote cellular uptake and prevent renal clearance after systemic delivery have involved conjugation of the siRNA to moieties such as cholesterol (35), or formulation in liposomes (36–38), although uptake and activity have only been demonstrated in the liver and lungs. The type of modification should be determined based on your particular need.

3.2. Validating miR-Mimics

The functional activities of miR-Mimics must be verified with several approaches including luciferase reporter gene assay, Western blot analysis, qRT-PCR quantification, and functional assays.

3.2.1. Luciferase Reporter Activity Assay

The dual luciferase system offers a relatively simple approach to assess miRNA inhibition. The psiCheck2 vector system from Promega encodes both *Photinus pyralis* and *Renilla reniformis* luciferase genes on a single plasmid with a multiple cloning site in the 3′UTR of *Renilla* luciferase for insertion of target 3′UTRs or synthetic oligonucleotides encoding the miRNA target sites. Alternatively, combination of the pMIR-REPORT™ luciferase miRNA expression reporter vector (Ambion) and PRL-TK (TK-driven *Renilla* luciferase expression vector) for dual luciferase activity assay has also been widely used for miRNA research. Below, the procedures using the psiCheck2 vector are described:

1. Cloning a miRNA target site Into psiCheck 2 Vector: There are two approaches for designing a miR-Mimic target site. First, identify the authentic binding site in the 3′UTR of the target gene and subclone the full length of the 3′UTR or the motif containing the binding site of interest into the psiCheck2 vector (say, at the XhoI and NotI sites). Second, design a standard miR-Mimic binding site with an oligonucleotide fragment containing at the 3′end 1–9 nts exactly complementary to the miR-Mimic and its reverse complement, and subclone the oligo fragment into the psiCheck2 vector. Add appropriate overhanging bases corresponding to the restriction enzyme(s) used to digest the psiCheck2 vector such that the mature miRNA's reverse complement is in the 5′–3′ orientation.
2. Transfecting the psiCheck2 vector into mammalian cells: Starve cells in serum-free medium for 24 h. Then, transfect cells (1×10^5/well) with 0.1 μg psiCheck2 vector for control and psiCheck2 vector (0.1 μg) + miRNA of interest (1 μg) for test with lipofectamine 2000 (Invitrogen), according to the manufacturer's instructions. For control experiments, cells can also be transfected with the antisense to the miRNA (AMO) alone, miRNA + AMO, or mismatched miR-Mimic.
3. Performing luciferase activity assay: Twenty-four hours (mammalian cells) after transfection of the constructs, cells can be

measured for luciferase activity with a dual luciferase reporter assay kit (Promega) on a luminometer (e.g., Lumat LB9507), according the manufacturer's instructions.

4. Assessing miRNA Inhibition: The relative levels of *Renilla* luciferase should reveal first, whether the miR-Mimic has the miRNA-like functional activity, and second, whether the AMO is able to block this activity. In control cells pretreated with AMO or mismatched miR-Mimics, no changes of *Renilla* luciferase activities should be seen. In test cells treated with the miR-Mimic alone, *Renilla* luciferase activities should be reduced and when cotransfected with its AMO, reduction of luciferase activities should be prevented.

3.2.2. Verification of Downregulation of Target Protein

miR-Mimics, like natural and endogenous miRNAs, silence target genes posttranscriptionally when introduced into cells. Downregulation of target genes at the protein level is a characteristic of miR-Mimic actions, but this needs to be confirmed for each miR-Mimic under study. Western blot analysis is a mostly commonly used method for verification of downregulation of cognate target protein, though other techniques like immunostaining are also handy for the purpose (1, 39–42). Western blotting procedures are just those routines you can find nowadays in any of the laboratories involved in biomedical research. Just do not forget to distinguish among the membrane protein sample, cytosolic protein sample, and nuclear protein sample for your particular need. In addition to Western blot and immunostaining, functional assays should also be employed when needed, such as enzyme activity assay (43), cell growth and death, patch-clamp recordings for ion channels as target genes (1, 39–42).

Extract protein samples from target tissues, employing the procedures essentially the same as described in detail elsewhere (1, 41). Determine protein content with Bio-Rad Protein Assay Kit (Bio-Rad, Mississauga, ON, Canada), using bovine serum albumin as the standard. Fractionate protein sample (~30 μg) by SDS–PAGE (7.5–10% polyacrylamide gels) and then transfer to PVDF membrane (Millipore, Bedford, MA). Incubate the sample overnight at 4°C with primary antibodies. Detect bound antibodies using the chemiluminescent substrate (Western Blot Chemiluminescence Reagent Plus, NEN Life Science Products, Boston, MA). Quantify Western blot bands using QuantityOne software by measuring the band intensity (Area × OD) for each group and normalizing to GAPDH (anti-GAPDH antibody from Research Diagnostics Inc.) as an internal control. Express the final results as fold changes by normalizing the data to the control values.

3.2.3. Quantification of mRNA and miRNA

The purposes of RNA quantification described in this section are twofolds: one for verifying successful transfection of the miR-Mimic introduced into cells and the other for measuring the possible

alterations of the cognate target mRNA produced by the miR-Mimic. The traditional method of validation has been northern blotting analysis, an effective method to visualize both the precursor and the mature miRNA. The sensitivity of northern blotting, however, is not sufficiently high for detecting some low-abundance mRNAs. As an alternative to northern blotting, real-time quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR) has been extremely effective in these validation studies (1, 2, 39–42). The methods can be used for quantifying both the miR-Mimics and target mRNAs. Another type of validation is to visualize the expression of the miRNA to locate the cellular and subcellular distribution using in situ hybridization.

Perform qRT-PCR on a thermocycler for 40 cycles. Determine the appropriate cycle threshold (Ct) using the automatic baseline determination feature. Conduct dissociation analysis (melt-curve) on the reactions to identify the characteristic peak associated with primer–dimers to separate from the single prominent peak representing the successful PCR amplification of miRNAs. Calculate fold variations in expression of miRNAs between RNA samples.

3.2.4. Functional Assays

Functional assays for investigating the role of miR-Mimics, in particular pathophysiological processes such as tumorigenesis, should be conducted. For the assays related to cancer research, effects of miR-Mimics on cell proliferation, cell cycle, cell death, cell invasion, etc., either in cell lines or in nude mice, may be evaluated (1, 43–45).

4. Notes

As an alternative to synthetic canonical miRNA technology (exogenously applied canonical miRNA), miR-Mimic strategy offers a couple of advantages that supplement the limitations of the former:

1. miR-Mimics can be designed to direct either solely translational process repression by sole "seed site" base-paring or both translational process repression and target mRNA degradation by larger degree of complementarity.
2. miR-Mimic strategy offers gene-specific targeting through miRNA mechanisms of action. This unique property makes it differ from miRNA and siRNA.
3. For a target gene, one can create as many as miR-Mimics at will to enhance the posttranscriptional repression, as long as the target sequence contains sufficient gene-specific stretches for designing that many miR-Mimics.

It should be noted that it is unclear yet as to what advantages the miR-Mimic technology may have over the siRNA approach. And it should also be recognized that finding gene-specific stretches of sequences restricted in the 3′UTR of protein-coding genes for designing miR-Mimics may represent a difficult task and in some cases may even be unrealistic.

Acknowledgments

This work was supported in part by the Canadian Institute of Health Research, Heart and Stroke Foundation of Quebec, and Fonds de la Recherche de l'Institut de Cardiologie de Montreal. Dr. Z. Wang is a Changjiang Scholar Professor of the Ministry of Education of China and a Longjiang Scholar Professor of Heilongjiang, China. The authors thank XiaoFan Yang for her excellent technical supports.

References

1. Xiao J, Yang B, Lin H, Lu Y, Luo X, Wang Z (2007) Novel approaches for gene-specific interference via manipulating actions of microRNAs: examination on the pacemaker channel genes HCN2 and HCN4. J Cell Physiol 212:285–292.
2. Xiao J, Lin H, Luo X, Chen G, Wang Z (2009) *microRNA-605* joins p53:Mdm2 network to form a positive feedback loop in cell-fate decision. EMBO J (accepted).
3. Xia H, Mao Q, Paulson HL, Davidson BL (2002) siRNA-mediated gene silencing in vitro and in vivo. Nat Biotechnol 20:1006–1010.
4. Golden DE, Gerbasi VR, Sontheimer EJ (2008) An inside job for siRNAs. Mol Cell 31:309–312.
5. Pushparaj PN, Aarthi JJ, Manikandan J, Kumar SD (2008) siRNA, miRNA, and shRNA: in vivo applications. J Dent Res 87:992–1003.
6. Wang Z, Luo X, Lu Y, Yang B (2008) miRNAs at the heart of the matter. J Mol Med 86:771–783.
7. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. Cell 115:787–798.
8. Doench JG, Sharp PA (2004) Specificity of microRNA target selection in translational repression. Genes Dev 18:504–511.
9. Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha I, Baillie DL, Fire A, Ruvkun G, Mello CC (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. Cell 106:23–34.
10. Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC (1999) The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. Cell 99:123–132.
11. Doench JG, Petersen CP, Sharp PA (2003) siRNAs can function as miRNAs. Genes Dev 17:438–442.
12. Hutvágner G, Zamore PD (2002) A microRNA in a multiple-turnover RNAi enzyme complex. Science 297:2056–2060.
13. Zeng Y, Yi R, Cullen BR (2003) MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. Proc Natl Acad Sci USA 100:9779–9784.
14. Brennecke J, Stark A, Russell RB, Cohen SM (2005) Principles of microRNA–target recognition. PLoS Biol 3:404–418.
15. Kiriakidou M, Tan GS, Lamprinaki S, De Planell-Saguer M, Nelson PT, Mourelatos Z (2007) An mRNA m(7)G cap binding-like motif within human Ago2 represses translation. Cell 129:1141–1151.
16. Kloosterman WP, Wienholds E, Ketting RF, Plasterk RH (2004) Substrate requirements for let-7 function in the developing zebrafish embryo. Nucleic Acids Res 32:6284–6291.

17. Elbashir SM, Lendeckel W, Tuschl T (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev 15:188–200.
18. Haley B, Zamore PD (2004) Kinetic analysis of the RNAi enzyme complex. Nat Struct Mol Biol 11:599–606.
19. Martinez J, Tuschl T (2004) RISC is a 5′ phosphomonoester-producing RNA endonuclease. Genes Dev 18:975–980.
20. Kiriakidou M, Nelson PT, Kouranov A, Fitziev P, Bouyioukos C, Mourelatos Z, Hatzigeorgiou A (2004) A combined computational–experimental approach predicts human microRNA targets. Genes Dev 18:1165–1178.
21. Ha I, Wightman B, Ruvkun G (1996) A bulged lin-4/lin-14 RNA duplex is sufficient for *Caenorhabditis elegans* lin-14 temporal gradient formation. Genes Dev 10:3041–3050.
22. Chiu YL, Rana TM (2003) siRNA function in RNAi: a chemical modification analysis. RNA 9:1034–1048.
23. Czauderna F, Fechtner M, Dames S, Aygun H, Klippel A, Pronk GJ, Giese K, Kaufmann J (2003) Structural variations and stabilising modifications of synthetic siRNAs in mammalian cells. Nucleic Acids Res 31:2705–2716.
24. Prakash TP, Allerson CR, Dande P, Vickers TA, Sioufi N, Jarres R, Baker BF, Swayze EE, Griffey RH, Bhat B (2005) Positional effect of chemical modifications on short interference RNA activity in mammalian cells. J Med Chem 48:4247–4253.
25. Kraynack BA, Baker BF (2006) Small interfering RNAs containing full 2′-Omethylribonucleotide-modified sense strands display Argonaute2/eIF2C2-dependent activity. RNA 12:163–176.
26. Allerson CR, Sioufi N, Jarres R, Prakash TP, Naik N, Berdeja A, Wanders L, Griffey RH, Swayze EE, Bhat B (2005) Fully 2′-modified oligonucleotide duplexes with improved in vitro potency and stability compared to unmodified small interfering RNA. J Med Chem 48:901–904.
27. Layzer JM, McCaffrey AP, Tanner AK, Huang Z, Kay MA, Sullenger BA (2004) In vivo activity of nuclease-resistant siRNAs. RNA 10:766–771.
28. Harborth J, Elbashir SM, Vandenburgh K, Manninga H, Scaringe SA, Weber K, Tuschl T (2003) Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing. Antisense Nucleic Acid Drug Dev 13:83–105.
29. Braasch DA, Jensen S, Liu Y, Kaur K, Arar K, White MA, Corey DR (2003) RNA interference in mammalian cells by chemically-modified RNA. Biochemistry 42:7967–7975.
30. Morrissey DV, Blanchard K, Shaw L, Jensen K, Lockridge JA, Dickinson B, McSwiggen JA, Vargeese C, Bowman K, Shaffer CS, Polisky BA, Zinnen S (2005) Activity of stabilized short interfering RNA in a mouse model of hepatitis B virus replication. Hepatology 41:1349–1356.
31. Amarzguioui M, Holen T, Babaie E, Prydz H (2003) Tolerance for mutations and chemical modifications in a siRNA. Nucleic Acids Res 31:589–595.
32. Altmann KH, Dean NM, Fabbro D, Freier SM, Geiger T, Haener R, Huesken D, Martin P, Monia BP, Muller M, Natt F, Nicklin P, Phillips J, Pieles U, Sasmor H, Moser H (1996) Second generation of antisense oligonucleotides. From nuclease resistance to biological efficacy in animals. Chimia 50:168–176.
33. Hoke GD, Draper K, Freier SM, Gonzalez C, Driver VB, Zounes MC, Ecker DJ (1991) Effects of phosphorothioate capping on antisense oligonucleotide stability, hybridization and antiviral efficacy versus herpes simplex virus infection. Nucleic Acids Res 19:5743–5748.
34. Braasch DA, Paroo Z, Constantinescu A, Ren G, Oz OK, Mason RP, Corey DR (2004) Biodistribution of phosphodiester and phosphorothioate siRNA. Bioorg Med Chem Lett 14:1139–1143.
35. McCaffrey AP, Meuse L, Pham TT, Conklin DS, Hannon GJ, Kay MA (2002) RNA interference in adult mice. Nature 418:38–39.
36. Zimmermann TS, Lee AC, Akinc A, Bramlage B, Bumcrot D, Fedoruk MN, Harborth J, Heyes JA, Jeffs LB, John M, Judge AD, Lam K, McClintock K, Nechev LV, Palmer LR, Racie T, Rohl I, Seiffert S, Shanmugam S, Sood V, Soutschek J, Toudjarska I, Wheat AJ, Yaworski E, Zedalis W, Koteliansky V, Manoharan M, Vornlocher HP, MacLachlan I (2006) RNAi-mediated gene silencing in non-human primates. Nature 441:111–114.
37. Morrissey DV, Lockridge J.A, Shaw L, Blanchard K, Jensen K, Breen W, Hartsough K, Machemer L, Radka S, Jadhav V, Vaish N, Zinnen S, Vargeese C, Bowman K, Shaffer CS, Jeffs LB, Judge A, MacLachlan I, Polisky B (2005) Potent and persistent in vivo anti-HBV activity of chemically modified siRNAs. Nat Biotechnol 23:1002–1007.
38. Ge Q, Filip L, Bai A, Nguyen T, Eisen HN, Chen J (2004) Inhibition of influenza virus production in virus-infected mice by RNA interference. Proc Natl Acad Sci USA 101:8676–8681.

39. Yang B, Lin H, Xiao J, Luo X, Li B, Lu Y, Wang H, Wang Z (2007) The muscle-specific microRNA miR-1 causes cardiac arrhythmias by targeting GJA1 and KCNJ2 genes. Nat Med 13:486–491.
40. Xiao J, Luo X, Lin H, Xu C, Gao H, Wang H, Yang B, Wang Z (2007) MicroRNA miR-133 represses HERG K^+ channel expression contributing to QT prolongation in diabetic hearts. J Biol Chem 282:12363–12367.
41. Luo X, Lin H, Lu Y, Li B, Xiao J, Yang B, Wang Z (2007) Transcriptional activation by stimulating protein 1 and post-transcriptional repression by muscle-specific microRNAs of I_{Ks}-encoding genes and potential implications in regional heterogeneity of their expressions. J Cell Physiol 212:358–367.
42. Luo X, Lin H, Pan Z, Xiao J, Zhang Y, Lu Y, Yang B, Wang Z (2008) Overexpression of Sp1 and downregulation of miR-1/miR-133 activates re-expression of pacemaker channel genes HCN2 and HCN4 in hypertrophic heart. J Biol Chem 283:20045–20052.
43. Xu C, Lu Y, Lin H, Xiao J, Wang H, Luo X, Li B, Yang B, Wang Z (2007) The muscle-specific microRNAs miR-1 and miR-133 produce opposing effects on apoptosis via targeting HSP60/HSP70 and caspase-9 in cardiomyocytes. J Cell Sci 120:3045–3052.
44. Lu Y, Xiao J, Lin H, Bai Y, Luo X, Wang Z, Yang B (2009) Complex antisense inhibitors offer a superior approach for microRNA research and therapy. Nucleic Acids Res 37:e24–e33.
45. Xiao L, Xiao J, Luo X, Lin H, Wang Z, Nattel S (2008) Feedback remodeling of cardiac potassium current expression. A novel potential mechanism for control of repolarization reserve. Circulation 118:983–992.

Chapter 16

Analysis of Targets and Functions Coregulated by MicroRNAs

Shu-Jen Chen and Hua-Chien Chen

Abstract

MicroRNAs (miRNAs) are small, nonprotein-coding RNAs that negatively regulate their mRNA target genes in a sequence-specific manner. While their specific impact on biological processes and cellular functions remain largely unknown, dysregulated miRNAs have been implicated in numerous diseases, including cancers. Several large-scale profiling studies using tissue samples have revealed a consistent yet complex pattern of miRNA dysregulation in human cancer. In particular, global alteration of multiple miRNAs is common in human tumorigenesis. Systemic analysis of pathways and functions coregulated by these dysregulated miRNAs is a crucial step to understand the role of miRNAs in tumorigenesis. This chapter provides an integrated pipeline to identify cellular pathways and functions specifically regulated by multiple dysregulated miRNAs. Protocols described in this chapter include (1) miRNA target prediction using TargetScan algorithm, (2) data compilation to identify target genes coregulated by multiple miRNAs, and (3) pathway enrichment analysis of coregulated targets using MetaCore pathway and network database.

Key words: MiRNA, Tumorigenesis, MiRNA targets, TargetScan, MetaCore pathway

1. Introduction

MicroRNAs (miRNAs) are small, nonprotein-coding RNA regulators involved in numerous biological and developmental processes (1, 2). MiRNAs bind to the 3′UTR of their target genes to induce translational blockade or transcript degradation. Cumulative evidence suggests that dysregulation of miRNA plays an important role in many human disorders, including cancer. Approximately 50% of human miRNAs are encoded in genome regions frequently altered in cancer (3, 4). In vitro and in vivo studies also demonstrated that many miRNAs may behave as tumor suppressors or oncogenes (5, 6). Recently, multiple large-scale profiling studies using tissue samples from different tumor types confirmed that global alteration of miRNA expression pattern is a common phenomenon associated with human cancer (7–14) (Table 1).

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_16, © Springer Science+Business Media, LLC 2011

Table 1
MiRNA differentially expressed in human cancer

Tumor type	Samples	Criteria	Up	Down	Reference
Breast cancer	76T, 10N	$p<0.05$	17	12	Iorio et al. (11)
Pancreatic ductal adenocarcinoma	8T, 5N	2-Fold	32	41	Szafranska et al. (12)
Pancreatic ductal adenocarcinoma	65T, 42N	2-Fold	21	4	Bloomston et al. (7)
Hepatocellular carcinoma	52T	$p<0.05$, 2-Fold	13	19	Varnholt et al. (13)
Pituitary adenoma	32T, 6N	$p<0.05$	7	23	Bottoni et al. (8)
Brain tumor (GBM)	9T, 9N	$p<0.05$, 1.8-fold	9	4	Ciafre et al. (10)
Lung cancer	104T, 104N	$p<0.01$	15	28	Yanaihara et al. (14)
Nasopharyngeal carcinoma	13T, 9N	FDR < 0.05, 3-fold	11	24	Chen et al. (9)

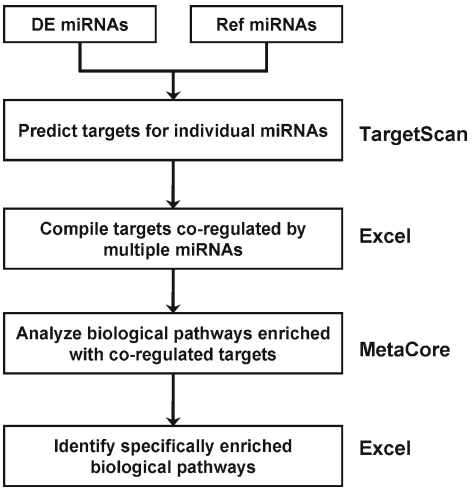


Fig. 1. Steps and tools involved in predicting targets and functions specifically enriched by co-regulated miRNAs.

Understanding the impact of global miRNA dysregulation on cellular functions and pathways is the first step toward uncovering the role of miRNA in tumor development (15). This chapter describes an integrated approach to predict target genes coregulated by multiple miRNAs and methods to analyze biological processes specifically enriched with coregulated miRNA targets (Fig. 1).

Predicting targets for individual miRNAs is a crucial step in the analysis framework. Various algorithms have been developed

in recent years to predict the target genes regulated by miRNAs. Several target prediction tools are publically available, either in a web server format or as a stand-alone application. One of the most popular target prediction tools is TargetScan, a web-based service developed and maintained by the Bartel Laboratory at the Whitehead Institute, which provides precompiled target genes for all human miRNAs (16). TargetScan algorithm predicts potential targets of miRNAs by searching for the presence of conserved 8mer and 7mer sites that match the seed region of each miRNA and uses context score to estimate site efficacy for each miRNA:mRNA pair (16, 17). Context score was calculated based on general features of site context, including site type (8mer, 7mer-M8, 7mer-A1), local AU composition, additional 3′ base-pairing to miRNA nt 13–16, and site position within the 3′ UTR. Data from multiple in vivo and in vitro studies were used to model the correspondence between context score and miRNA efficacy. Several recent studies using microarray and proteomics data have confirmed that many of the TargetScan predicted miRNA–mRNA interactions are correct (18–20).

Recent microarray and proteomic approaches have revealed that although individual miRNA can regulate hundreds of target genes, the degree of repression on individual target is usually very mild (18–20). To exert its biological function, miRNA tends to regulate genes located within the same pathway. For example, miR-16 family has been shown to trigger cell cycle arrest by silencing multiple cell cycle-regulated genes (21). Similarly, miR-17-5p has been shown to coordinately suppress more than 20 genes involved in the G1/S transition to regulate cell cycle progression (22). Thus, bioinformatic analysis of pathways enriched by predicted miRNA targets is a powerful method to interrogate potential biological functions of miRNA (9, 15).

Although pathway enrichment analysis has been widely applied to assess the biological function of large gene sets generated from high-throughput expression studies, no unified method or gold standard has emerged as the enrichment tools are still in an actively growing and improving stage. Publically available tools for pathway enrichment analysis include the KEGG pathway database (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.jp/kegg), BioCarta database (http://www.biocarta.com), Panther classification system (http://www.pantherdb.org/pathway), and DAVID gene functional classification tools (http://www.david.abcc.ncifcrf.gov). Integrated suites for pathway enrichment analysis provided by commercial suppliers include the PathwayStudio (Ariadne Genomics, Inc., http://www.ariadnegenomics.com), IPA (Ingenuity Systems, Inc., http://www.ingenuity.com) and MetaCore (GeneGo, Inc., http://www.genego.com). This chapter describes protocols for pathway enrichment analysis using the MetaCore database.

2. Materials

2.1. Select Differentially Expressed miRNAs

Currently, several high-throughput methods have been used to profile miRNA expression in tissues and cultured cells. These technologies have been widely used to profile miRNA expression in various cancer types and identified numerous miRNAs dysregulated in cancer. Differentially expressed miRNAs (DE-miRs) between normal and tumor were selected based on the fold change in expression level, or p values of statistical analysis, or both. List of DE-miRs can be generated through in-house profiling study or through mining of published data. In this chapter, we use the 11 miRNAs up-regulated in nasopharyngeal carcinoma previously published by our laboratory as an example (9) (Table 2). The DE-miRs were selected based on the criteria of fold change ≥3 and false discovery rate <0.05.

2.2. Select Reference miRNAs

In order to calculate the specificity of the miRNA coregulated pathways, a set of reference miRNAs (Ref-miRs) should be selected and analyzed in parallel to estimate the pathways enriched by chance. Ideally, Ref-miRs should be randomly selected from a pool of miRNAs whose expression levels are not significantly altered (e.g., less than 1.5-fold difference between normal and tumor) in the same expression study. The number of Ref-miRs should be the same as the DE-miRs, and their seed sequences should not overlap with the seed sequences of the DE-miRs. In addition, the number of predicted targets for Ref-miRs should be similar to the number of targets predicted for DE-miRs (Table 2).

2.3. Software

1. Target prediction: TargetScan (http://www.targetscan.org).
2. Internet browser: Internet Explorer or FireFox.
3. Spreadsheet program: Microsoft Excel (2003 and 2007 versions) or OpenOffice Calc.
4. Pathway enrichment analysis tool: MetaCore (GeneGO, St. Joseph, MI, USA).

3. Methods

Protocols described in this section include (1) predicting targets for individual miRNA using the TargetScan web server, (2) compiling coregulated miRNA targets using Excel pivot table, (3) analyzing pathways enriched with coregulated miRNA targets using GeneGO MetaCore, and (4) identifying specifically enriched biological pathways using Excel. The overall scheme for data analysis is depicted in Fig. 1.

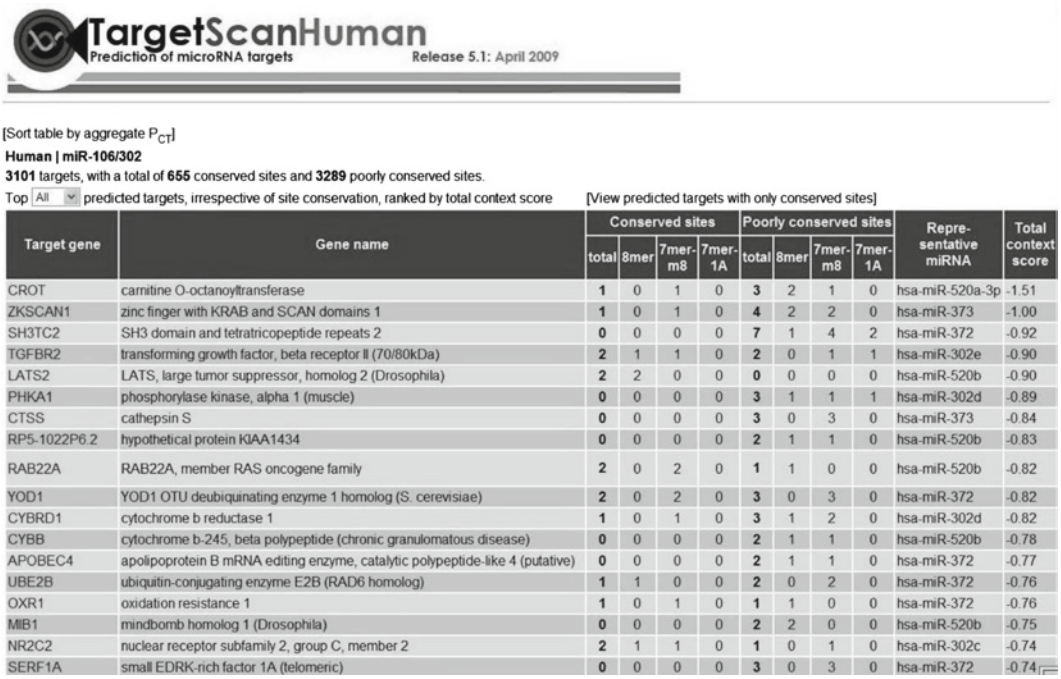
Table 2
MicroRNAs selected for target prediction

	Seed sequence	Predicted target		Seed sequence	Predicted target
DE-miRs			Ref-miRs		
miR-106a	AAAGUGC	1,593	miR-124	AAGGCAC	1,209
miR-138	GCUGGUG	750	miR-129-3p	AGCCCUU	569
miR-142-3p	GUAGUGU	472	miR-146a	GAGAACU	788
miR-155	UAAUGCU	991	miR-193b	ACUGGCC	530
miR-15b*	GAAUCAU	884	miR-27b	UCACAGU	1,500
miR-17	AAAGUGC	1,593	miR-299-5p	GGUUUAC	677
miR-17*	AAAGUGC	995	miR-324-5p	GCAUCCC	436
miR-18a	AAAGUGC	527	miR-483-3p	CACUCCU	552
miR-196b	AGGUAGU	474	miR-500	AAUCCUU	991
miR-205	CCUUCAU	870	miR-516a-5p	UCUCGAG	86
miR-25	GGCGGAG	117	miR-654-3p	AUGUCUG	873

3.1. Predicting Targets for Individual miRNAs Using TargetScan

TargetScan is a web-based server that provides precompiled targets for human miRNAs. The current version (Release 5.1: April 2009) contains predicted targets for 677 mature human miRNAs in miRBase (Version 12.0). List of potential targets for each miRNA with gene symbol, gene name, numbers of conserved and nonconserved sites, and sum of context score can be obtained using the following steps:

1. Go to the TargetScan Human website at http://www.targetscan.org/vert_50/ (release 5.1).
2. Enter the name of miRNA (e.g., hsa-miR-106a) or select the desired miRNA from the dropdown list, and click the "SUBMIT" button. Predicted targets for the miRNA will be displayed in a tabular format (see Note 1).
3. By default, TargetScan displays genes with conserved mRNA::-miRNA sites. To download all predicted targets, click the link labeled "View top predicted targets, irrespective of site conservation." TargetScan will display the top "100" predicted targets.
4. To retrieve all predicted target, use the dropdown list and choose display "all" predicted targets. TargetScan will display all predicted targets in a table, irrespective of site conservation (Fig. 2).



TargetScanHuman
Prediction of microRNA targets Release 5.1: April 2009

[Sort table by aggregate P_{CT}]
Human | miR-106/302
3101 targets, with a total of **655** conserved sites and **3289** poorly conserved sites.
Top All predicted targets, irrespective of site conservation, ranked by total context score [View predicted targets with only conserved sites]

Target gene	Gene name	Conserved sites				Poorly conserved sites				Repre-sentative miRNA	Total context score
		total	8mer	7mer-m8	7mer-1A	total	8mer	7mer-m8	7mer-1A		
CROT	carnitine O-octanoyltransferase	1	0	1	0	3	2	1	0	hsa-miR-520a-3p	-1.51
ZKSCAN1	zinc finger with KRAB and SCAN domains 1	1	0	1	0	4	2	2	0	hsa-miR-373	-1.00
SH3TC2	SH3 domain and tetratricopeptide repeats 2	0	0	0	0	7	1	4	2	hsa-miR-372	-0.92
TGFBR2	transforming growth factor, beta receptor II (70/80kDa)	2	1	1	0	2	0	1	1	hsa-miR-302e	-0.90
LATS2	LATS, large tumor suppressor, homolog 2 (Drosophila)	2	2	0	0	0	0	0	0	hsa-miR-520b	-0.90
PHKA1	phosphorylase kinase, alpha 1 (muscle)	0	0	0	0	3	1	1	1	hsa-miR-302d	-0.89
CTSS	cathepsin S	0	0	0	0	3	0	3	0	hsa-miR-373	-0.84
RP5-1022P6.2	hypothetical protein KIAA1434	0	0	0	0	2	1	1	0	hsa-miR-520b	-0.83
RAB22A	RAB22A, member RAS oncogene family	2	0	2	0	1	1	0	0	hsa-miR-520b	-0.82
YOD1	YOD1 OTU deubiquinating enzyme 1 homolog (S. cerevisiae)	2	0	2	0	3	0	3	0	hsa-miR-372	-0.82
CYBRD1	cytochrome b reductase 1	1	0	1	0	3	1	2	0	hsa-miR-302d	-0.82
CYBB	cytochrome b-245, beta polypeptide (chronic granulomatous disease)	0	0	0	0	2	1	1	0	hsa-miR-520b	-0.78
APOBEC4	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 4 (putative)	0	0	0	0	2	1	1	0	hsa-miR-372	-0.77
UBE2B	ubiquitin-conjugating enzyme E2B (RAD6 homolog)	1	1	0	0	2	0	2	0	hsa-miR-372	-0.76
OXR1	oxidation resistance 1	1	0	1	0	1	1	0	0	hsa-miR-372	-0.76
MIB1	mindbomb homolog 1 (Drosophila)	0	0	0	0	2	2	0	0	hsa-miR-520b	-0.75
NR2C2	nuclear receptor subfamily 2, group C, member 2	2	1	1	0	1	0	1	0	hsa-miR-302c	-0.74
SERF1A	small EDRK-rich factor 1A (telomeric)	0	0	0	0	3	0	3	0	hsa-miR-372	-0.74

Fig. 2. Screen shot of a search result of the TargetScanHuman database. The screen displays part of conserved and nonconserved targets of a human miRNA, hsa-miR-106a. A total of 3,101 targets are predicted for hsa-miR-106a. miR-106a is a member of the miR-106/302 miRNA family. The "Representative miRNA" column lists the miRNA in the family member with the most favorable "Total context score" for each target.

5. To download the predicted targets into Excel, place the cursor within the target table in the browser and right click to export the entire table to Microsoft Excel (see Note 2).
6. Excel will retrieve all of the predicted targets in a new workbook. Save the new workbook based on the name of the miRNA (e.g., "miR-106a target.xls"). The Excel contains the following columns: "Target gene," "Gene name," "Conserved sites," "Poorly conserved sites," "Representative miRNA," "Total context score," "Aggregate PCT," "Previous TargetScan publication(s)," and "Links to sites in UTRs."
7. Continue to predict and save targets for individual miRNAs by repeating steps 2–6 until predicted targets for all miRNAs are saved.

3.2. Compiling Coregulated miRNA Targets Using Excel Pivot Table

Upon the completion of protocol 3.1, predicted targets for individual miRNAs are saved in separate workbooks. In order to identify high-efficacy targets coregulated by DE-miRs, all targets for individual DE-miR need to be assembled in a single worksheet and filtered by context score to remove low-efficacy targets. The number of regulating DE-miR for each high-efficacy target can then be analyzed using the Excel pivot function. The same

compiling protocol is also used to identify high-efficacy targets coregulated by Ref-miRs:

1. Open a blank Excel workbook and named the file "Total DE-miR target." Name the first worksheet as "raw data" and label the first three columns as "Target gene," "Total context score," and "miRNA name."
2. Open the first DE-miR target-containing workbook (e.g., "miR-106a target.xls" in this example), copy data in "Target gene" and "Total context score" columns (the first and 12th columns) into the first and second columns of the "raw data" worksheet (Fig. 3).
3. Type the name of the first DE-miR (e.g., miR-106a) into "C2" cell in the "raw data" worksheet and copy the miRNA name to the remaining cells data (see Note 3).
4. Open the second DE-miR target-containing workbook (e.g., "miR-138 target.xls" in this example), copy data in "Target gene" and "Total context score" columns and paste them into the "raw data" worksheet below the data of the first DE-miR targets.
5. Fill in the name of the second DE-miR (miR-138) (Fig. 3).

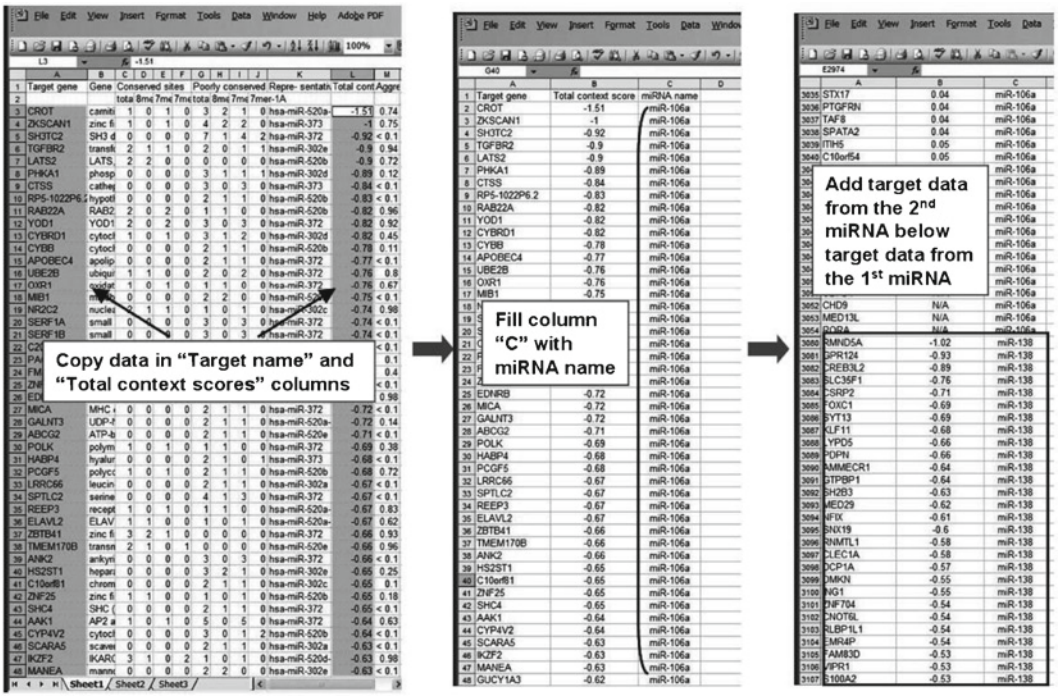


Fig. 3. Assemble predicted targets for all miRNAs in Excel (Subheading 3.2, steps 1–5). Copy "Target name" and "Total context score" data from the first miRNA workbook (steps 1 and 2). Paste data into the "total DE-miR target" workbook and add miRNA name in column C (step 3). Assemble data from the second miRNA (steps 4 and 5).

6. Continue to assemble the target name, context score, and miRNA name for the remaining DE-miRs into the same worksheet. In the example used in this chapter, the finished worksheet contains 50,716 rows of unfiltered targets for 11 DE-miRs.
7. To select high-efficacy targets for DE-miRs, sort the target table by "total context score" column in ascending order. Use "autofilter" to select and delete targets with context score greater than −0.2 (see Note 4).
8. Excel pivot function is then used to generate a summary table with miRNA counts for individual targets. Place the cursor within the data table range. Click the "Insert" ribbon to insert a pivot table in a new worksheet (Fig. 4).
9. Go to the worksheet with pivot table, drag "Target gene" field into the row title, and drag "miRNA name" field into the column title. Drag "miRNA name" field into the sum area. The resulting pivot table will list all unique targets and their interacting miRNAs. The number of interacting miRNAs for each target are summarized in the "Grand total" column (see Note 5) (Fig. 4).

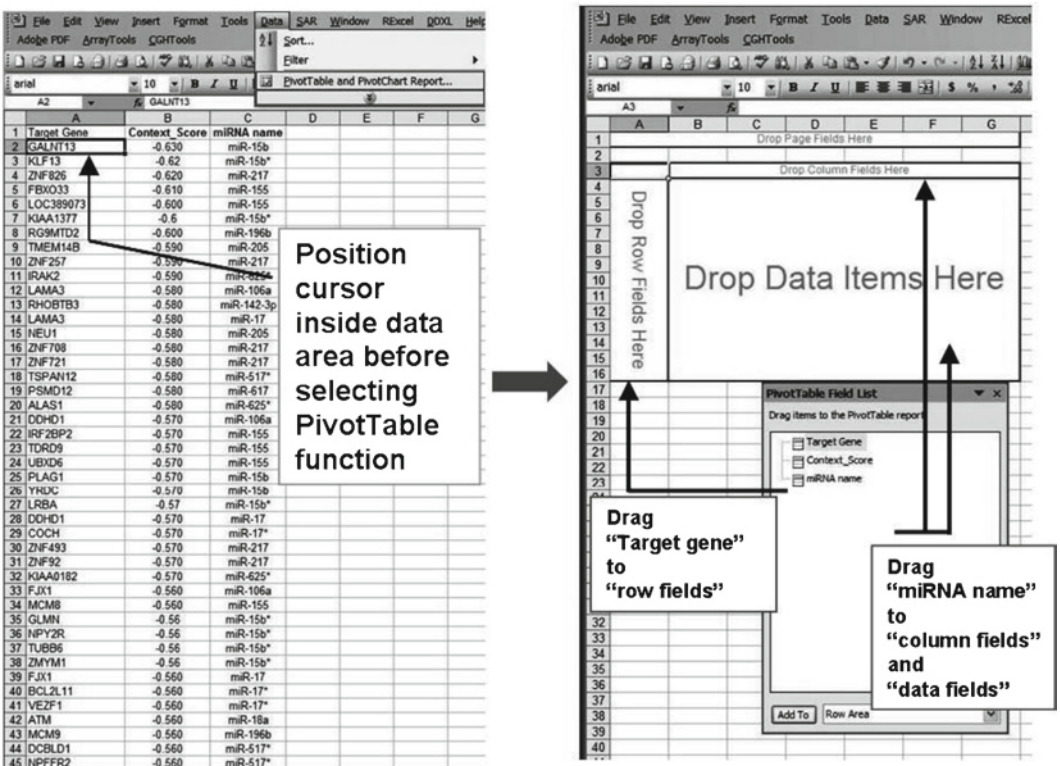


Fig. 4. Compile coregulated targets using Excel PivotTable function (Subheading 3.2, steps 8 and 9). Position cursor within the range of data table and activate the PivotTable function (step 8). Drag data into appropriate region to perform the pivot analysis (step 9).

10. Make sure that the data display is set to "count of miRNA name" in the pivot table. To select the targets coregulated by DE-miRs, copy the "target gene" and "grand total" columns into a new worksheet, rename the new worksheet as "compiled DE-miR target" (Fig. 5).
11. In the "compiled DE-miR target" worksheet, sort the grand total column in descending order. Select targets coregulated by at least three miRNAs by filtering out targets with grand total value ≤2 (Fig. 5).
12. Copy and paste the filtered gene list from step 11 into a text editor, e.g., NotePad, and save the gene list as "DE-miR coregulated targets.txt" in a text file format. This text file will be used as the gene list for pathway enrichment analysis described in Subheading 3.3 (see Note 6).
13. Repeat the above steps to generate gene list coregulated by Ref-miRs and save as "Ref-miR coregulated targets.txt."

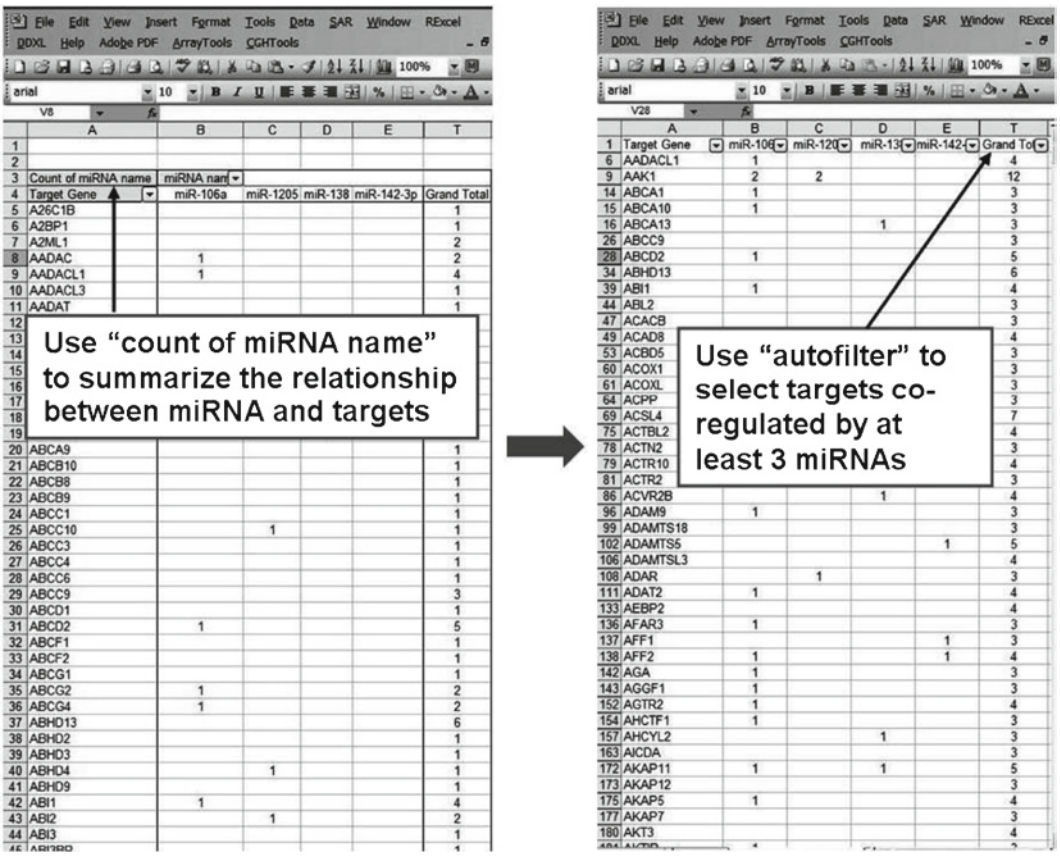


Fig. 5. Identify target genes coregulated by multiple miRNAs (Subheading 3.2, steps 10 and 11). Set the data display in PivotTable to "Count of miRNA name" (step 10). Copy pivoted data into a new worksheet. Filter the "Grand Total" column to select targets coregulated by more than two miRNAs (step 11).

3.3. Analyzing Biological Pathways Enriched with Coregulated miRNA Targets Using GeneGO MetaCore

To analyze biological pathways enriched by DE-miR coregulated targets, statistical enrichment analysis is performed using the MetaCore pathway enrichment tool. MetaCore uses the formula of hypergeometric distribution to calculate *p* value, which reflects the probability of a pathway to arise by chance. The *p* value for pathways enriched with targets coregulated by Ref-miRs is calculated using the same method:

1. In MetaCore "Data Manager" page, upload the target list as "General" data type.
2. Browse to select the text file (e.g., DE-miR coregulated targets.txt).
3. Specify "Gene Symbol" and "*Homo sapiens*" for uploaded data.
4. Once data are uploaded, refresh the "Data Manager" to view active experiment data.
5. From the "View" menu, select "Functional Ontology Enrichment" tab and select the "GeneGo Pathway Maps" item to display all pathways enriched with miRNA targets (Fig. 6).
6. In the "GeneGo Pathway Maps" result page, click on "Maps" to display the list of statistically significantly enriched pathway maps (Fig. 6). By default, MetaCore displays the top 50 items on the list. To display all pathways, scroll down to the bottom

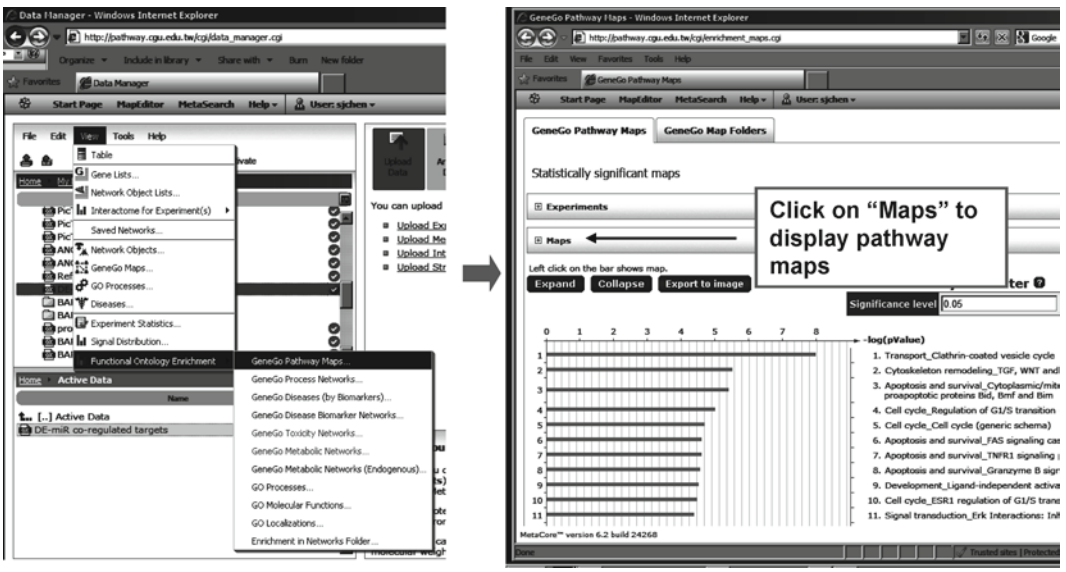


Fig. 6. View pathways enriched with coregulated miRNA targets in MetaCore (Subheading 3.3, steps 5 and 6). In MetaCore Data Manager page, make sure that "DE-miR co-regulated target" is the only list in the active data area. Use "view functional ontology enrichment" to analyze GeneGO pathway maps enriched with coregulated miRNA targets (step 5). The pathway enrichment analysis result will be displayed in graphic format. Click on "Maps" to display pathway data in tabular format (step 6).

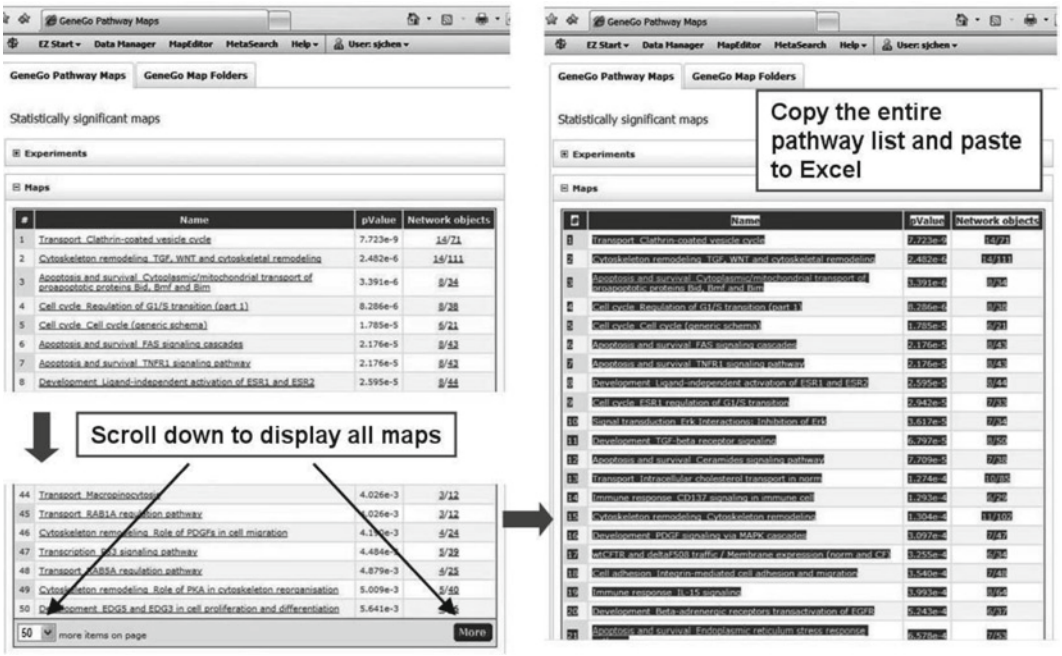


Fig. 7. Display all GeneGO pathways enriched by DE-miR targets (Subheading 3.3, steps 6 and 7). By default, MetaCore only display the top 50 enriched pathways. Scroll to the bottom of the table list, select "all" items on page and click "More" to display the complete pathway list (step 6). Copy the entire pathway table to an Excel workbook (step 7).

of the maps list table, choose "all" items on page, and press "More" to display all pathways in the list (Fig. 7).

7. Once the entire pathway list is displayed on the browser, right click in the table to export the data into Excel. Alternatively, use mouse to select the entire table and copy the data to Excel. Save the pathway list as "DE-miR pathway list" (Fig. 7).
8. From the MetaCore "Data Manager" page, delete the target gene list from the active experiment area.
9. Repeat steps 1–7 to upload data and retrieve the enriched pathway list for the reference gene list. Save the pathway list as "Ref-miR pathway list."

3.4. Identifying Specifically Enriched Biological Pathways Using Excel

To identify pathways specifically regulated by DE-miR targets, the ratio of *p* values between DE-miRs and Ref-miRs were calculated and used as specificity index:

1. In Excel, copy data in "Name" and "pValue" columns from the "DE-miR pathway list" to a new worksheet named "Combined pathways." In column "C," type "DE" as the list name (Fig. 8).
2. Copy data in "Name" and "pValue" columns from the "Ref-miR pathway list" and paste the data into the "Combined pathway"

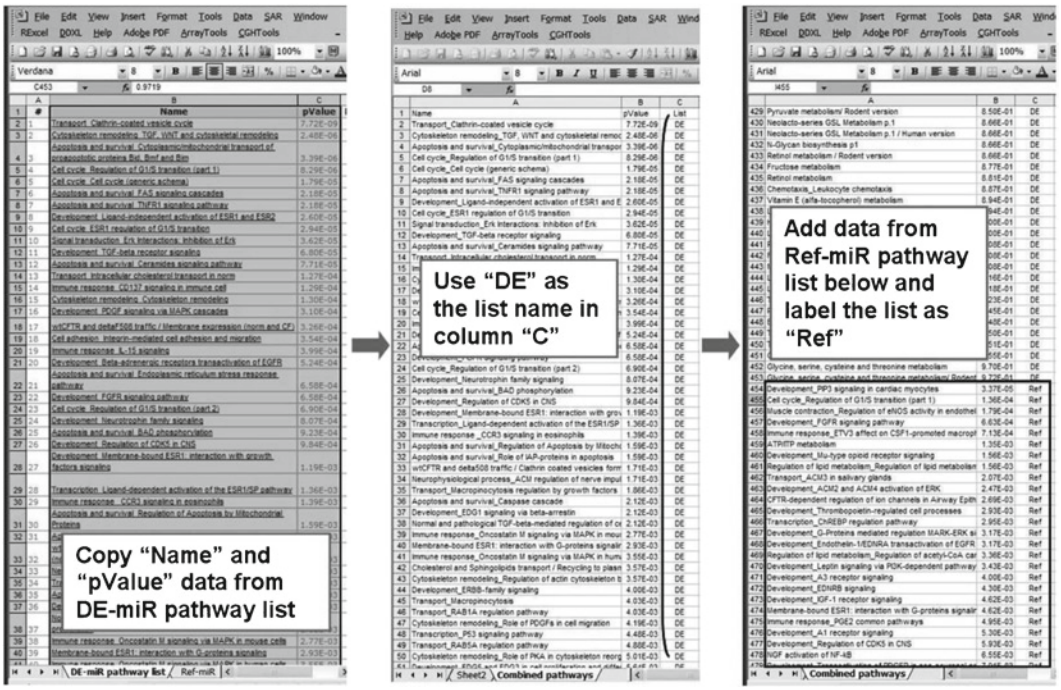


Fig. 8. Assemble the enriched pathway lists in Excel (Subheading 3.4, steps 1 and 2). Copy pathway name and pValue data from the DE-miR pathway list into the "combined pathways" worksheet. Add "DE" as the list name in column C (step 1). Copy and paste pathway name and pValue data from the Ref-miR pathway list into the "combined pathway" worksheet. Use "Ref" as the list name in column C (step 2).

worksheet below the DE-miR pathway data and label the list name "Ref" as shown in Fig. 8.

3. Position the cursor inside the pathway data area and activate the PivotTable function in Excel. Drag the "Name" field into the "Row field" area. Drag the "List" field into "Column field" area. Drag the "pValue" field into the "data field" area. Excel pivot function will automatically align pathway with associated pValues from each lists (Fig. 9).
4. Display pivoted data as "Sum of pValue." Copy the pathway and pValue data into a new worksheet named "enriched pathway" (Fig. 9). Use "autofilter" function to select all blank cells in "Ref" column. Fill in "1" as the *p* value for all blank cells. Use "autofilter" to select all blank cells in "DE" column. Fill in "1" as the *p* value for all blank cells (see Note 7) (Fig. 10).
5. Label column "D" as "Specificity ratio." Divide the pValue in Ref column with pValue in DE column to generate the specificity ratio in column "D" (Fig. 10).
6. Significantly enriched pathways can be selected by filtering the *p* value of "DE" column. Typically, significantly enriched

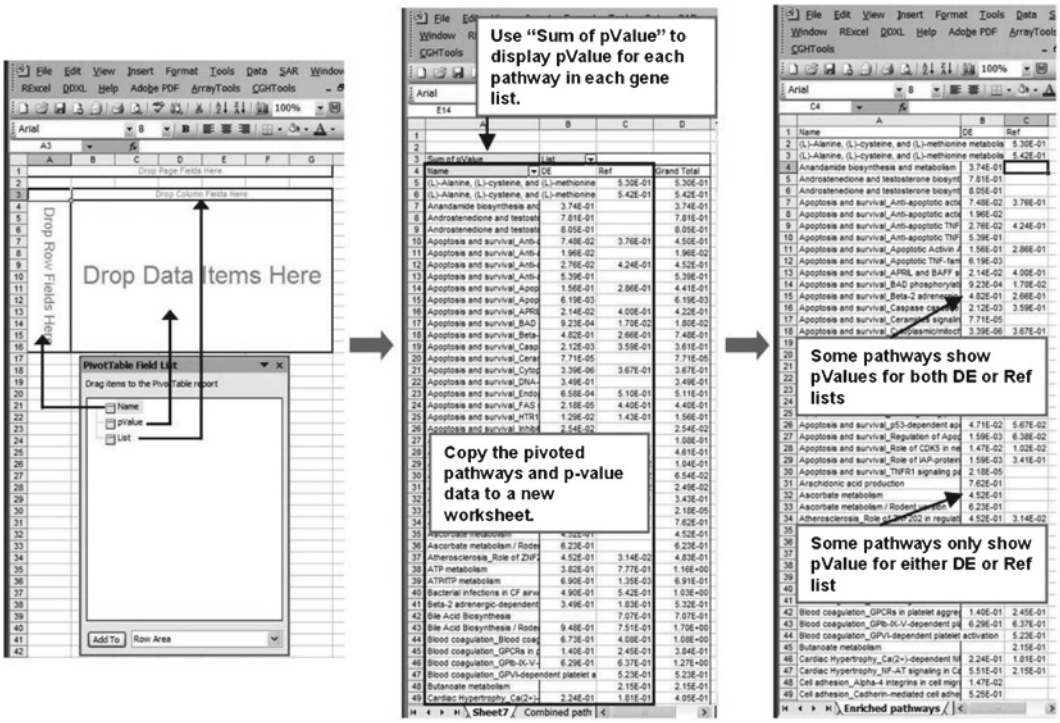


Fig. 9. Use Excel PivotTable to analyze the pValues in "DE" and "Ref" list for individual pathways (Subheading 3.4, steps 3 and 4). With cursor positioned within the data range in the "combined pathway" worksheet, activate the PivotTable function in Excel. Drag data into appropriate fields to perform pivot analysis (step 3). Display data as "Sum of pValue." Copy the pivoted pathways and *p* value data into "enriched pathways" worksheet (step 4). Note that not all pathways will have pValues for both "DE" and "Ref" lists.

pathways are selected by setting the *p* value threshold at $p < 0.01$. Alternatively, false discovery rate can be used to select the pathway list. Typically, false discovery rate threshold is set at 0.1 (see Note 8) (Fig. 10).

7. Specifically enriched pathway list can be further generated by filtering the specificity index. Threshold for selecting specifically enriched pathways can be set at "specificity index" greater than 100 (Fig. 10).

4. Notes

1. For miRNAs not included in the precompiled miRNA list, TargetScan offers an online "Custom Prediction" interface. Users can input the seed sequence (nt 2–8) of mature miRNA to run custom prediction. However, the TargetScan online custom prediction only displays targets with conserved sites. No context score will be provided in the custom prediction.

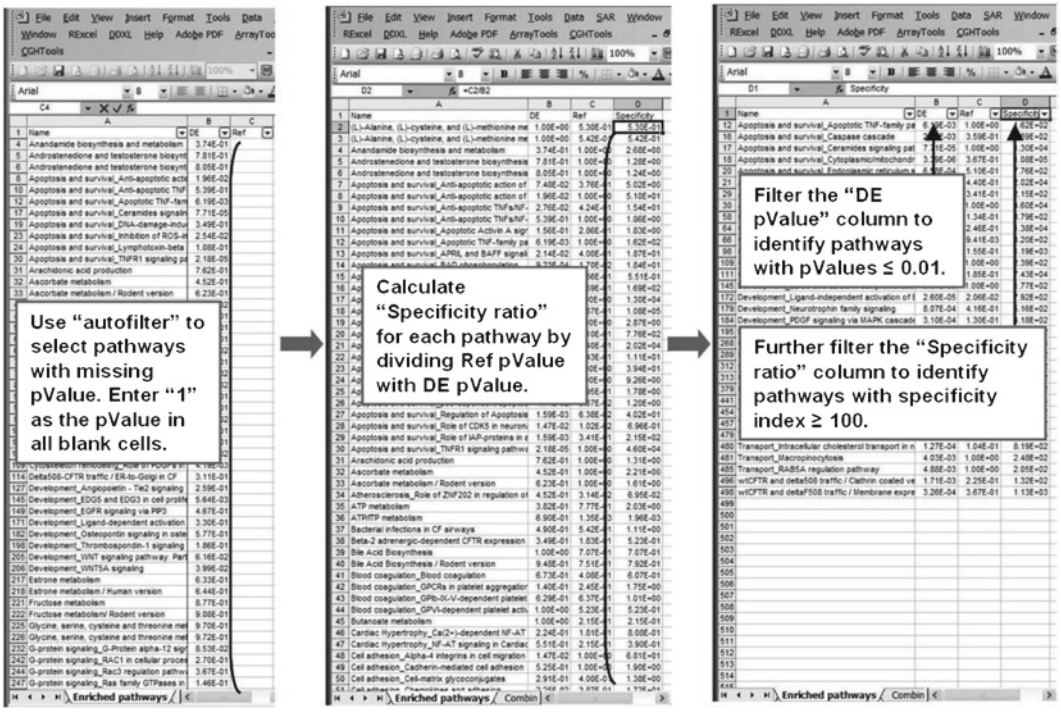


Fig. 10. Calculate specificity index and select significantly enriched pathways (Subheading 3.4, steps 4–7). Use "autofilter" to select pathways with missing *p* value and enter "1" as the *p* value (step 4). Calculate specificity ratio for all pathways (step 5). Use "autofilter" to identify pathways significantly enriched with targets co-regulated by DE-miRNAs (step 6). Specifically enriched pathways can be selected by further filter the specificity index column (step 7).

Alternatively, users can download Perl scripts and 3′UTR data from the TargetScan website and run the custom prediction locally. To identify all conserved and nonconserved sites using a custom set of data, use the targetscan_50.pl script. To calculate context scores for a set of predicted miRNA sites in a custom set of data, use the targetscan_50_context_scores.pl script. The 3′UTR database contains UTR sequences from 23-way alignment (UTR_Sequences.txt.zip). Setting up TargetScan locally for custom prediction is beyond the scope of this article.

2. FireFox does not provide the right click button to directly export data in table format into Excel. To copy data from FireFox into Excel, use the mouse to highlight the data you want to copy, and then click on the "EDIT" function on the Toolbar and select "COPY." Then, you can paste that data into Excel. FireFox does not preserve the original table format. Therefore, the column labels appeared on the first row in Excel do not match the labels displayed in the TargetScan result interface. Manually relabeling the columns in Excel is needed to do the next step correctly.

3. TargetScan groups miRNAs into families and displays context score for miRNA:mRNA pair using a "Representative miRNA" from that family. The "representative miRNA" listed in the table is the miRNA with the most favorable total context score in that family. For a miRNAs with multiple members in the family (e.g., miR-106a is a member of miR-106/302 family), the representative miRNA may be different from the miRNA in your list. Therefore, do not use the data in "Representative miRNA" column as the name of individual miRNA.
4. TargetScan uses "context score" to reflect differential site efficacy. A more negative context score is associated with a more favorable site. The context score threshold for high-efficacy target was set at −0.2. Sites with a total context score greater than −0.2 are filtered out from the target list.
5. Similar pivot function can be found in the OpenOffice Calc application. "DataPilots" is OpenOffice Calc's equivalent of Excel's pivot table. First, select the data containing miRNA target. Then select "Data" → "DataPilot" → "Start" from the menu. Calc will then ask whether you want to base your analysis on the current selection or you want to base the analysis on another data source such as a database or an external interface. In this case, leave the "Current selection" option checked and click "OK." The next screen will show a graphical representation of the pivot table. You will also see buttons representing fields (in this case, the fields correspond to the columns in the underlying spreadsheet). To create a pivot table for all genes targeted by all miRNAs, drag and drop the "Target gene" button to the "Row Fields" area, and then drag and drop the "miRNA name" to the "Column Fields" area. Also, drag and drop the "miRNA name" to the "Data Fields" area. The calculation shows a count of the miRNA field.
6. This example calculates targets coregulated by at least three miRNAs in the DE-miR list which contains 11 miRNAs. The threshold number of coregulated miRNAs may be adjusted based on the total number of DE-miRs.
7. MetaCore only displays pValue for pathways which contains at least one gene from the submitted gene list. Therefore, the "DE-miR targets" gene list and the "Ref-miR targets" gene list will produce different enriched pathway lists. The degree of overlap between the two pathway lists depends on the number of genes in each gene list. After pivoting, some pathways will show pValues in both "DE" and "Ref" gene lists while other pathways will show pValue for only one of the gene list. Before calculating the specificity ratio for each pathway, all missing pValues must be replaced by 1 to avoid error in calculation.

8. There is no standard approach for estimating the statistical significance in the enrichment analysis. Although *p* values may be used as a guide, biological functions should also be considered when selecting pathways for follow-up validation.

References

1. Ambros, V. (2004) The functions of animal microRNAs, *Nature 431*, 350–355.
2. Bartel, D. P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function, *Cell 116*, 281–297.
3. Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., and Croce, C. M. (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers, *Proceedings of the National Academy Sciences of United States of America 101*, 2999–3004.
4. Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R., and Golub, T. R. (2005) MicroRNA expression profiles classify human cancers, *Nature 435*, 834–838.
5. Chen, C. Z. (2005) MicroRNAs as oncogenes and tumor suppressors, *New England Journal of Medicine 353*, 1768–1771.
6. Croce, C. M. (2008) Oncogenes and cancer, *New England Journal of Medicine 358*, 502–511.
7. Bloomston, M., Frankel, W. L., Petrocca, F., Volinia, S., Alder, H., Hagan, J. P., Liu, C. G., Bhatt, D., Taccioli, C., and Croce, C. M. (2007) MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis, *JAMA 297*, 1901–1908.
8. Bottoni, A., Zatelli, M. C., Ferracin, M., Tagliati, F., Piccin, D., Vignali, C., Calin, G. A., Negrini, M., Croce, C. M., and Degli Uberti, E. C. (2007) Identification of differentially expressed microRNAs by microarray: a possible role for microRNA genes in pituitary adenomas, *Journal of Cellular Physiology 210*, 370–377.
9. Chen, H. C., Chen, G. H., Chen, Y. H., Liao, W. L., Liu, C. Y., Chang, K. P., Chang, Y. S., and Chen, S. J. (2009) MicroRNA deregulation and pathway alterations in nasopharyngeal carcinoma, *British Journal of Cancer 100*, 1002–1011.
10. Ciafre, S. A., Galardi, S., Mangiola, A., Ferracin, M., Liu, C. G., Sabatino, G., Negrini, M., Maira, G., Croce, C. M., and Farace, M. G. (2005) Extensive modulation of a set of microRNAs in primary glioblastoma, *Biochemical and Biophysical Research Communications 334*, 1351–1358.
11. Iorio, M. V., Ferracin, M., Liu, C. G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M., Menard, S., Palazzo, J. P., Rosenberg, A., Musiani, P., Volinia, S., Nenci, I., Calin, G. A., Querzoli, P., Negrini, M., and Croce, C. M. (2005) MicroRNA gene expression deregulation in human breast cancer, *Cancer Research 65*, 7065–7070.
12. Szafranska, A. E., Davison, T. S., John, J., Cannon, T., Sipos, B., Maghnouj, A., Labourier, E., and Hahn, S. A. (2007) MicroRNA expression alterations are linked to tumorigenesis and non-neoplastic processes in pancreatic ductal adenocarcinoma, *Oncogene 26*, 4442–4452.
13. Varnholt, H., Drebber, U., Schulze, F., Wedemeyer, I., Schirmacher, P., Dienes, H. P., and Odenthal, M. (2008) MicroRNA gene expression profile of hepatitis C virus-associated hepatocellular carcinoma, *Hepatology 47*, 1223–1232.
14. Yanaihara, N., Caplen, N., Bowman, E., Seike, M., Kumamoto, K., Yi, M., Stephens, R. M., Okamoto, A., Yokota, J., Tanaka, T., Calin, G. A., Liu, C. G., Croce, C. M., and Harris, C. C. (2006) Unique microRNA molecular profiles in lung cancer diagnosis and prognosis, *Cancer Cell 9*, 189–198.
15. Gusev, Y., Schmittgen, T. D., Lerner, M., Postier, R., and Brackett, D. (2007) Computational analysis of biological functions and pathways collectively targeted by co-expressed microRNAs in cancer, *BMC Bioinformatics 8 Suppl 7*, S16.
16. Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets, *Cell 120*, 15–20.
17. Grimson, A., Farh, K. K., Johnston, W. K., Garrett-Engele, P., Lim, L. P., and Bartel, D. P. (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing, *Molecular Cell 27*, 91–105.
18. Wang, Y. P., and Li, K. B. (2009) Correlation of expression profiles between microRNAs

and mRNA targets using NCI-60 data, *BMC Genomics 10*, 218.

19. Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008) Widespread changes in protein synthesis induced by microRNAs, *Nature 455*, 58–63.
20. Baek, D., Villen, J., Shin, C., Camargo, F. D., Gygi, S. P., and Bartel, D. P. (2008) The impact of microRNAs on protein output, *Nature 455*, 64–71.
21. Liu, Q., Fu, H., Sun, F., Zhang, H., Tie, Y., Zhu, J., Xing, R., Sun, Z., and Zheng, X. (2008) miR-16 family induces cell cycle arrest by regulating multiple cell cycle genes, *Nucleic Acids Research 36*, 5391–5404.
22. Cloonan, N., Brown, M. K., Steptoe, A. L., Wani, S., Chan, W. L., Forrest, A. R., Kolle, G., Gabrielli, B., and Grimmond, S. M. (2008) The miR-17-5p microRNA is a key regulator of the G1/S phase cell cycle transition, *Genome Biology 9*, R127.

Chapter 17

Utilization of SSCprofiler to Predict a New miRNA Gene

Anastasis Oulas and Panayiota Poirazi

Abstract

Experimental identification provides a valuable yet slow and expensive method for predicting novel miRNA genes. With the advent of computational procedures, it is now possible to capture characteristic features of miRNA biogenesis in an in silico model, resulting thereafter in the fast and inexpensive prediction of multiple novel miRNA gene candidates. These computational tools provide valuable clues to experimentalists, allowing them to narrow down their search space, making experimental verification less time consuming and less costly. Furthermore, the computational model itself can provide biological information as to which are the dominant features that characterize these regulatory units. Moreover, large-scale, high-throughput techniques, such as deep sequencing and tiling arrays, require computational methods to analyze this vast amount of data. Computational miRNA gene prediction tools are often used in synergy with high-throughput methods, aiding in the discovery of putative miRNA genes. This chapter focuses on a recently developed computational tool (SSCprofiler) for identifying miRNA genes and provides an overview of the methodology undertaken by this tool, and defines a stepwise guideline on how to utilize SSCprofiler to predict novel miRNAs in the human genome.

Key words: Hidden Markov model, Human microRNA prediction, Expression data, Web tool

1. Introduction

The recent discovery of miRNAs has opened up a whole new world of noncoding regulation (1). The biogenesis of miRNAs is initiated by the transcription of a large precursor that folds into an imperfect stem–loop or hairpin structure. The mature miRNA, which is the functional part, is derived from this precursor and is usually 19–27 nucleotides long. The regulatory function of the mature miRNA in mammalian systems is dependent highly on an elaborate complex of proteins that bind to the mature miRNA. Driven by complementary base paring of the mature miRNA and its target, this complex binds primarily to the 3′-UTR region of

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_17, © Springer Science+Business Media, LLC 2011

specific target mRNA(s), ultimately leading to the inhibition of translation and/or the degradation of the targeted mRNA.

While experimental approaches have contributed significantly toward the prediction of novel miRNA genes, they are often limited by their complexity and suboptimality (2). More specifically, inefficiencies of experimental techniques arise from low expression of miRNAs, making isolation and cloning very difficult, that is, the tissue specificity of miRNAs as well as technical difficulties with the cloning procedure itself. Often the most tedious problem faced by experimentalists is the identification of genomic regions in which to search for novel miRNAs. Computational prediction of miRNA genes within such regions provides a valuable alternative technique which offers a fast, cheap, reliable, and effective way of identifying novel miRNA genes. By predicting genomic regions that are likely to contain miRNA genes, computational tools enable experimentalists to direct their efforts to these "hotspots," hence facilitating the search process (see Note 1).

Effective computational prediction of putative miRNA genes requires taking under consideration specific characteristic features, derived either from experimental (3–5) or from computational evidence (6–10), that define these molecules. These features include sequence, secondary structure, and species conservation. Computational prediction of miRNA genes has been addressed via the use of both supervised and unsupervised algorithms. The first are trained on known miRNA biological features to build a classification scheme or predictive model which is then used to identify putative miRNAs, while the second consist of alignment tools or conservation detection methods (see Note 2).

Recently, a sophisticated miRNA gene prediction tool (SSC profiler (11)) that relies on supervised machine learning algorithms (hidden Markov models) has proved to be very effective, achieving high performance on identifying miRNAs in the human genome.

SSCprofiler is a successful example of a next generation tool that combines multiple biological features of miRNAs (sequence, structure, and conservation) and also takes into consideration high-throughput methods in the form of tiling array expression data to further provide greater credence to computational predictions (see Note 3). Moreover, SSCprofiler is one of the few computational tools that provided raw material for wet-lab experiments, the results of which were positive with respect to four miRNA gene candidates. For these reasons, the next section will provide a detailed stepwise outline on how to use the scanning interface of SSCprofiler to search for novel human miRNAs.

2. Materials

SSCprofiler works on a quad core xeon server (2.66GH) with 2 × 2GB dual rank DIMMs supported by the Institute of Molecular Biology and Biotechnology (IMBB). The scanning interface of SSCprofiler can be found at: http://mirna.imbb.forth.gr/SSCprofiler.html (see Fig. 1). The software was implemented using

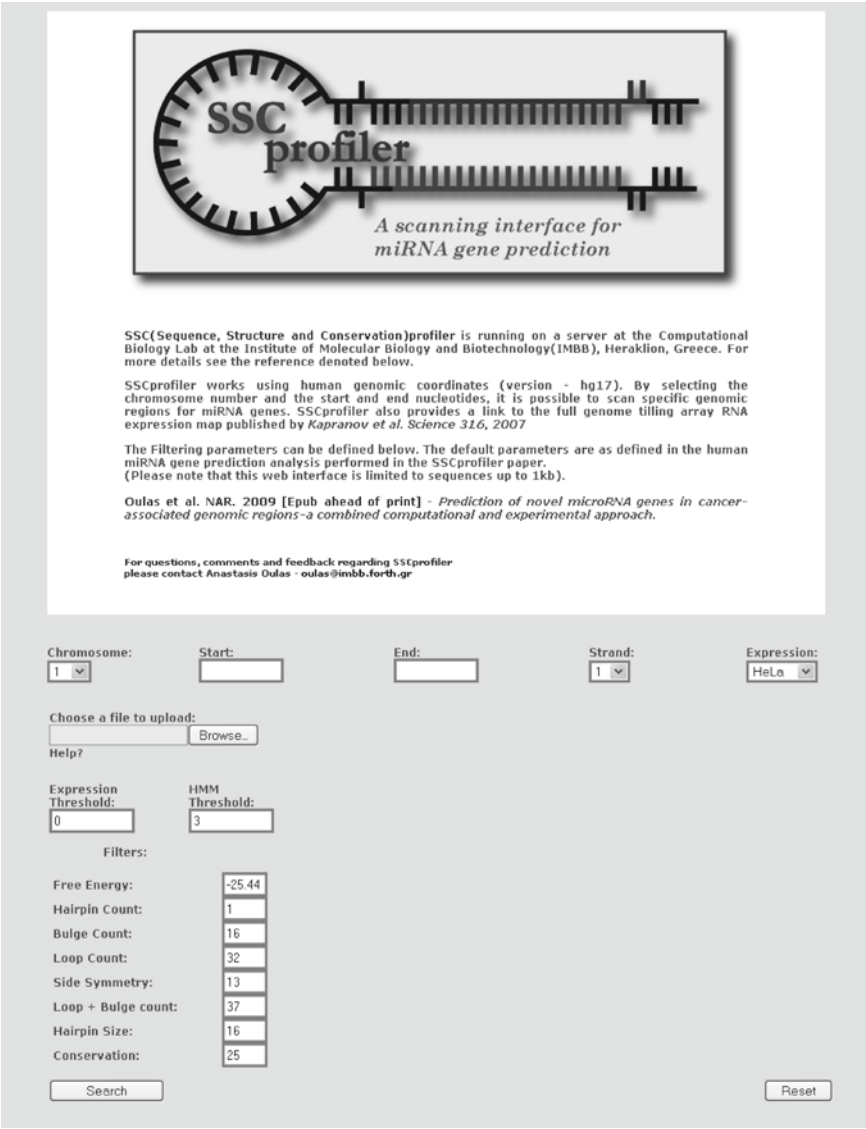


Fig. 1. Screenshot of the scanning interface of SSCprofiler http://mirna.imbb.forth.gr/SSCprofiler.html.

the high-level programming language JAVA; however, it also utilizes various publically available tools developed in the C programming language. These include the biosequence analysis tool, HMMER (12), based on profile hidden Markov models, which can be found as a web server at: http://bioweb2.pasteur.fr/motif/intro-en.html#hmm and also the RNA secondary structure prediction tool RNAfold found as part of the Vienna RNA package (13) and supported as a web server at: http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi. The SSCprofiler scanning interface makes use of the human genome version hg17 – May 2004 for mapping genomic coordinates to sequence and the multiz full genome alignment files of the human May 2004 hg17 genome assembly to obtain conservation information for human and seven other vertebrate genomes: Mouse May 2004 (mm5), Rat June 2003 (rn3), Dog July 2004 (canFam1), Chicken February 2004 (galGal2), Fugu August 2002 (fr1), Zebrafish November 2003 (danRer1). Note: Chimp data were not included owing to high percentage similarity (~95%) with humans.

3. Methods

3.1. Scanning Genomic Regions for Profiles

The process of scanning genomic regions for miRNA precursors using SSCprofiler is approached by scanning the genome, and subsequently, the resulting candidates are ranked according to their expression value derived from large-scale expression data. This approach entails the following steps.

1. The user defines a nucleotide start and end position, which encompass the search space (limited to 1,000 nt) in a specific chromosome of the human genome. A sliding window of selected length is passed along the genomic sequence shifting 1-nucleotide at a time.
2. The user is also permitted to define the DNA strand which will be scanned, namely, the positive (1) or the negative (−1) strand.
3. For every window shift, sequence, structure, and conservation information is retrieved. According to the tool's prior training specifications, all three of the above features are utilized, unless otherwise specified by the user. Structure prediction is performed by the computational tool RNAfold (13), and conservation across multiple species is obtained from the multiz files.
4. Each sequence is passed through a set of filters utilizing the user-defined filtering parameters. The filtering parameters

provided by the SSCprofiler scanning interface (including their default values) are shown below.

(a) The number of hairpins (default value = 1).

(b) The number of bulges (default value ≤16).

(c) The number of loops (default value ≤32).

(d) The degree of asymmetry, often defined as the difference in loops + bulges on either side of the hairpin (default value ≤13).

(e) The combination of bulges and loops, i.e., the sum of loop and bulge count (default value ≤37).

(f) The length of the hairpin (default value ≤16).

(g) The minimum folding energy as defined by RNAfold (default value ≤−25.44 kcal/mol).

(h) The percentage of conservation across species – usually according to multiz full genome alignment files (default value ≥25% of nucleotides conserved).

The default values for these parameters are as defined in the SSCprofiler study (11) and represent optimum cut-off values between positive (miRNA) and negative sequences that were used to build the classifier. These parameters can be adjusted by the user.

5. For each sequence, the features used during training (sequence, structure, and conservation) are considered simultaneously for every nucleotide position in the genomic sequence.
6. The trained HMM classifier is used to assign a score to each genomic sequence within the sliding window. The score threshold can be selected by the user. The default score (value of 3) was defined as the score where sensitivity and specificity from the training/validation process were optimal in the respective study (11).
7. Candidates that overlap by ≤50 nt (average minimum genomic distance between two miRNA genes) are grouped together and the candidate with the highest HMM score is used to represent the cluster.
8. The candidates can also be assessed according to their expression in specific cell cultures using tiling array data. The user can specify the cell culture that will be used (HeLa or HepG2) as well as the expression threshold or cutoff by which the data will be further filtered. The default value for the expression threshold is initially set to 0.
9. For batch jobs, the user can specify a file that contains multiple genomic regions each limited to 1,000 nt. The format of this file is defined in the SSCprofiler scanning interface.

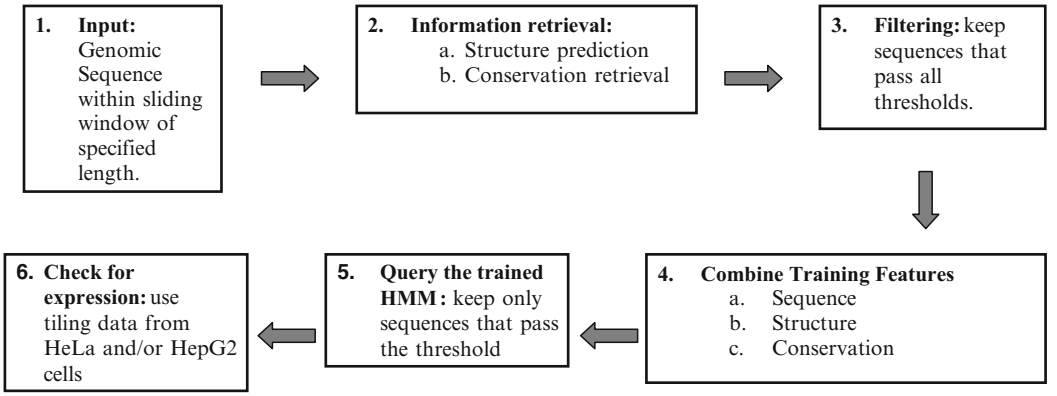


Fig. 2. Flowchart of the scanning procedure. The process of scanning genomic regions for miRNA precursor profiles involves six steps: *Step 1*: A sliding window of selected length is passed along the genomic sequence shifting 1 nt at a time. *Step 2*: For every window shift, sequence structure and conservation information is retrieved according to the selected training features, i.e., structure prediction is performed and conservation is obtained from the multiz files. *Step 3*: Each sequence within the sliding window is passed through the filters utilizing the predefined filtering parameters (i.e., Hairpin Length, Asymmetry). *Step 4*: For each sequence, the features used during training (sequence, structure, and/or conservation) are generated according to the 16-letter key described in (11). This allows the simultaneous consideration of information for every nucleotide position in the genomic sequence. *Step 5*: The trained HMM model is used to assign a likelihood score to each genomic sequence within the sliding window. The HMM score threshold can be selected by the user. It is usually defined as the score where sensitivity and specificity from the training/validation process are optimal. *Step 6*: Candidates that overlap by ≤50 nt are grouped and the candidate with the highest score is used to represent the cluster. Thereafter, the candidates are assessed according to their expression in HeLa or HepG2 cells using tiling array data.

The scanning procedure of SSCprofiler is summarized in five steps in Fig. 2.

3.2. Output of the Program

The output of SSCprofiler is a table with predicted miRNA gene candidates that exceeded the user-defined thresholds. The table displays the chromosome number of the predicted miRNA gene candidate, the exact genomic coordinates (according to human genome versions hg17 – May 2004), the DNA strand that the predicted candidate resides on, and the maximum expression as well as its nucleotide location in the candidate miRNA gene (this is most likely the start of the location of the mature miRNA region). Moreover, the results display the score given by SSCprofiler to the predicted candidate as well as the cell line (HepG2 or HeLa) used to obtain expression information for the specific candidate miRNA gene. An example of SSCprofiler output is shown in Table 1. The result page of SSCprofiler further allows for user interaction by displaying a link for a secondary structure view of any predicted candidate. By clicking on this link, the user can view a pdf image of the secondary structure of a selected candidate as predicted by the program RNAfold (13) (as shown in Fig. 3).

Table 1
A representative example of a typical SCCprofiler query output

chr	nucStart	nucEnd	Strand	Max_Expression	Max_Exp_location	HMMScore	Cell_Line	2ry Structure
5	149984684	149984787	bottom_strand	264.0	149984697	26.6	HeLa	2ry Structure
9	123327358	123327461	bottom_strand	1667.5	123327416	24.4	HeLa	2ry Structure
5	148958951	148959054	bottom_strand	363.5	148959006	29.8	HeLa	2ry Structure
22	40863894	40863997	top_strand	345.0	40863943	23.8	HeLa	2ry Structure

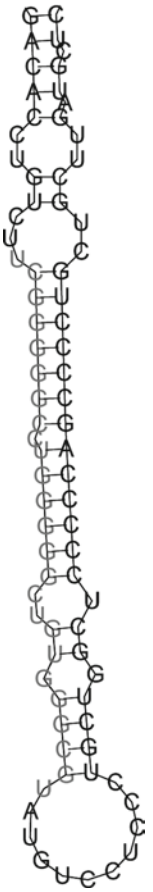


Fig. 3. The secondary structure of a predicted miRNA generated after following the link shown in Table 1. This allows the user to examine the secondary structure of a specific miRNA gene candidate predicted by SSCprofiler as well as observe the location of the mature miRNA according to its expression data.

4. Notes

1. As a cautionary note, it should be stressed that for any computational tool to prove its effectiveness, a subset of the high-scoring miRNA gene candidates should be experimentally verified.
2. Only a few studies (11, 14–16) have performed experimental verification of their miRNA gene candidates, mostly because experimental verification requires solving various difficult technical issues. One such issue is whether or not the predicted miRNA gene candidate is expressed in the tissue culture under investigation.

3. High-throughput methods such as tiling arrays (17) and deep sequencing (16, 18) are anticipated to address issues like tissue specificity and allow for a global, genome-wide expression profile of miRNAs.

Acknowledgments

This work was supported by the action 8.3.1 (Reinforcement Program of Human Research Manpower – "PENED 2003", [03ED842]) of the operational program "competitiveness" of the Greek General Secretariat for Research and Technology, a Marie Curie Fellowship of the European Commission [PIOF-GA-2008-219622] and the National Science Foundation [NSF 0515357].

References

1. Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14, *Cell 75*, 843–854.
2. Huttenhofer, A., and Vogel, J. (2006) Experimental approaches to identify non-coding RNAs, *Nucleic Acids Res 34*, 635–646.
3. Khvorova, A., Reynolds, A., and Jayasena, S. D. (2003) Functional siRNAs and miRNAs exhibit strand bias, *Cell 115*, 209–216.
4. Lee, Y., Jeon, K., Lee, J. T., Kim, S., and Kim, V. N. (2002) MicroRNA maturation: stepwise processing and subcellular localization, *EMBO J 21*, 4663–4670.
5. Lim, L. P., Glasner, M. E., Yekta, S., Burge, C. B., and Bartel, D. P. (2003) Vertebrate microRNA genes, *Science 299*, 1540.
6. Helvik, S. A., Snove, O., Jr., and Saetrom, P. (2006) Reliable prediction of Drosha processing sites improves microRNA gene prediction, *Bioinformatics 23*, 142–149.
7. Hertel, J., and Stadler, P. F. (2006) Hairpins in a Haystack: recognizing microRNA precursors in comparative genomics data, *Bioinformatics 22*, e197–e202.
8. Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., and Yekta, S. (2003b) The microRNAs of *Caenorhabditis elegans*, *Genes Dev 16*, 991–1008.
9. Sewer, A., Paul, N., Landgraf, P., Aravin, A., Pfeffer, S., Brownstein, M. J., Tuschl, T., van Nimwegen, E., and Zavolan, M. (2005) Identification of clustered microRNAs using an ab initio prediction method, *BMC Bioinformatics 6*, 267–281.
10. Yousef, M., Nebozhyn, M., Shatkay, H., Kanterakis, S., Showe, L. C., and Showe, M. K. (2006) Combining multi-species genomic data for microRNA identification using a Naive Bayes classifier, *Bioinformatics 22*, 1325–1334.
11. Oulas, A., Boutla, A., Gkirtzou, K., Reczko, M., Kalantidis, K., and Poirazi, P. (2009) Prediction of novel microRNA genes in cancer-associated genomic regions – a combined computational and experimental approach, *Nucleic Acids Res 37*, 3276–3287.
12. Eddy, S. R. (1998) Profile hidden Markov models, *Bioinformatics 14*, 755–763.
13. Hofacker, I. L. (2003) Vienna RNA secondary structure server, *Nucleic Acids Res 31*, 3429–3431.
14. Berezikov, E., Guryev, V., van de Belt, J., Wienholds, E., Plasterk, R. H., and Cuppen, E. (2005) Phylogenetic shadowing and computational identification of human microRNA genes, *Cell 120*, 21–24.
15. Nam, J. W., Shin, K. R., Han, J., Lee, Y., Kim, V. N., and Zhang, B. T. (2005) Human microRNA prediction through a probabilistic co-learning model of sequence and structure, *Nucleic Acid Res 33*, 3570–3581.
16. Friedlander, M. R., Chen, W., Adamidi, C., Maaskola, J., Einspanier, R., Knespel, S., and Rajewsky, N. (2008) Discovering microRNAs from deep sequencing data using miRDeep, *Nat Biotechnol 26*, 407–415.
17. Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., Stadler, P. F., Hertel, J., Hackermuller, J., Hofacker, I. L.,

Bell, I., Cheung, E., Drenkow, J., Dumais, E., Patel, S., Helt, G., Ganesh, M., Ghosh, S., Piccolboni, A., Sementchenko, V., Tammana, H., and Gingeras, T. R. (2007) RNA maps reveal new RNA classes and a possible function for pervasive transcription, *Science 316*, 1484–1488.

18. Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A. O., Landthaler, M., Lin, C., Socci, N. D., Hermida, L., Fulci, V., Chiaretti, S., Foa, R., Schliwka, J., Fuchs, U., Novosel, A., Muller, R. U., Schermer, B., Bissels, U., Inman, J., Phan, Q., Chien, M., Weir, D. B., Choksi, R., De Vita, G., Frezzetti, D., Trompeter, H. I., Hornung, V., Teng, G., Hartmann, G., Palkovits, M., Di Lauro, R., Wernet, P., Macino, G., Rogler, C. E., Nagle, J. W., Ju, J., Papavasiliou, F. N., Benzing, T., Lichter, P., Tam, W., Brownstein, M. J., Bosio, A., Borkhardt, A., Russo, J. J., Sander, C., Zavolan, M., and Tuschl, T. (2007) A mammalian microRNA expression atlas based on small RNA library sequencing, *Cell 129*, 1401–1414.

Chapter 18

MicroRNA Profiling Using Fluorescence-Labeled Beads: Data Acquisition and Processing

Thomas Streichert, Benjamin Otto, and Ulrich Lehmann

Abstract

The discovery of small regulatory RNA molecules during the last years has changed our understanding of many biological and pathological processes. The most prominent and best analyzed class of these small regulatory noncoding RNAs is composed of the microRNAs. The analysis of microRNA expression patterns is now widely used in biology and pathology employing a range of methodologies. However, procedures for data processing and calculations are far from standardized and differ considerably between published studies. This makes comparisons and meta-analyses still quite difficult.

In this chapter, we describe a modified method for normalization and processing of raw data obtained if utilizing fluorescence-labeled bead technology from Luminex. Inc.

Key words: MicroRNA, data normalization, Luminex. Inc. FlexmiR, fluorescence-labeled beads

1. Introduction

Regulatory or noncoding (i.e., not for protein coding) RNAs have emerged during the recent years as a new class of key regulators involved in development, normal physiology, and many different types of diseases (1, 2). Some studies suggest that microRNA profiling is superior to mRNA profiling (e.g., (3)), most probably due to the pleiotropic effects of every microRNA and the inherent greater stability of the microRNA molecules in comparison to mRNA molecules.

In breast cancer, deregulation of microRNA expression and alteration of microRNA loci are well described. Numerous diagnostic, prognostic, and therapeutic options are discussed in the literature. (See (4) for a recent review.)

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8_18, © Springer Science+Business Media, LLC 2011

There are several options for the detection and quantification of microRNA expression levels: northern blotting, semiquantitative in situ hybridization, quantitative real-time PCR, fluorescence-labeled bead technology, and array hybridization (with various probe sets). Each approach has its specific advantages and shortcomings. Advantages of the fluorescence-labeled bead technology are the omission of any amplification step, which might introduce a bias of unknown extent, and the hybridization of probe and target in a homogeneous solution which avoids any boundary surface artifacts, a potential problem for every array hybridization method.

Alternative protocols and methods for microRNA profiling of FFPE specimens (based on real-time PCR or array-hybridization) have been published (5–11). A detailed discussion of the merits and shortcomings of these studies can be found in our recent freely available publication Hasemeier et al. (12).

Besides the wet-laboratory steps, the statistical analysis plays a crucial role. The different tools are well characterized for normal gene expression or genomic analysis. Here, we describe the use of selected tools for miRNA analysis.

After the array scan and the readout of the raw signals, the measured signals have to be corrected for systematic bias and noise variation effects, a step called postprocessing. This step could be performed according to manufacturer's descriptions, or using standard procedures like median adjustment or quantile normalization (13). To detect differential miRNA expression, standardized procedures used for gene expression analysis could be applied. These procedures could comprise supervised approaches like *t*-test variations, correlation tests, bayesian networks, or principal component analysis. In addition, unsupervised methods like hierarchical clustering could be applied to identify groups within the sets.

In this chapter, we describe the normalization and processing of data obtained with the Luminex miRNA expression platform (FlexmiR), which represents a modification of the procedures suggested by Blenkiron et al. (14). The raw data were taken from a project in which the microRNA expression profile of male breast cancer was analyzed (15). A detailed description of all materials and methods required for the wet-laboratory work can be found in a recently published book chapter from the same series (16).

In general, miRNA analysis and gene expression analysis are very similar. However, the miRNA assay construction introduces a necessity for a specific choice of the methods and algorithms used. These single steps are here discussed while introducing the most common approaches and pointing out their strengths and weaknesses.

2. Materials

2.1. R Statistical Platform and BioConductor Project

R is an open-source programming platform suited for statistical analysis and visualization tasks. Due to the modular construction the integration of new functions, methods and datasets can be realized quite simply in the form of so-called packages. High flexibility, simple extensibility, and the provided interfaces to other programming languages and external programs qualify R to one of the most popular analysis platforms in microarray analysis.

BioConductor is an open-source project providing an enormous variety of extension packages for R especially designed for microarray analysis (see Note 1).

R is available for all three common operating systems. While it is already provided with the most Linux distributions, it can be downloaded as simple self-installing file for MS Windows and Mac OS. Independent of the running operating system, the basic BioConductor packages can be directly installed with two code lines within R:

```
>source(http://bioconductor.org/biocLite.R)
>biocLite()
```

Single BioConductor package, e.g., *made4* package, can be installed with:

```
>bioc("made4")
```

More information about further package installation can be found on the official BioConductor website (http://www.bioconductor.org). A very good introduction to analyzing microarrays with R and BioConductor can be found in the *R & BioConductor Manual* (http://faculty.ucr.edu/~tgirke/Documents/R_BioCond/R_BioCondManual.html).

2.2. TM4 Microarray Software Suite

The TM4 Microarray Software Suite is a modular program collection for microarray analysis (http://www.tm4.org/mev.html). The *MeV* module provides the most common statistical approaches such as hypothesis tests, clustering methods, classification procedures, or normalization steps with a user-friendly graphic interface. The new program version features an interface to the R statistical platform. The programs are Java-based, so the modules run on all common operating systems where the Sun Java Running Environment (http://www.java.com/de/download/manual.jsp) is installed. The Principal Component Analysis requires the additional installation of Sun Java 3D (http://java.sun.com/javase/technologies/desktop/java3d/).

2.3. Cluster and TreeView

Cluster and TreeView are two programs published by Michael Eisen et al. (http://rana.lbl.gov/EisenSoftware.htm). Cluster offers hierarchical clustering, self-organizing maps, k-means clustering, and principal component analysis. The program has been used for clustering in several frequently cited publications such as the initial classification of breast cancer into subtypes by Sorlie et al. (17). The results calculated by Cluster can be visualized and investigated with TreeView. Both programs perform very well but unfortunately are only available for MS Windows.

3. Methods

Due to the multiplex nature of microarray assays, the resulting signals are quite error prone. On the one hand, biological variability is frequently observed; on the other hand, scanned raw signals usually contain systematic bias and noise variability introduced by technical and experimental artifacts. The nonbiological noise components should be removed before any statistical analysis steps to ensure significance of final results.

3.1. Data Post-processing

Nonspecific hybridization effects, background fluorescence, or unusual effects such as scratches and bubbles might introduce a signal component usually termed as background noise. These variations incorporate information of the single array only. The corresponding correction procedures are thus called background correction. A second variance type can appear when several arrays, well plates, or samples are analyzed. Measured signals might be globally weaker on one array than on the other due to a different setting of the scanner or influences of the wet laboratory. Correction of this variation type is called normalization.

3.1.1. Background Correction

In miRNA expression experiments, the number of measured signals is usually much too small to assume certain conditions such as Gaussian distribution of measured signals. Therefore, background corrections as the method used by RMA (13) are not suited here.

Using bead coupled capture probes and several vials ("pools") will directly return signals for each pool and include several samples on each plate. Here, a potential approach would run a negative control sample on each plate and use the signals of that sample as a measure for background noise. The advantage of this method is that a really measured signal for unspecific hybridization and background fluorescence is used. However, effects such as regional unequal hybridization will not be included.

3.1.2. Normalization

As mentioned before, the normalization of arrays is the process of removing nonbiological data variance that occurs among arrays or samples. The variations could, e.g., be a result of slightly differing settings or conditions during the scan or differences in the processing of the samples. The problem of this step is the introduction of new artificial variances to the data. Some of the common methods performing nicely for gene expression arrays such as RMA depend on a high number of probes on each array. Because the number of probes in miRNA analysis is usually rather small, these approaches are not suitable. To overcome this problem most manufacturers include internal controls, spike-in probes, or use housekeeping genes, which can be utilized for normalization. Two common methods not depending on internal controls (ICs) are variance stabilization normalization VSN (18) and median centering. VSN incorporates a signal transformation such that the global signal variance among the samples or arrays is equalized. Median centering shifts the sample or array signals to a common median value. The following three examples show how the mentioned procedures could be applied to the results of a Luminex multivial assay. The internal control based procedure follows the recommendation of the manufacturer (http://www.luminexcorp.com/microrna/documents/templates/89-00002-00-166_FMIR_Hum_Panel_Manual_revB.pdf), while the median centering example is performed according to a procedure used by Blenkiron et al. (14) modified slightly to accommodate the used platform.

3.1.2.1. Background Correction and Normalization as Recommended by Luminex. Inc

1. *Background correction*. The Median Fluorescence Intensities MFI of the background control sample (water treated in the same manner as an RNA sample throughout the experiment) are subtracted from the corresponding other sample intensities.
2. Figure 1 demonstrates this calculation using Excel.
3. *Inter-pool normalization*. One of the pools is selected as reference pool. The assay includes four normalization microspheres

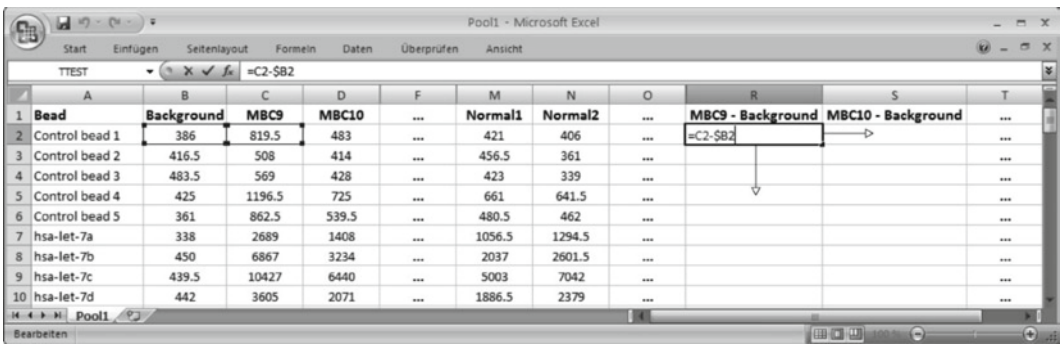


Fig. 1. Screenshot of an Excel sheet used for background correction.

in each pool. For each of the four normalization microspheres, the ratios between the reference pool signals and the other pool signals are determined. For a total of five pools including the reference this would return 4×4 ratios. For each of the four nonreference pools, the median of the four normalization ratios are determined, and the result is used as scaling factor for all signals of that pool.

4. *Intersample normalization*. Similar to the interpool normalization with the difference that the median normalization factors are determined within each sample and all samples are scaled one sample chosen as reference (see Note 2).

3.1.2.2. Median Centering like Normalization

1. *Signal scale adjustment*. All mean fluorescence intensity signals smaller 1 are set to a value of 1 and all values log2-transformed. The log2 scale features a more robust intensity range for further analysis.
2. *Background correction*. The Median Fluorescence Intensities MFI of the background control sample are subtracted from the corresponding other sample intensities.
3. *Interpool normalization*. Skipped (see Note 3).
4. *Inter-sample normalization*. To adjust for sample specific bias the median intensity in each sample is determined and subtracted from the sample signals.
5. For a better highlighting of the intensity differences over the samples the signals of each bead is median centered over all samples (see Note 4).

3.1.2.3. VSN Normalization

The following steps are performed on the R statistical platform. Signals need to be formatted in a tab-delimited file with the first column containing unique probe identifiers.

4. CODE

```
>library("vsn")
## here the signals will be read in from your tab-delimited file
>exp <- read.table("mysignalfilename.txt",sep="\t",header=T,row.
    names=1)
## normalization step
>exp.norm <- exprs(vsn(exp))
## write back to new file
>write.table(exp.norm, file="normalizedsignalsfile.txt",sep="\t",
    quote=FALSE)
```

Figure 2 shows a screenshot of the start of the normalization procedure using the statistical platform R.

Like Median centering VSN usually is a gentle normalization procedure. The advantage is that unlike median centering this procedure can correct noise affecting signal scaling not only signal shifts. Yet it should be remembered, that it assumes comparable variance across the samples. If due to experiment design that assumption does not hold, then probably a different normalization might perform better.

Subsequently, check the normalization effect with signal Box-Whisker plots:

```
## check the raw signals. The "las" flag just sets the type
## of axis-labellings to always perpendicular to the axis
>boxplot(exp,las=2)
## and now the normalized
>boxplot(as.data.frame(exp.norm),las=2)
```

The effects of this normalization procedure on our data set are illustrated in Fig. 3.

4.1. Statistical Analysis

Today, a variety of potential and freely available statistical tools and platforms for expression analysis exist. Some of them are easy to use such as TIGR (http://www.tm4.org/mev.html) or Cluster and TreeView (http://rana.lbl.gov/EisenSoftware.htm), while more experienced users might prefer more potent solutions as the R statistical platform (http://www.r-project.org) assisted by the packages of the BioConductor project (http://www.bioconductor.org).

Basically two general approaches can be distinguished, the supervised and the non-supervised analysis. While in a supervised

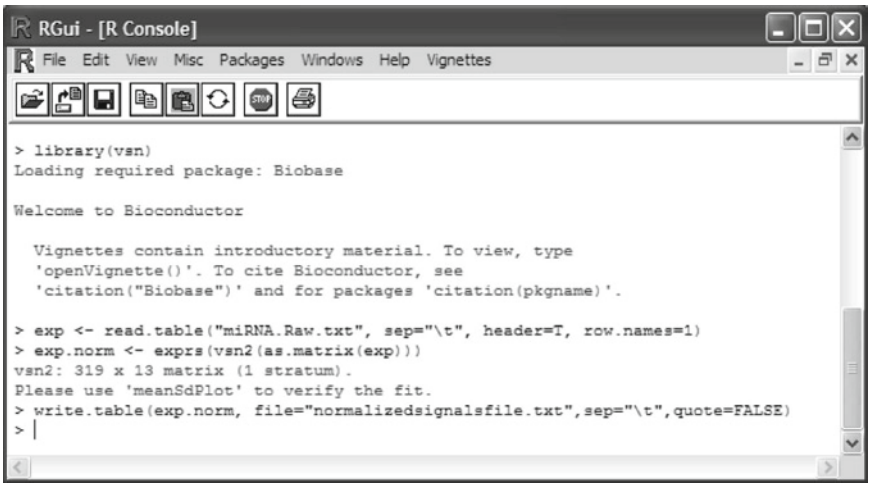


Fig. 2. Screenshot of the start of the normalization procedure using the statistical platform R.

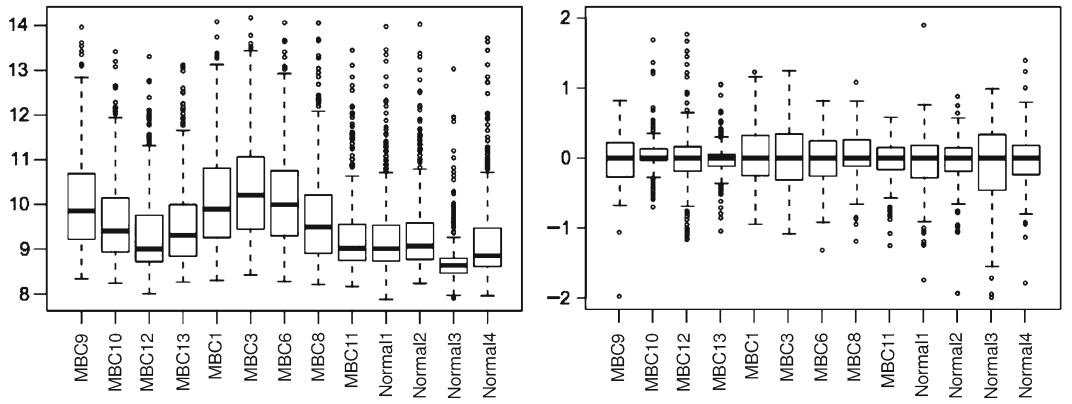


Fig. 3. Effect of the normalization on the distribution of the raw data (i.e., miRNA expression signals for 13 samples). Left: Raw signals. Right: Signals processed with median normalization.

analysis the statistical method is based on previous knowledge of the data composition, such as sample groupings, unsupervised methods provide an unbiased outlook. A typical representative of unsupervised analysis is hierarchical clustering, a method that shuffles samples and/ or genes according to signal similarity and might display the data relationships as dendrogram. Principal Component Analysis PCA or Self Organizing Maps are other approaches searching for relationships in the data. A typical method associated with supervised analysis is the simple hypothesis driven student's *t*-Test which can be used for differential expression search. But, there are a number of further approaches like Significance Analysis of Microarrays SAM or k-means clustering frequently used for gene expression investigations. Last but not the least, correlation methods such as Pearson correlation or Wilcoxon Signed-Rank Test can be applied to integrate different data sets or types, e.g., comparing expression values derived from array experiments with quantitative RT-PCR results.

4.1.1. Hierarchical Clustering

The application of a hierarchical clustering procedure is a quick and easy option for reviewing the relations in data. The method's name refers to the sequential merging of single clusters (samples most similar to each other) by building up a relational hierarchy. The measure used to quantify this relationship is called similarity distance. Typically used distances are Euclidean and Pearson correlation distance. While Euclidean distance weights differences in signals directly, the Pearson correlation distance values signal profiles across the samples.

When two clusters get merged to a new one, the distance of that cluster to all other clusters has to be calculated from the distances of the two former single clusters. This step is called linkage. The linkage is called complete linkage when the new cluster

distance is set to the maximum of the two old distances, single linkage when the minimum distance is used and average linkage when it is set to the average.

There is a variety of programs and tools that can be used for data clustering. Aside packages for the R statistical platform like "gplots" programs like "Custer" and TreeView from EisenLabs (http://rana.lbl.gov/EisenSoftware.htm) and the "MeV" module of the "TM4 Microarray Software Suite" are popular. The R platform usually requires coding or programming and thus is more suited for more experienced users. However after getting familiar with the platform it offers a large amount of freedom and tweaking possibilities. An easy example is provided below. The other three programs offer a nice self-explanatory user interface with easy handling. These are a good choice for the beginning. "Cluster" and "TreeView" build upon each other. While "Cluster" performs the clustering procedure, "TreeView" is provided to display the results (see Note 5).

4.1.1.1. Hierarchical Clustering with MeV

The easiest way to perform a hierarchical clustering is using MeV. The function is under "Clustering"→ "Hierarchical Clustering." In this menu, the distance metric and linkage method can be chosen.

Figure 4 displays a screenshot of the first window of the MeV module, which is used for performing the hierarchical clustering.

In Fig. 5, the result of the clustering of a subset of microRNAs and the samples is shown. "Cluster" version 2.11 and "TreeView" version 1.60 were used (using an uncentered correlation metric).

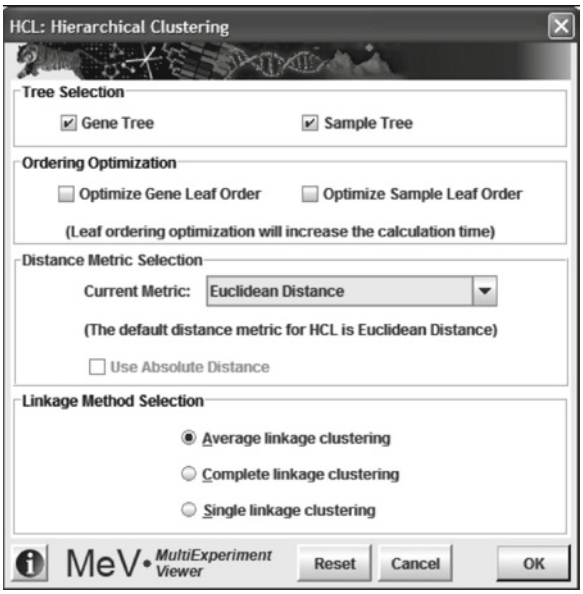


Fig. 4. Screenshot of the starting window of the module MeV used for hierarchical clustering.

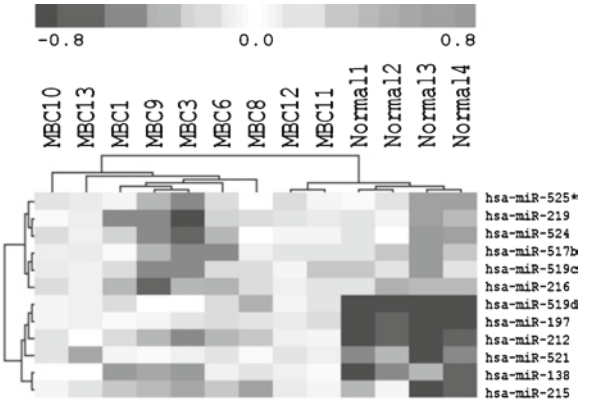


Fig. 5. Hierarchical clustering of a subset of miRNA expression values. The clustering was performed with the MeV module of the TM4 Microarray Software Suite using pearson correlation distance and average linkage.

4.1.2. Principal Component Analysis

Principal component analysis (PCA) is a mathematical transformation procedure identifying components in a provided vector of signals correlating with the major data variability. PCA is an unsupervised approach just as hierarchical clustering and thus is frequently used to investigate the composition of the data set signals. Plotting the main components on the coordinate axes, the relationship between data can be displayed as distance between set members in two- or three-dimensional space. Sample relations can be analyzed, as well as gene relations. Given that well-defined classes exist within the data, the main principal components, describing the biggest part of variation, should correlate with them. Graphically, they might separate these classes in two- or three-dimensional space. This analysis can be performed by several programs including R statistical platform, the MeV module, or the EisenLabs' Cluster program.

4.1.2.1. PCA with MeV

Principal component analysis can be calculated with the TM4 MeV module. The procedure can be found under "Data Reduction"→ "Principal Component Analysis". This analysis will require an additional installation of Sun Java 3D (11) for visualization.

4.1.3. Differential Expression with Welch t-Test

t-Test approaches are very common in differential expression analysis. The best known procedure is the Student's *t*-Test. Student's *t*-Test is a hypothesis-driven and parametric procedure comparing the mean values of two groups, taking the group variances into account. If the *p*-value calculated by the test falls below a given threshold, the difference between the mean values is regarded as significant and thus the two groups as different. The significance increases with increasing mean difference and decreasing variances. The reliability of the statistic increases with increasing group sizes.

The test is called hypothesis driven because the classification of the data into two groups hypothesizes some major differences between these two groups. The denotation "parametric" refers to the assumption underlying the Student's *t*-Test that the value distribution of the provided data follows certain standard parameters and thus an equal variance in both groups. These standard distribution assumptions might hold if the group sizes are big enough. Yet such dimensions are rather seldom in microarray experiments. To overcome this problem in differential expression analysis, usually, a derivated procedure called Welch Test is used. The Welch Test can handle unequal variances und hence much smaller groups (see Note 6).

Correction for multiple testing

In microarray analysis, a multitude of genes, miRNAs, or other probes are analysed simultaneously. The risk of finding a few significant candidates by pure chance rises with increasing number of analyzed probes, which is particularly often true with large gene expression arrays. To solve this problem, a *p*-value correction procedure such as Bonferroni, Benjamini-Hochberg or Holm, which should reduce the number of false significance assignments are appended to the initial test. They can, however, be very harsh in some cases and affect real significant candidates in turn.

Because of the usually small number of investigated probes in miRNA analyses, it is advisable to use a less-stringent correction like Benjamini-Hochberg, if any. An alternative easy way to restrict results to highly significant candidates is setting a minimum threshold for the difference between signal means.

4.1.3.1. Welch Test with MeV

In MeV, the *t*-Test is provided under Statistics-> *t*-Test. If needed (see above), the test type has to be set to Welch approximation (unequal variance).

Figure 6 displays the starting window of the MeV module used for theWelch *t*-Test calculation. Comments are inserted to guide the new user.

4.1.3.2. Welch Test with MS Excel

MS Excel can perform a *t*-Test by inserting the function:

=TTEST(A1:E1,F1:J1,2,3)

where the fields A1 to E1 are the values of group 1, and F1 to J1 are the values of group 2. The 2 as third parameter denotes a two-sided test and the fourth parameter "3" an unequal variance based test.

4.1.4. Differential Expression with SAM

Another common method used for differential expression is Significance Analysis of Microarrays by Tusher et al. (19). This is a statistical technique measuring the strength of the relationship between expression signals and the assigned groups by calculating a statistic d_i during several repeated assignment permutations.

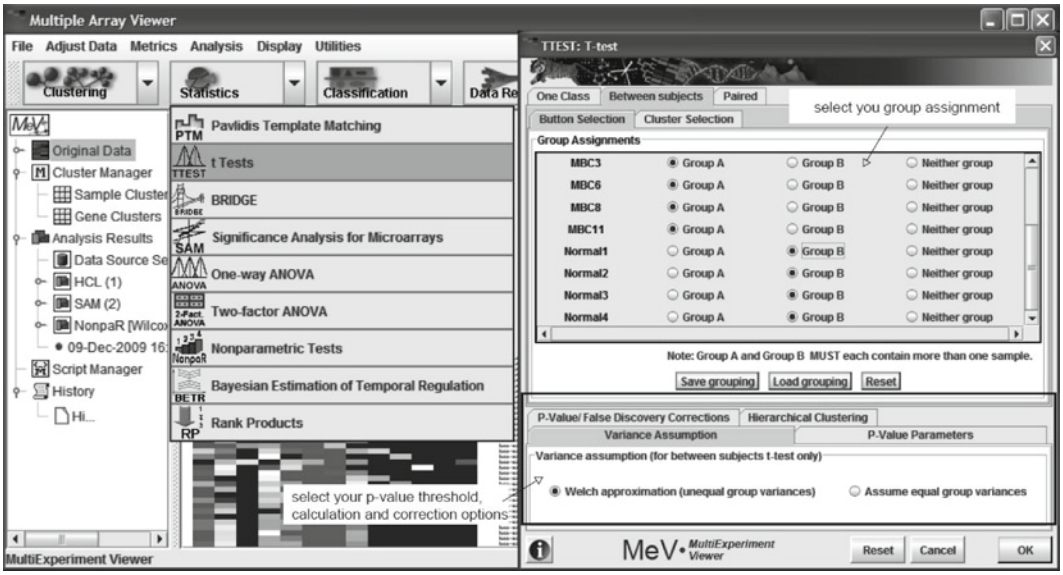


Fig. 6. Screenshot of the starting window of the MeV module used for performing Welch *t*-test calculations.

If the signal difference between groups is high and the variance within groups small, then this statistic will tend to grow. The cutoff for significance is determined by a parameter delta chosen based on the false positive rate. Significant candidates would have to pass this delta cutoff with the statistic d_i.

To speak from our experience, the application of different methods for differential expression analysis and checking the overlap between the results can enhance the interpretation. Highly significant candidates tend to be detected by most methods. However, it cannot conversely be assumed that candidates who are only detected by one method are insignificant.

4.1.4.1. Performing SAM with MeV

Performing a SAM analysis with MeV is quite easy. The function is found under "Statistics"→"Significance Analysis for Microarrays". The higher the number of permutations set, the more reliable the results are. But, the number of permutations is limited by the number of samples.

4.1.4.2. Performing SAM with MS Excel

There is a SAM macro for MS Excel. Instructions can be found in this manual (http://www-stat.stanford.edu/~tibs/SAM/).

4.2. Correlation of Data for Validation

There are several methods for correlating different data sources. One possible approach here is the application of a correlation test such as Pearson's correlation or Spearman's rank correlation test. Performing a linear regression would be another one. The difference between Pearson's correlation and Spearman's rank correlation is the data type used. The former uses the pure values, while

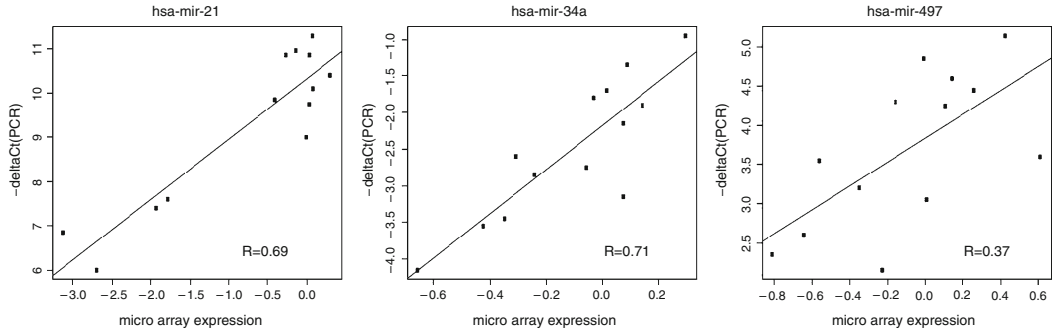


Fig. 7. Correlation of miRNA expression values calculated from the measurements using fluorescence-labeled beads with C_T-values of a quantitative real-time PCR measurement. Shown are as an example the data for three miRNAs. The miRNA C_T-values are normalized by subtraction from C_T-values of a reference gene measured along with the miRNA sample.

the latter digitalizes these values into their discrete ranks within the value set. This feature provides the Spearman's correlation with a higher robustness against outliers in bargain for less precision. If the number of samples in the experiment is small or the expected variability in the signals due to experiment design rather high, then a rank-based approach might be the safer choice. For example, the biological variance expected in human samples is much higher than in cell-culture samples.

Figure 7 shows correlation analysis for selected microRNAs using quantitative real-time PCR as a second independent method.

4.2.1. Correlation with MS Excel

There are limitations if MS Excel is used for correlation calculation. According to the author's knowledge, only Pearson's correlation is provided with the standard installation. The function "CORREL" can be used like this:

=CORREL(A1:E1,F1:J1)

where again group 1 is given by the fields A1 to E1 and group 2 by the fields F1 to J1.

4.3. Annotation of Data and Data Mining

Having investigated the data in the last steps, the analysis reaches a point where the results have to be annotated and further data mining must be performed. There is a variety of different online resources for accomplishing this task.

The most prominent resource is the *miRBase* of the Sanger Institute (8). This database provides information about miRNA annotations and tools for sequence and target search. The sequence database contains all published miRNA sequences, genomic locations, and associated annotations. The target database allows searching for predicted miRNA target genes.

The *ARGONAUTE* database (20) of the University of Heidelberg provides many tools for miRNA annotation, gene-based pathway and target miRNA information, information about

disease or organ associations with miRNAs, Motif searches, and further statistics (see Note 7). The "*microRNA*" database (21) is another site providing information about miRNAs, targets, and sequences, and in addition a tissue-based search database for miRNA expression. Similarly, the *TargetScan* database (http://www.targetscan.org) and *miRDB* (http://mirdb.org/miRDB/faq.html) allow mining for miRNA targets in a variety of organisms.

CORNA is a recently published package by Wu and Watson (22) for the R statistical platform, introducing a procedure for testing gene lists for regulation by miRNAs.

5. Notes

1. The advantage of R and BioConductor is the great potential and the active community. Unfortunately R requires at least some basic programming skills and some initial exercise. Therefore, some other programs might prove more user-friendly for beginners.
2. Spike-in and internal controls provide high potential for normalization. This holds true as long as the control signals are only affected by global variance influencing all measured probes and not by control probe specific bias. Experience has shown, however, that in some experiments internal controls, like all the other probes on an array, can sometimes be subject to nonglobal influences. If that is the case, IC-based normalization might be outperformed by other approaches. In general, a quick look at the single IC raw signals is recommended.
3. Because every plate of the Luminex Assay features different beads, it could be argued that a common median intensity on each plate would not be expected, and a correction would introduce more artificial noise than the plate specific bias. However, if this step is to be performed, for each pool, the global median over all beads and all samples could be determined. Then, the intensities are scaled such that the global median values converge.
4. Median centering is a gentle method without big demands to the data. As already mentioned above, it could be restricted to just centering samples, pools, or probes. Yet its simplicity can turn out to be a disadvantage if the stretching of signal scale is affected by artificial variance across samples. This variance would not be corrected; moreover, if the signal scales are only affected by different stretching factors, but not shifted, the median centering might introduce even more bias.

5. For the interested user, it is worth to have a look at the Coupled Two Way Clustering CTWC server of the Weizman Institute (http://ctwc.weizmann.ac.il/).
6. It should be stressed that for mathematical reasons, group sizes smaller than three samples definitely cannot be handled by such an approach and that the minimum group size recommended is always five samples.
7. During the preparation of this manuscript, the *ARGONAUTE* database was incorporated in a new internet-based resource named "miRWalk" (http://www.ma.uni-heidelberg.de/apps/zmf/mirwalk/). The validated target genes of the former ARGONAUTE database are now combined with in silico target gene prediction data.

Acknowledgements

The authors would like to thank Britta Hasemeier for indispensable help in preparing the figures.

References

1. O'Rourke, J. R., Swanson, M. S., and Harfe, B. D. (2006) MicroRNAs in mammalian development and tumorigenesis, *Birth Defects Res C Embryo Today 78*, 172–179.
2. Boyd, S. D. (2008) Everything you wanted to know about small RNA but were afraid to ask, *Lab Invest 88*, 569–578.
3. Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R., and Golub, T. R. (2005) MicroRNA expression profiles classify human cancers, *Nature 435*, 834–838.
4. Khoshnaw, S. M., Green, A. R., Powe, D. G., and Ellis, I. O. (2009) MicroRNA involvement in the pathogenesis and management of breast cancer, *J Clin Pathol 62*, 422–428.
5. Nelson, P. T., Baldwin, D. A., Scearce, L. M., Oberholtzer, J. C., Tobias, J. W., and Mourelatos, Z. (2004) Microarray-based, high-throughput gene expression profiling of microRNAs, *Nat Methods 1*, 155–161.
6. Lawrie, C. H. (2008) MicroRNA expression in lymphoid malignancies: new hope for diagnosis and therapy? *J Cell Mol Med. 12*, 1432–1444
7. Li, J., Smyth, P., Flavin, R., Cahill, S., Denning, K., Aherne, S., Guenther, S. M., O'Leary, J. J., and Sheils, O. (2007) Comparison of miRNA expression patterns using total RNA extracted from matched samples of formalin-fixed paraffin-embedded (FFPE) cells and snap frozen cells, *BMC Biotechnol 7*, 36.
8. Xi, Y., Nakajima, G., Gavin, E., Morris, C. G., Kudo, K., Hayashi, K., and Ju, J. (2007) Systematic analysis of microRNA expression of RNA extracted from fresh frozen and formalin-fixed paraffin-embedded samples, *Rna 13*, 1668–1674.
9. Wang, H., Ach, R. A., and Curry, B. (2007) Direct and sensitive miRNA profiling from low-input total RNA, *Rna 13*, 151–159.
10. Hoefig, K. P., Thorns, C., Roehle, A., Kaehler, C., Wesche, K. O., Repsilber, D., Branke, B., Thiere, M., Feller, A. C., and Merz, H. (2008) Unlocking pathology archives for microRNA-profiling, *Anticancer Res 28*, 119–123.
11. Zhang, X., Chen, J., Radcliffe, T., Lebrun, D. P., Tron, V. A., and Feilotter, H. (2008) An array-based analysis of microRNA expression comparing matched frozen and formalin-fixed paraffin-embedded human tissue samples, *J Mol Diagn 10*, 513–519.
12. Hasemeier, B., Christgen, M., Kreipe, H., and Lehmann, U. (2008) Reliable microRNA profiling in routinely processed formalin-fixed

paraffin-embedded breast cancer specimens using fluorescence labelled bead technology, *BMC Biotechnol 8*, 90.

13. Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data, *Biostatistics 4*, 249–264.
14. Blenkiron, C., Goldstein, L. D., Thorne, N. P., Spiteri, I., Chin, S. F., Dunning, M. J., Barbosa-Morais, N. L., Teschendorff, A. E., Green, A. R., Ellis, I. O., Tavare, S., Caldas, C., and Miska, E. A. (2007) MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype, *Genome Biol 8*, R214.
15. Lehmann U, Streichert T, Otto B, Albat C, Hasemeier B, Christgen H, Schipper E, Hille U, Kreipe HH, Länger F. (2010) Identification of differentially expressed microRNAs in human male breast cancer, *BMC Cancer*. Mar 23; 10:109.
16. Lehmann, U. (2009) MicroRNA profiling in formalin-fixed paraffin-embedded specimens. In: MicroRNAs and the Immune System, Ed.: S. Monticelli, Methods in Molecular Biology, Humana Press/Springer.
17. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Thorsen, T., Quist, H., Matese, J. C., Brown, P. O., Botstein, D., Eystein Lonning, P., and Borresen-Dale, A. L. (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications, *Proc Natl Acad Sci USA 98*, 10869–10874.
18. Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression, *Bioinformatics 18 Suppl 1*, S96–S104.
19. Tusher, V. G., Tibshirani, R., and Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response, *Proc Natl Acad Sci U S A 98*, 5116–5121.
20. Shahi, P., Loukianiouk, S., Bohne-Lang, A., Kenzelmann, M., Kuffer, S., Maertens, S., Eils, R., Grone, H. J., Gretz, N., and Brors, B. (2006) Argonaute – a database for gene regulation by mammalian microRNAs, *Nucleic Acids Res 34*, D115–D118.
21. Griffiths-Jones, S., Saini, H. K., van Dongen, S., and Enright, A. J. (2008) miRBase: tools for microRNA genomics, *Nucleic Acids Res 36*, D154–D158.
22. Wu, X., and Watson, M. (2009) CORNA: testing gene lists for regulation by microRNAs, *Bioinformatics 25*, 832–833.

INDEX

A

A549 human non-small cell lung cancer cells..... 147
Acute myeloid leukemias (AML)..... 181–183
Akt 42, 143–147, 150–151
Algorithms 23, 24, 28, 30, 31, 33, 36, 194, 199–201, 222, 223, 240, 250
AML. *See* Acute myeloid leukemias
AMO. *See* Anti-*microRNA* antisense oligodeoxyribonucleatide
Antagomir 63
Anti-*microRNA* antisense oligodeoxyribonucleotide (AMO) 41–46, 49–54, 61–63, 213–215
Anti-miRNA antisense 41, 45, 49–54
Antibody.....
anti-Akt..... 145
anti-β-actin..... 145
anti-EGFR..... 145
anti-HER2 ..... 64, 145
anti-phospho-Akt..... 145
anti-phospho-HER2 145
Antisense oligodeoxynucleotide (ASO)..... 44, 45
Apoptosis..... 42, 50, 57, 59, 61, 144, 182, 184, 186–187, 194
Argonaute (Ago) protein 9, 10, 82, 92, 93
Argonaute- 2, 4, 6–7, 10, 81–94
5-azacytidine 5, 166–169

B

Bayesian classifier 23
Bcl-2.....2, 42, 59
BioCarta database..... 223
Bioconductor 251, 255, 262
Biogenesis..... 1–12, 28, 30, 33, 58, 64, 81, 82, 92, 208, 239
Bioinformatics 32
Biosensor 107–116
Bisulphite modification 165, 176–178
Bladder cancer 161, 162, 168, 169
Breast cancer.....3, 32, 42, 50, 51, 59, 61, 64, 249, 250, 252

C

c-Myc 5, 10, 22, 59, 63
Cancer 1–12, 21–36, 42, 51, 52, 57–64, 69, 81, 82, 107, 108, 143–158, 161–179, 182, 194, 195, 209, 210, 216, 221, 222, 224
Cancer cell reprogramming 63–64
Cancer stem cells 58, 59, 63–64
Caspase- 3, 42, 184, 187, 189
Caspase- 7, 42, 184, 187, 189
CBF leukemia 183
Cell culture 70, 144–145, 149, 152, 154, 162–163, 167–169, 178, 185, 213, 243, 261
Cell viability 157, 182, 184, 187, 189
Coding gene 58, 60
Colorectal cancer (CRC).....32, 42, 194, 195
Computational approach 26, 120
Cycle threshold (Ct)..... 100, 216

D

Data normalization..... 250
DAVID..... 223
Di George syndrome Critical Region 8 (DGCR8)..... 5
Diana-MicroT.....194, 198–200
Dicer..... 4, 6–8, 10–12, 82, 193
Differentially expressed miRNAs (DE-miRs)..... 224, 226–235
Differentiation.....57, 59, 64, 107, 194
Dimethyl sulfoxide (DMSO) 85, 89, 171, 179, 187
DNA methylation 5, 167
DNA methyltransferase inhibitor.....162, 166, 167, 169
Drosha..... 4–7, 10–12, 27, 31, 36, 82
Drug cocktail..... 52
Dulbecco's modified Eagle's medium (DMEM)..... 70, 144, 162, 167, 169, 185, 187

E

EcoRI 186
Epidermal growth factor receptor (EGFR)..... 64, 143–147, 150–153
Epigenetic 3, 161–179
Ethidium bromide (EtBr) 83, 85, 87, 88, 122, 166

Wei Wu (ed.), *MicroRNA and Cancer: Methods and Protocols*, Methods in Molecular Biology, vol. 676,
DOI 10.1007/978-1-60761-863-8, © Springer Science+Business Media, LLC 2011

Etoposide 182, 184, 187
Exportin- 5, 4, 6, 82, 92
Expression data 23, 240, 242, 246
Expression profiling 59, 64, 98, 181
Extended dynamic range (XDR) 174

F

FGFR3 161
FlexmiR 250
Fluorescence-labeled beads 249–263

G

Gene expression 2, 44, 45, 50, 58, 64, 69, 107, 108, 119, 120, 149, 152, 154, 156, 157, 161, 163, 164, 173, 179, 184, 194, 207, 208, 250, 253, 256, 259
Gene prediction 24, 27, 30, 32, 33, 36, 240, 263
Glioblastoma 3, 32, 42, 61, 144, 147

H

Haemocytometric chamber 169
Hairpin 5, 7, 23–26, 30, 31, 82, 85, 92, 154, 164, 174, 193, 239, 243, 244
HCN2 ....43, 51, 209, 212
HCN4 ....43, 209
HEK293T cells 184, 185
HeLa cell line 42, 110, 115, 210, 243, 244
HepG2 cell line 243, 244
HER2 ....64, 143–147, 150–151
Hidden Markov models (HMMs) 23, 27, 28, 240, 242–244
Hierarchical clustering 186, 250, 252, 256–258
High-throughput 2, 23, 92, 97–106, 119–140, 223, 224, 240, 247
Human microRNA prediction 25, 239–247
Hybridization 69–78, 83, 84, 87, 88, 94, 98, 108, 111–115, 201, 216, 250, 252
Hypermethylation 162, 166, 167

I

Immunoblotting 144–151, 156, 157
In situ hybridization (ISH) 69–78, 216, 250
In vitro colony-forming/replating assay 183, 189
Induced pluripotent stem (iPS) 63, 64
IPA 223
Isopropanol 121, 125

J

JM109 high efficiency competent *E. coli* cells 135

K

KEGG pathway database 223
Keratinocyte serum-free medium (KSFM) 163
KIT 183

L

Label-free detection 111
Let- 7, 1, 3, 8, 10, 11, 22, 57, 59–61, 64, 69, 100–102, 104, 130, 132, 133, 208
Let-7a- 3, 5
Let-7b 105, 110, 112–115
Let-7c 105, 110, 112–114
Lin- 28, 11, 12
Lin- 4, 1, 32, 57, 69, 208, 212
Lipofectamine 2000 ....144, 147, 153, 155, 157, 214
LNA antimiR 63
LNCaP human prostate cancer 147, 154
Locked nucleic acid (LNA) 53, 61–63, 70, 73, 155, 213
Luciferase reporter gene assay 46, 147–149, 214
Luminex.Inc 253–254

M

MathCAD program 138, 139
Mature microRNA 2, 3, 5–10, 12, 26, 27, 31, 33, 58, 60, 61, 70, 81, 83, 86, 87, 89, 94, 98, 119, 154, 194, 199, 214, 216, 225, 233, 239, 244, 246
MCF-7 ....42, 51, 70, 75
MetaCore pathway 230
Metastasis 11, 58, 61
Microarray 70, 108, 162–165, 167, 171–175, 223, 251, 252, 256–260
Microprocessor 5, 11, 31, 36
Microprocessor SVM 31
MicroRNA (miRNA) 3, 23, 59, 85, 111, 123
- delivery 58, 63, 64
- library 120, 123, 124
- mimic therapy 61
- mimics 209
- miR-101 11
- miR-106a 32, 225–227, 235
- miR-125a 12
- miR-126 181–189
- miR-127 5
- miR-142 12
- miR-143 ....3, 32
- miR-145 ....3, 32, 70, 75
- miR-151 12
- miR-155 ....32, 51–53, 61, 62
- miR-18a 59
- miR-21 ....11, 42, 43, 52, 53, 59, 61, 62
- miR-29b 10
- miR-302 ....63, 64
- miR-31 3
- miR-331-3p 144, 146, 147, 151, 154, 156

miR-34 5
miR-430 43
miR-7 ..... 3, 144, 146, 147, 151–153
miR-99a/100 161
miR17-5p 51–53, 59, 62, 223
miR17-92 cluster 3, 5, 11, 22, 59
modulation 57–64
network 51, 64
precursor cloning 81–94
replacement therapy 60–61, 63
signaling pathways 144
sponges 50, 62
targets 2, 50, 51, 61, 62, 120, 149, 155, 156, 194, 195, 199, 200, 208, 214, 222–224, 226–231, 235, 261, 262
miR-21 inhibitor 11, 42, 43, 52, 53, 59, 61, 62
miR-mask 42–46
miRAnda 194, 198, 199
miRBase 57, 162, 163, 194, 198, 199, 225, 261
miRDB 262
miRDeep 30
miRNA induced pluripotent stem cells (mirPS) 58, 64
miRNA interference (miRNAi) 81–94, 208, 210
miRNA-masking antisense oligonucleotides 41–46
miRNA:mRNA interaction 44, 46, 58, 156, 208, 223
miRNome 58
mirPS. *See* miRNA induced pluripotent stem cells
miRRim 23, 27–29
MiRseeker 24, 25
MT-AMO. *See* Multiple-target AMO technology
MT-AMO$_{1/133}$ 51
MT-AMO$_{21/155/17}$ 51, 52
Multiple-AMO 50, 51, 53, 54
Multiple-target AMO technology (MT-AMO) 51–53

N

Nanofluidic 98
Nanoparticles 61, 108
Non-coding genes 2
Northern blots(ing) 26, 28–30, 32, 81–94, 108, 216, 250

O

Oncogenes 3, 5, 10, 11, 58, 60, 61, 70, 181, 209, 212, 221
OncomiRs 58, 59, 62
One drug with multiple-targets 52
OpenArrays 97–106
Opti-MEM reduced serum media 144
Optimem 147, 148
Oven 74, 76, 84, 87, 88, 122, 173

P

P-bodies 9, 10
P-element induced wimpy testis (PIWI) 82
P53 ..... 5, 11, 198
Paraffin-embedded tissue specimens 69–78
Pasha 5
PathwayStudio 223
PAZ domain 7, 8, 82
Peptide nucleic acids (PNAs) 108–111, 113–115
Phosphatidylinositol 3-kinase (PI3K) 144, 147
PI3K. *See* Phosphatidylinositol 3-kinase
PicTar 194, 198, 199
Plasmids 85, 91, 94, 134–137, 139, 144–145, 147–149, 152–155, 157, 182, 183, 186, 214
Polyacrylamide gel electrophoresis (PAGE) 120–122, 124, 127–131, 133–134, 139, 140
Polybrene 185, 188
Polycistronic transcription unit 2
Polymerase chain reaction (PCR) 70, 75, 85, 89–91, 98–100, 102–106, 108, 121–123, 130, 133–136, 139, 164–166, 172, 175, 177, 178, 184, 186, 216
Polymorphisms 193–205
Posttranscriptional regulation 3, 11, 12, 82
Pre-miRNA. *See* Precursors microRNA
Precursors microRNA (Pre-miRNA) 3–7, 10, 11, 24, 27, 28, 30, 31, 33, 36, 58, 82, 84–85, 88–91, 94, 98, 144–145, 147–149, 152, 154–157, 242, 244
Prehybridization 74, 76, 83, 87
Pri-miRNA. *See* primary microRNAs
Primary microRNAs (Pri-miRNAs) 2, 5, 6, 58
Principal component analysis (PCA) 250–252, 256, 258
Probes 71–74, 76, 78, 83, 87, 88, 94, 98, 100, 109, 113, 163, 164, 174, 188, 250, 252–254, 259, 262
Processing 3, 6, 7, 10–12, 31, 36, 70, 72, 74, 76, 81, 82, 100, 171, 208, 249–263
Prostate cancer 32, 64, 144, 147, 154
Protein activator of PKR (PACT) 6–7
Protein K 72, 76, 78, 165, 175
PTEN 42, 43, 60, 196
Puromycin 185, 188

Q

QIAquick column 178
QIAquick PCR purification kit 178
Quantitative PCR (qPCR) 98, 100, 106, 174–175, 182, 184–186

R

R statistical platform 251, 254, 255, 257, 258, 262
Rat1a cells 184, 187

Real-time PCR (RT-PCR) 70, 75, 98, 99, 147, 169, 174, 175, 182, 184, 186, 188, 211, 250, 256, 261
RECK 60
Recombinant epidermal growth factor (rEGF) 163
Reference miRNAs (Ref-miRs) 224, 225, 227, 229–232, 235
Retrovirus 184, 187
Reverse transcriptase 85, 102, 105, 133, 165, 174, 185
Reverse transcription (RT) 84–85, 98, 100, 102, 104–105, 123, 132–133, 164, 165, 174, 175, 185
RISC-loading complex (RLC) 6, 7
RMA 252, 253
RNA interference (RNAi) 3, 12, 81–94, 207, 208
RNAi-based drugs 3, 12
RNA-editing 11–12
RNA-induced silencing complex (RISC) 4, 6–9, 58, 82, 193, 207
RNAcofold 198, 201
RNase III endonuclease 5
RNase inhibitor 102, 105, 165, 175, 186
RNU44 ..... 164, 175
RNU48 175
RPMI-1640 ..... 70, 185
RT112 cells 162, 168, 179
RT4 cells 162, 168

S

SAM 256, 259–260
Sanger miRBase 163, 261
SDS-PAGE 145, 146, 149–150
Seed sequence 50, 194, 224, 225, 233
Self organizing maps 252, 256
Sequence detection system (SDS) 84, 88, 93, 145, 150, 175, 211
Short hairpin RNAs (shRNAs) 6, 10, 82, 85, 91–93
Silicon nanowire (SiNWs) 107–116
Single nucleotide polymorphisms (SNPs) 11–12, 108, 194, 198, 200–205
siRNA-based therapy 3, 64
SMAD proteins 11, 12
Small interfering RNA (siRNA) 3, 7–10, 54, 82, 85, 91–93, 155, 194, 207–210, 212–214, 216, 217
Sprouty2 60
SSCprofiler 23, 27–29, 32, 239–247
Stemness miRNAs 64

T

T-A cloning 134–135
T4 RNA ligase 85, 89, 123, 129, 130, 132, 164, 172, 179
Taq DNA polymerase 166, 177
Taqman® 98, 100, 102, 104, 106, 184
Tar RNA binding protein (TRBP) 6–7. 10
TargetScan 144, 198, 200, 223–226, 233–235, 262
TGF-β 11, 43
Theory of single-agent and multiple targets 51
THP-1 cells 182, 184–187
TIMP3 60
Total RNA abundance 123–124
Transfections 51, 54, 86, 92, 93, 144–149, 152, 154–157, 184–187, 189, 214, 215
Treeview 252, 255, 257
Tropomyosin 1 (TPM1) 60
Trypsin 144, 168, 169
Trypsin inhibitor 163
TSmiRs. *See* Tumor suppressors
Tumor suppressor miRNA (TSG miRNA) 50, 167
Tumor suppressors (TSmiRs) 3–5, 22, 42, 58, 59, 70, 181, 221
Tumorigenesis 10, 12, 22, 42, 61, 194, 209, 216
Tumors 3, 10, 50, 59, 61, 64, 70, 77, 81, 97, 120, 143, 147, 198, 221, 222, 224

U

U6 RNA 186
U87MG human glioblastoma 147
UCSC genome browser 162, 198, 200
Universal PCR master mix 165, 174, 175, 184, 186
Untranslated regions (UTRs) 199, 226
Uracil N-glycosylase (UNG) 165
Urea 83, 84, 86, 93, 127, 128, 130, 131, 134, 139
Urothelial cells 162
UV 87, 88, 103, 122, 129–131, 134, 136

V

Vector
  pcDNA3.1 85
  pGEM-T Easy 135
  pGL3-based firefly luciferase 147
  pMIR-REPORT-based firefly luciferase 147
  Renilla luciferase 147, 152, 154, 214
Vertical slide holder 100

Virus
adenovirus 54
lentivirus 54
Votex 129, 131, 135–137, 176, 177, 179

W

Web based service 223
Web tool 223
Welch test 259
Western blotting 150, 215
Wilms tumor 10
WT-1 10

X

XhoI 85, 91, 186, 214

Y

YM–50 Montage spin column 109, 110

Z

Zebrafish 43, 73, 242