

Commentary

Strengths and Limitations of Laboratory Procedures for MicroRNA Detection

Jill Koshiol¹, Ena Wang³, Yingdong Zhao², Francesco Marincola³, and Maria Teresa Landi¹

Abstract

Background: MicroRNAs (miR) are endogenous, noncoding RNAs involved in many cellular processes and have been associated with the development and progression of cancer. There are many different ways to evaluate miRs.

Methods: We described some of the most commonly used and promising miR detection methods.

Results: Each miR detection method has benefits and limitations. Microarray profiling and quantitative real-time reverse-transcription PCR are the two most common methods to evaluate miR expression. However, the results from microarray and quantitative real-time reverse-transcription PCR do not always agree. High-throughput, high-resolution next-generation sequencing of small RNAs may offer the opportunity to quickly and accurately discover new miRs and confirm the presence of known miRs in the near future.

Conclusions: All of the current and new technologies have benefits and limitations to consider when designing miR studies. Results can vary across platforms, requiring careful and critical evaluation when interpreting findings.

Impact: Although miR detection and expression analyses are rapidly improving, there are still many technical challenges to overcome. The old molecular epidemiology tenet of rigorous biomarker validation and confirmation in independent studies remains essential. *Cancer Epidemiol Biomarkers Prev*; 19(4); OF1-5. ©2010 AACR.

In 1993, Rosalind Lee, Rhonda Feinbaum, and Victor Ambros reported the discovery of a gene, *lin-4*, which coded for small RNAs rather than a protein (1). This discovery led to the identification of an entirely new class of RNA: microRNA (miR). Mature miRs are small, single-stranded RNAs ~22 nucleotides in length that are highly conserved across species (2). By degrading mRNA transcripts or inhibiting protein translation, miRs negatively regulate gene expression for a variety of fundamental biological processes, such as apoptosis, development, differentiation, and proliferation (2, 3). It is estimated that miRs regulate ~30% of human genes (4), and miR dysregulation has been associated with the development and

progression of cancer (5, 6). In fact, the same miR can act as an oncogene in some tissues and as a tumor suppressor in others (5).

These discoveries have sparked a great deal of interest in miR research. Because of their unique posttranscription and protein translation regulatory functions, miRs are important epigenetic modulators. For example, because miRs can inhibit protein translation, gene expression may be high whereas the encoded protein expression is low (7). In addition, although mRNA is not stable in formalin-fixed paraffin-embedded tissue, miR expression profiles seem to correlate well between fresh and formalin-fixed paraffin-embedded samples, possibly due to the small size and resistance to RNA degradation (8-13) of miRs. Stable miRs have also been detected in serum, plasma, urine, and other biological fluids and may be associated with cancer (14-22). These features make miRs extremely attractive for epidemiologic research, where archived formalin-fixed paraffin-embedded tissue, blood, or other biological fluids is most often available.

MiRs have been evaluated through several different methods, and each method has its own limitations. Some of the most commonly used and promising methods are listed in Table 1. The cloning method was originally used to discover miRs, which were subsequently confirmed

Authors' Affiliations: ¹Division of Cancer Epidemiology and Genetics and ²Biometric Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, and ³Infectious Disease and Immunogenetics Section, Department of Transfusion Medicine, Clinical Center and Center for Human Immunology, NIH, Department of Health and Human Services, Bethesda, Maryland

Corresponding Author: Jill Koshiol, Infections and Immunepidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Boulevard, Room 7070, Rockville, MD 20852-7248. Phone: 301-402-9508; Fax: 301-402-0817. E-mail: koshiolj@mail.nih.gov

doi: 10.1158/1055-9965.EPI-10-0071

©2010 American Association for Cancer Research.

with Northern blot (1, 23-26). Although most new miRs are still discovered through cloning, these methods are time consuming, low throughput, and biased toward the discovery of highly abundant miRs (27, 28). Similarly, other miR profiling methods have benefits and limitations. For example, although *in situ* hybridization is low throughput and has limited sensitivity and specificity, it shows the cellular localization of miRs, which is useful for characterizing the biology of individual miRs (28, 29). More generally, methods that directly detect miRs have low sensitivity because of the extremely short sequences and relatively low copy numbers of miRs. For these reasons, methods that do not involve miR amplification require more input total RNA. However, methods that do use amplification can be error prone due to the extremely short and inflexible template characteristics and similarity in sequences within miR families. Amplified samples are also more greatly affected by handling errors (30).

Although there is currently no gold standard for measuring miR expression (29), oligonucleotide microarray (microchip) and quantitative real-time reverse-transcription PCR (qRT-PCR) are two of the most common methods for evaluating known miRs (7, 27, 29-31). Although some studies of cancer cell lines (32, 33) or human tissue (34) found good correlation between microarray and qRT-PCR for selected miRs, one recent study compared semihigh-throughput microarray and qRT-PCR in proliferating murine myoblast cells and concluded that there was low correlation across platforms (27). Similarly, we found poor overall correlation between microarray- and qRT-PCR-based miR expression in 49 samples from lung cancer cases in the Environmental And Genetics in Lung cancer Etiology population-based case-control study. Microarray and qRT-PCR miR expression were significantly correlated for only 4 of 9 (44%) human miRs evaluated. Other studies of miR expression in cancer have also reported a relatively poor replication of microarray miR expression by qRT-PCR (35-37), and studies with 100% validation often report only one to three miRs (38-43).

There are several reasons why the results from qRT-PCR and microarray might differ. First, the larger dynamic range of stem-loop qRT-PCR (seven logs versus three to four logs for microarray) may provide greater sensitivity (27). qRT-PCR may also have higher specificity compared with microarray in distinguishing miRs with bases that differ at the 3'-end because stem-loop primers can distinguish between miRs that differ by one nucleotide (27, 44). In addition, because miRs vary slightly in length and guanine-cytosine (GC) content, they have different melting temperatures (30). Yet all miR probes on a microarray must undergo the same hybridization conditions because they are all on the same microchip. These homogenized hybridization conditions can lead to sequence-dependent differential hybridization affinities that may result in either false positives due to nonspecific hybridization or false negatives due to hybridization signals that do not exceed the background

threshold (32). Dual-channel (color) hybridization is less affected by this limitation than single-channel hybridization because the ratio between the two channels is used for data analysis rather than signal intensity. On the other hand, qRT-PCR relies exclusively on the success of cDNA synthesis, which is initiated by a stem-loop primer primed to short sequences at the 3'-end of the miR. Failure to initiate cDNA synthesis could result in false negatives. Readers should take care when reviewing older studies because stem-loop primers designed based on older versions of miR sequence databases, such as miRBase 9.2 (45), may not correctly prime to natural miR sequences due to inaccuracies in miR sequences from earlier versions compared with the current version, miRBase 14 (46). However, most modern commercially available stem-loop primers are designed based on later versions of the miRBase. In addition, qRT-PCR requires extreme care to avoid contamination or other technical errors and can produce variable results even in expert laboratories, suggesting that it is not the ideal gold standard (29).

It is considered good practice to profile miRs by microarray followed by validation with qRT-PCR (5). However, there are no standard guidelines for conducting and reporting such validation. For example, some authors report validation by qRT-PCR for some miRs and by Northern blot for other miRs, or report validation of precursor miRs but not mature miRs, without any explanation as to why these tests were chosen (47-49). In addition, when authors report that a few miRs were validated by qRT-PCR, it is often unclear if other miRs were also tested but not validated by qRT-PCR. Standardized guidelines would aid the interpretation of miR data by creating transparency in reporting. Furthermore, relative quantification of miR expression by qRT-PCR depends on the small nRNA used as an internal control. There is no standard as to which internal control should be used for the normalization of qRT-PCR data, and inappropriate normalization can result in erroneous conclusions (50). Clarity in describing how standardization controls are chosen would also aid data interpretation.

Because the full complement of human miRs has not been ascertained (29), platforms such as microarray and qRT-PCR that can only identify known sequences are limited. Emerging sequencing technologies provide a new discovery approach and have already been used to study small RNA, of which miR is one of the main components. Next-generation high-resolution deep sequencing allows both discovery of new miRs and confirmation of known miRs (7) in a high-speed, high-throughput fashion without the need for gels (51) or the ambiguity in data interpretation inherited by other methods. These new methods primarily include three platforms: the Roche (454) Genome Sequencer, which uses pyrosequencing to simultaneous sequence over 1 million reads in excess of 400 bp (52); the Illumina (Solexa) Genome Analyzer, which uses sequencing-by-synthesis to produce ~200 million 36- to 100-bp reads

Table 1. Characteristics of miR profiling methods

Technology	Throughput	Sensitivity	Specificity	Use*	Cost	Time requirement	Limitations
Cloning	Low	Low	High	Discovery and confirmation	High/miR	High	Need sequence confirmation. Limited quantitative ability. High cost and time.
<i>In situ</i> hybridization	Low	Low	Low	Confirmation	High/miR	1 d [†]	High background, low sensitivity and specificity. Can only detect high abundance miRs.
Microarray (oligonucleotide microchip)	High	Low	Low for subfamily	Confirmation	Low/miR	1 d	Requires 0.2-2 µg total RNA. Potential cross-hybridization of related miRs. Can only measure relative abundance.
Bead-based flow cytometry	High	Medium	High	Confirmation	Low/miR	Low	High complexity: requires removal of genomic DNA and recovery of small RNA, amplification, hybridization, and flow cytometry. More prone to external contamination due to amplification.
Northern blot	Low	Relatively low	High	Confirmation	Low/miR	High	High complexity: requires many labor-intensive steps, including radiolabeled oligonucleotide probes. Not all laboratories certified to handle radioactive probes. Requires a large amounts of total RNA (5-25 µg/sample).
qRT-PCR	Semihigh	High	High	Confirmation	High/miR	1 d	Primer for cDNA is based on complementarity over short sequences at the 3'-end. More prone to external contamination due to amplification.
Next generation sequencing	High	High	High	Discovery and confirmation	Low/miR, High/sample	2-5 d	High complexity: requires several gel purifications. Requires at least 2-10 µg very high quality total RNA and costly, specialized equipment. Potential underrepresentation of lower copy miRs.

*Discovery of new miRs and/or confirmation of the presence of already known miRs.

[†]Requires 1 d for testing after the assays have been optimized.

(53); and the Applied Biosystems SOLiD system, which uses sequencing by oligo ligation and detection to produce 400 million 50-bp reads (54).

In brief, these methods determine the nucleotide sequence by taking a picture every time a new nucleotide is added to the growing strand, thus emitting light (51). To ensure sufficient light signal intensity for accurate detection of each added nucleotide, these methods typically amplify the fragments through emulsion PCR or library generation followed by PCR-based cluster amplification. However, amplification can result in sequence errors and some sequences may be preferentially amplified, limiting the ability to accurately quantify relative abundance. These methods can also be less accurate in areas of homopolar (identical) bases. New techniques to read the sequence derived from a single molecule are currently under development. Limitations of next-generation sequencing include bioinformatic challenges due to large quantities of data and the high cost of instruments and reagents, although each sample can be bar coded, allowing samples to be mixed and run simultaneously to reduce cost. The third generation of sequencing technologies currently under development could eventually provide lower cost options (51, 55).

References

- Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 1993;75:843–54.
- Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004;116:281–97.
- Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev* 2006;20:515–24.
- Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 2005;120:15–20.
- Croce CM. Causes and consequences of microRNA dysregulation in cancer. *Nat Rev Genet* 2009;10:704–14.
- Esquela-Kerscher A, Slack FJ. Oncomirs—microRNAs with a role in cancer. *Nat Rev Cancer* 2006;6:259–69.
- Creighton CJ, Reid JG, Gunaratne PH. Expression profiling of microRNAs by deep sequencing. *Brief Bioinform* 2009;10:490–7.
- Lawrie CH, Soneji S, Marafioti T, et al. MicroRNA expression distinguishes between germinal center B cell-like and activated B cell-like subtypes of diffuse large B cell lymphoma. *Int J Cancer* 2007;121:1156–61.
- Li J, Smyth P, Flavin R, et al. 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* 2007;7:36.
- Nelson PT, Baldwin DA, Kloosterman WP, Kauppinen S, Plasterk RH, Mourelatos Z. RAKE and LNA-ISH reveal microRNA expression and localization in archival human brain. *RNA* 2006;12:187–91.
- Nelson PT, Baldwin DA, Searce LM, Oberholtzer JC, Tobias JW, Mourelatos Z. Microarray-based, high-throughput gene expression profiling of microRNAs. *Nat Methods* 2004;1:155–61.
- Wang H, Ach RA, Curry B. Direct and sensitive miRNA profiling from low-input total RNA. *RNA* 2007;13:151–9.
- Xi Y, Nakajima G, Gavin E, et al. Systematic analysis of microRNA expression of RNA extracted from fresh frozen and formalin-fixed paraffin-embedded samples. *RNA* 2007;13:1668–74.
- Chen X, Ba Y, Ma L, et al. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res* 2008;18:997–1006.
- Huang Z, Huang D, Ni S, Peng Z, Sheng W, Du X. Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer. *Int J Cancer* 2009. E-pub doi: 10.1002/ijc.25007.
- Wang J, Chen J, Chang P, et al. MicroRNAs in plasma of pancreatic ductal adenocarcinoma patients as novel blood-based biomarkers of disease. *Cancer Prev Res (Phila Pa)* 2009;2:807–13.
- Hanke M, Hoefig K, Merz H, et al. A robust methodology to study urine microRNA as tumor marker: microRNA-126 and microRNA-182 are related to urinary bladder cancer. *Urol Oncol* 2009. E-pub doi: 10.1016/j.urolonc.2009.01.027.
- Xie Y, Todd NW, Liu Z, et al. Altered miRNA expression in sputum for diagnosis of non-small cell lung cancer. *Lung Cancer* 2010;67:170–6.
- Habbe N, Koorstra JB, Mendell JT, et al. MicroRNA miR-155 is a biomarker of early pancreatic neoplasia. *Cancer Biol Ther* 2009;8:340–6.
- Gillad S, Meiri E, Yogev Y, et al. Serum microRNAs are promising novel biomarkers. *PLoS ONE* 2008;3:e3148.
- Lodes MJ, Caraballo M, Suci D, Munro S, Kumar A, Anderson B. Detection of cancer with serum miRNAs on an oligonucleotide microarray. *PLoS ONE* 2009;4:e6229.
- Mitchell PS, Parkin RK, Kroh EM, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A* 2008;105:10513–8.
- Grishok A, Pasquinelli AE, Conte D, et al. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 2001;106:23–34.
- Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. *Science* 2001;294:853–8.
- Lau NC, Lim LP, Weinstein EG, Bartel DP. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 2001;294:858–62.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Dr. Aaron Schetter for his critical review of this manuscript.

Received 01/15/2010; revised 02/02/2010; accepted 02/03/2010; published OnlineFirst 03/23/2010.

26. Lee RC, Ambros V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 2001;294:862–4.
27. Chen Y, Gelfond JA, McManus LM, Shireman PK. Reproducibility of quantitative RT-PCR array in miRNA expression profiling and comparison with microarray analysis. *BMC Genomics* 2009;10:407.
28. Li W, Ruan K. MicroRNA detection by microarray. *Anal Bioanal Chem* 2009;394:1117–24.
29. Nelson PT, Wang WX, Wilfred BR, Tang G. Technical variables in high-throughput miRNA expression profiling: much work remains to be done. *Biochim Biophys Acta* 2008;1779:758–65.
30. Kong W, Zhao JJ, He L, Cheng JQ. Strategies for profiling microRNA expression. *J Cell Physiol* 2009;218:22–5.
31. Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nat Rev Cancer* 2006;6:857–66.
32. Beuvink I, Kolb FA, Budach W, et al. A novel microarray approach reveals new tissue-specific signatures of known and predicted mammalian microRNAs. *Nucleic Acids Res* 2007;35:e52.
33. Mattie MD, Benz CC, Bowers J, et al. Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies. *Mol Cancer* 2006;5:24.
34. Ach RA, Wang H, Curry B. Measuring microRNAs: comparisons of microarray and quantitative PCR measurements, and of different total RNA prep methods. *BMC Biotechnol* 2008;8:69.
35. Avissar M, Christensen BC, Kelsey KT, Marsit CJ. MicroRNA expression ratio is predictive of head and neck squamous cell carcinoma. *Clin Cancer Res* 2009;15:2850–5.
36. Baffa R, Fassan M, Volinia S, et al. MicroRNA expression profiling of human metastatic cancers identifies cancer gene targets. *J Pathol* 2009;219:214–21.
37. Mees ST, Mardin WA, Wendel C, et al. EP300-a miRNA-regulated metastasis suppressor gene in ductal adenocarcinomas of the pancreas. *Int J Cancer* 2010;126:114–24.
38. Chow TF, Youssef YM, Lianidou E, et al. Differential expression profiling of microRNAs and their potential involvement in renal cell carcinoma pathogenesis. *Clin Biochem* 2010;43:150–8.
39. Huang ZM, Yang J, Shen XY, et al. MicroRNA expression profile in non-cancerous colonic tissue associated with lymph node metastasis of colon cancer. *J Dig Dis* 2009;10:188–94.
40. Izzotti A, Calin GA, Arrigo P, Steele VE, Croce CM, De Flora S. Downregulation of microRNA expression in the lungs of rats exposed to cigarette smoke. *FASEB J* 2009;23:806–12.
41. Izzotti A, Calin GA, Steele VE, Croce CM, De Flora S. Relationships of microRNA expression in mouse lung with age and exposure to cigarette smoke and light. *FASEB J* 2009;23:3243–50.
42. Lebanony D, Benjamin H, Gilad S, et al. Diagnostic assay based on hsa-miR-205 expression distinguishes squamous from non-squamous non-small-cell lung carcinoma. *J Clin Oncol* 2009;27:2030–7.
43. Li J, Huang H, Sun L, et al. MiR-21 indicates poor prognosis in tongue squamous cell carcinomas as an apoptosis inhibitor. *Clin Cancer Res* 2009;15:3998–4008.
44. Chen C, Ridzon DA, Broomer AJ, et al. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* 2005;33:e179.
45. Available from: <ftp://mirbase.org/pub/mirbase/9.2/>.
46. Available from: <http://www.mirbase.org/ftp.shtml>.
47. Calin GA, Liu CG, Sevignani C, et al. MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc Natl Acad Sci U S A* 2004;101:11755–60.
48. Sorrentino A, Liu CG, Addario A, Peschle C, Scambia G, Ferlini C. Role of microRNAs in drug-resistant ovarian cancer cells. *Gynecol Oncol* 2008;111:478–86.
49. Yang H, Kong W, He L, et al. MicroRNA expression profiling in human ovarian cancer: miR-214 induces cell survival and cisplatin resistance by targeting PTEN. *Cancer Res* 2008;68:425–33.
50. Peltier HJ, Latham GJ. Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA* 2008;14:844–52.
51. Ansorge WJ. Next-generation DNA sequencing techniques. *N Biotechnol* 2009;25:195–203.
52. Available from: <http://www.454.com/products-solutions/system-features.asp>.
53. Available from: <http://www.illumina.com>.
54. Available from: http://www3.appliedbiosystems.com/AB_Home/applicationstechnologies/SOLiDSystemSequencing/overviewofsolidssystem/index.htm.
55. Mardis ER. Anticipating the 1,000 dollar genome. *Genome Biol* 2006;7:112.