Mol Biol Rep (2011) 38:4125–4135
DOI 10.1007/s11033-010-0532-1

The discovery approaches and detection methods of microRNAs

Yong Huang · Quan Zou · Sheng Peng Wang · Shun Ming Tang · Guo Zheng Zhang · Xing Jia Shen

Received: 17 July 2010 / Accepted: 15 November 2010 / Published online: 25 November 2010

Abstract MicroRNAs (miRNAs) are small, highly conserved, non-coding RNAs that regulate gene expression of target mRNAs through cleavage or translational inhibition. Computer-based approaches for miRNA gene identification are being considered as indispensable in miRNAs research. Similarly, experimental approaches for detection of miRNAs are crucial to the testing and validating of computational algorithms. The detection of miRNAs in tissues or cells can supply valuable information for investigating the biological function of these molecules. Selective and highly sensitive detection methods will pave the way for extended understanding of miRNA function within organisms. In this review, we summarize the various computational methods for identification of miRNAs as well as the methodologies that have been developed to detection miRNAs.

Y. Huang · S. P. Wang · S. M. Tang · G. Z. Zhang · X. J. Shen
Jiang Su University of Science and Technology, Zhenjiang 212018, Jiangsu, People's Republic of China

Q. Zou
School of Information Science and Technology of Xiamen University, Xiamen 361005, Fujian, People's Republic of China

Y. Huang · S. P. Wang · S. M. Tang · G. Z. Zhang · X. J. Shen (✉)
The Key Laboratory of Silkworm and Mulberry Genetic Improvement, Ministry of Agriculture, Sericultural Research Institute, Chinese Academy of Agricultural Sciences, Zhenjiang 212018, Jiangsu, People's Republic of China
e-mail: shenxj63@yahoo.com.cn

Keywords MicroRNAs · Computational methods · Experimental detection

Introduction

miRNAs are endogenously encoded and produced small (~22 nucleotides) non-coding RNA molecules, which negatively regulate expression of target genes at the post-transcriptional level [1, 2]. miRNAs are initially transcribed as longer primary transcripts (pri-miRNAs) and processed first by the RNase enzyme complex, and then by Dicer, leading to incorporation of a single strand into the RNA-induced silencing complex (RISC) [3]. miRNAs guide RISC to mRNAs that results in post-transcriptional repression. Recently, numerous studies have demonstrated that many miRNAs are involved in the regulation of gene expression through the targeting of mRNAs during cell proliferation, apoptosis, the control of stem cell self renewal, differentiation, metabolism, development, and tumor metastasis [4, 5]. Up to date, more than 10,000 miRNAs have been reported in animals, plants and viruses by using computational and experimental detected methods in miRNAs database (http://www.mirbase.org/) [6]. Since miRNA has the ability to regulate many vital biological functions, it formulates a plethora of opportunities for miRNA research [7].

The identification of new miRNAs is the foundation of miRNA research. This led to the development of increasingly sophisticated computational prediction approaches of miRNAs and the biological detection techniques [8]. An overview of the current advances in this area is presented here with a view to stimulate the reader to explore the diverse and exciting field of miRNA research.

Biochemical approaches to miRNA discovery

Classic forward genetics

Forward genetic experiments, whereby mutant genes are isolated from an organism showing abnormal phenotypic characteristics. The first known miRNA, lin-4, was discovered in 1993 by Victor Ambros and his colleagues through the study of heterochronic gene lin-14 in worms by classic forward genetics. They discovered that lin-4 in *C. elegans* did not code for a protein but instead produced a pair of short RNA transcripts that each regulate the timing of larval development by translational repression of lin-14, which encodes for a nuclear protein [9]. They postulated the regulation was due in part to sequence complementarity between lin-4 and unique repeats within the 3′ UTR of the lin-14 mRNA. In 2000, almost 7 years after the second miRNA, let-7, was discovered, also using forward genetics in worms. The let-7 encodes a temporally regulated 21-nucleotide small RNA that controls the developmental transition from the fourth larval into the adult stage [10–12]. Similar to lin-4, let-7 performs its function by binding to the 3′ UTR of lin-41 and hbl-1 (lin-57), and inhibiting their translation [13, 14]. The lin-4 and let-7 are nonhomologous and act in a similar manner to trigger the transition to late-larval and adult stages. However, this method is also tedious, time-consuming and costly for miRNA discovery, primarily owing to the enormous effort required to identify each gene responsible for a particular phenotype [15].

MiRNA cloning

Subsequently, a breakthrough in miRNA identification was made when directional cloning was used to construct a cDNA library for endogenous small RNAs [16]. Direct cloning and sequencing of short RNA molecules has enabled the identification of many miRNAs. Several labs identified more than a hundred miRNAs by cDNA cloning and subsequent sequence analysis [16, 17]. This approach has the advantage that it can be applied to any organism, even if little or no genomic information is available. In addition, miRNAs can be identified independent of their function, thus also allowing the identification of redundant miRNAs. By cDNA cloning, miRNAs have now been identified in diverse animals, plants and viruses [18–22].

The miRNA cloning procedure requires size selection of 18–24 nt RNA purified by polyacrylamide gel electrophoresis (PAGE) (Fig. 1). The small RNA is to enable directional ligation of an adapter sequence, first to the 5′ end of the RNA, and, after PAGE purification of ligated products and rephosphorylation of the 3′ terminus, ligation of a second adapter to the 3′ end. The resultant miRNA

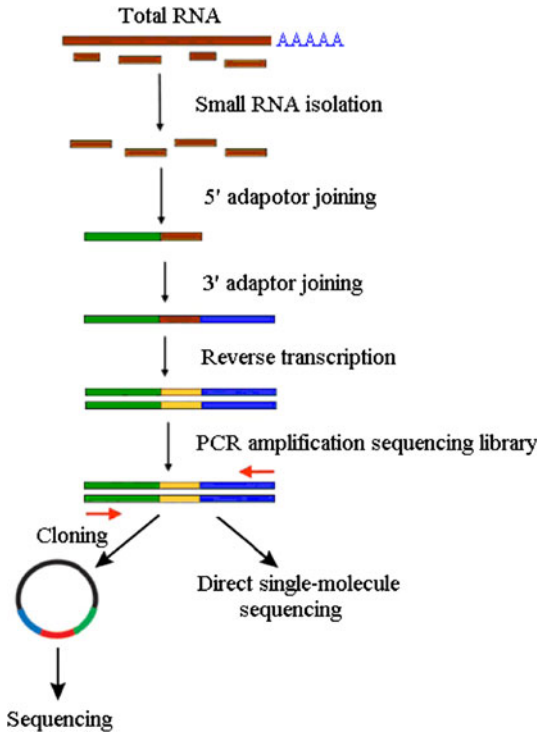

Fig. 1 Methods for cloning miRNA

sequence is inserted between two known sequences and can be amplified by standard RT-PCR from a cDNA library enriched for small RNAs [23]. Individual clones are then sequenced and analyzed to determine the genomic origin of the small RNA. To prevent self-circularization of the mature miRNAs and the adapter, small RNAs are usually dephosphorylated before ligation and the 3′-hydroxyl terminus of the 3′ adapter is blocked by incorporating a non-nucleotide group during chemical synthesis of the oligonucleotide [24]. A limitation of this approach, however, is that miRNA expressed at a low level or only in a specific condition or specific cell types would be difficult to find.

Next generation sequencing

Direct cloning, on the other hand, may not detect miRNAs that have very low expression levels or that are expressed only in specific conditions and tissue types. A new generation of sequencing technologies (also called deep sequencing), from Illumina/Solexa, ABI/SOLiD, 454/Roche, can detect many small RNAs with a higher degree of reliability and has provided unprecedented opportunities for high-throughput detection of miRNAs [25]. These technologies produce a number of sequence fragments in the 20–300 base pair range and open the door to detecting and profiling known and novel miRNAs at unprecedented sensitivity [26]. Their applications in the field have also included gene expression profiling, genome annotation, and detection of aberrant

transcription [27]. Current three main technologies for discovery of miRNAs sequencing platforms are pretty different and summarized in Table 1.

For example, the Illumina/Solexa technology is the most successful and widely-adopted next-generation sequencing platform worldwide. It supports massively parallel sequencing using a proprietary reversible terminator-based method that enables detection of single bases as they are incorporated into growing DNA strands. A fluorescently-labeled terminator is imaged as each dNTP is added and then cleaved to allow incorporation of the next base. The Illumina/Solexa utilizes a unique "bridged" amplification reaction that occurs on the surface of the flow cell [28]. In this process, one end of single DNA molecule is attached to a solid surface using an adapter; the molecules subsequently bend over and hybridize to complementary adapters (creating the "bridge"), thereby forming the template for the synthesis of their complementary strands. After the amplification step, a flow cell with more than 40 million clusters is produced, where the cluster station amplifies DNA on the flow cell surface to create clusters containing 500–1,000 clonal copies of each molecule (Fig. 2). Single-stranded, adapter-ligated fragments are bound to the surface of the flow cell exposed to reagents for polyermase-based extension. The first cycle of sequencing consists first of the incorporation of a single fluorescent nucleotide, followed by high resolution imaging of the entire flow cell. These images represent the data collected for the first base. Any signal above background identifies the physical location of a cluster (or polony), and the fluorescent emission identifies which of the four bases was incorporated at that position. And then, genome analyzer is capable of generating 26–50 bases reads and producing at least 1 Gb of sequence per run in 2–3 days [29, 30].

The Illumina/Solexa technology was recently used to sequence small RNA libraries from human embryonic stem cells before and after their differentiation into embryonic bodies [31]. The results generated more than six million short sequence reads from each library and identified 334 known and 104 novel miRNA genes in one of the most comprehensive miRNA profiling exercises to date [31]. The drawbacks in the use of this platforms sequencing include the extensive amount of data necessary to organize from a collection, because each sequence read can align to multiple areas on the genome even without consideration of a possible nucleotide error within the read [32, 33].

Computational approaches to miRNA discovery

The highly constrained tissue and time-specific expression patterns, presence of degradation products from mRNAs,

Table 1 Comparison of three sequencing platforms

Platform	Approach	Read length (bases)	Gb per run	Run time (days)	Company name	References
Illumina/Solexa	Sequencing by synthesis with reversible terminators	75–100	18–100	2–3	Illumina, Inc	[34]
ABI/SOLiD	Massively parallel sequencing by ligation	~50	1–3	8	Applied Biosystems	[35]
454/Roche	Pyrosequencing on solid support	200–300	0.08–0.12	0.4	Roche Applied Science	[36]

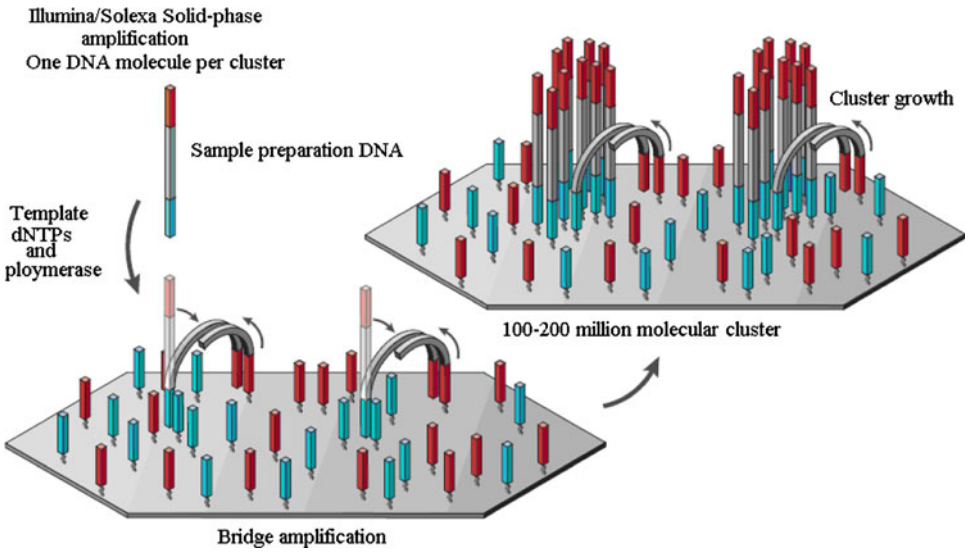

Fig. 2 Solid-phase amplification of the Illumina/Solexa approach

Table 2 Resources of the others computational approaches for miRNA predictions

Name of Program	Website	References
proMiR II	http://cbit.snu.ac.kr/~ProMiR2/	[47]
mir-abela	http://www.mirz.unibas.ch/cgi/pred_miRNA_genes.cgi	[48]
triplet-SVM	http://bioinfo.au.tsinghua.edu.cn/mirnasvm/	[49]
Vmir	http://www.hpi-hamburg.de/fileadmin/downloads/VMir.zip	[50]
RNA micro	http://www.bioinf.uni-leipzig.de/~jana/software/index.html	[51]
mirCoS	Based on LIBSVM library package	[52]
BayesMiRNAFind	https://bioinfo.wistar.upenn.edu/miRNA/miRNA/login.php	[53]
One-ClassMirnaFind	http://wotan.wistar.upenn.edu/OneClassmiRNA/	[54]
miPred	http://www.bioinf.seu.edu.cn/miRNA/	[55]
Srnaloop	http://arep.med.harvard.edu/miRNA/pgmlicense.html	[56]
findMiRNA	http://sundarlab.ucdavis.edu/mirna/	[57]

and other noncoding RNAs, has made it difficult and incomplete to clone miRNAs. Computational methods play important roles in the identification of new miRNAs [37]. This led to the development of increasingly sophisticated computational approaches to predict miRNAs. Computational approaches have been developed to complement experimental methods in discovery of miRNAs that express restrictively in specific environmental conditions or cell types. At present, several major computational finding methods of miRNAs were summarized below. The other approaches for finding miRNAs are summarized in Table 2.

MiRscan

Lim et al. [38] developed a method for the identification of miRNAs, called MiRscan to find miRNA genes conserved in more than one genome. Briefly, it slides a 110-nt window along both strands of the *C. elegans* genome and folding the window with RNAfold to identify predicted stem–loop structures (>25 bp) and a folding free energy of at least -25 kcal/mol. Each conserved hairpin, considered as a potential pre-miRNA, was further evaluated for the location of miRNA within it by passing a 21-nt window along each stem–loop, assigning a loglikelihood score to each position for its similarity to known miRNAs, the training set for the algorithm being 50 published miRNAs from *C. elegans* and *C. briggsae.* They identified 30 novel miRNAs in *C. elegans* and 38 novel human miRNAs [38]. The method uses an RNA folding algorithm RNAFold to locate potential hairpin structures in sequences that are evolutionarily conserved among *C. elegans* and *C. briggsae* [39]. Subsequently, the accuracy of MiRscan program was further improved by considering sequence conservation about 200 bp upstream of the miRNA fold-back and a highly significant sequence motif [40, 41].

miRSeeker

Lai et al. [42] using a detection program called miRseeker, estimated that *Drosophila* genomes contain around 110 miRNA genes and identified 48 miRNA candidates which does not simply use conservation but recognizes conservation patterns specific to miRNAs. The method represents the first attempt to identify conserved stem–loops due to selection, and not as an artefact of considering genomes that are not sufficiently distant. The authors aligned the non-annotated intergenic and intronic sequences of the genomes of *Drosophila*. The conserved regions were then folded in order to identify and score potential stem–loop structures. The miRseeker described so far and variations there of have been able to recover a substantial part of the known miRNAs and have been useful in identifying several new regulators [43]. The sensitivity of this method was also demonstrated by the presence of 75% (18 of 24) of previously identified *Drosophila* miRNAs within the top 124 candidates. In total, 48 novel miRNA candidates identified by miRseeker were strongly conserved in more distant insect, nematode, or vertebrate genomes. Of these, the expression of a total of 24 novel miRNA genes was experimentally verified. This approach effectively prevents the identification of non-conserved candidates, and makes several assumptions in the absence of a clear evolutionary model for these structures.

Phylogenetic shadowing

Berezikov et al. [44] identified 16 novel human miRNAs, using a phylogenetic-based approach. Phylogenetic shadowing is a powerful genomic-conservation assessment technique, which determines the level of conservation of each nucleotide in the assessed sequence [44]. Phylogenetic shadowing overcomes the limitation of classical

phylogenetic footprinting and allows unambiguous sequence alignments and accurate conservation determination at single nucleotide resolution level. Using this approach, Berezikov et al. found that nucleotides in the stem of miRNA hairpins precursors are significantly more conserved than in sequences flanking the hairpin, and in the hairpin_s loops. They then used this distinctive property, in conjunction with other known properties of miRNAs, as described above. The 976 candidate miRNAs by scanning whole-genome human/mouse and human/rat alignments, most of the novel miRNA candidates being conserved in other vertebrates (dog, cow, chicken, opossum, and zebrafish) were predicted. Northern blot analysis confirmed the expression of mature miRNAs for 16 out of 69 representative candidates.

miRank

Xu et al. propose and develop a novel miRNA prediction method, miRank, based on a random walk through a graph consisting of known miRNA examples and unknown candidate sequences [45]. Each miRNA is a vertex connected to its neighbor by an edge that is weighted by its similarity of miRNA features. The score or relevance of a vertex increases with its number of connections. The vertices are then ranked by relevance score, and an arbitrary cut-off of the ranked list includes both the positive examples and the most similar of the predicted unknowns. The strength of miRank is its ability to identify novel miRNAs in newly sequenced genomes where there are few annotated miRNAs (positive examples). The authors found miRank to be superior to SVM classifiers, and attribute its success to the fact that it structures the list and ranks the candidate examples, as well as the query sequences, during the training and classification steps. They first tested on *Homo sapiens* genome by their method; using a very few known human miRNAs as samples, the method achieved a prediction accuracy greater than 95%. They then applied this method to predict 200 miRNAs in *Anopheles gambiae*. The further study showed that 78 out of the 200 putative miRNA precursors encode mature miRNAs that are conserved in at least one other animal species.

miRDeep

miRDeep was developed which uses a probabilistic algorithm to score features of miRNA candidates by using bayesian statistics. First, the secondary RNA structure and energetic stability is investigated. Second, the compatibility between the sequenced RNA and miRNA biogenesis by enzymatic cleavage is scored. Third, since miRNAs are often conserved across large phylogenetic distances, candidate miRNA sequences can be compared with reference sequences or with candidate sequences from other animal species. It gets its accuracy and robustness from published *C. elegans* data and data generated by deep sequencing human and dog RNAs. miRDeep also reports altogether ~230 previously unannotated miRNAs, of which four novel *C. elegans* miRNAs are validated by northern blot analysis [25]. The miRDeep algorithm is that after the sequencing reads are aligned to the genome, the algorithm excises genomic DNA bracketing these alignments and computes their secondary RNA structure. Plausible miRNA precursor sequences are then identified and, in the core part of the miRDeep algorithm, scored for their likelihood to be real miRNA precursors. The output is therefore a scored list of known and novel miRNA precursors and mature miRNAs in the deep-sequencing sample, as well as estimates for the number of false positives.

MiRanalyzer

MiRanalyzer implements all necessary methods for a comprehensive analysis of deep-sequencing experiments of small RNA molecules. It detects all known miRNA sequences annotated in miRBase, finds all perfect matches against other libraries of transcribed sequences. The algorithm is based on the random forest classifier and was trained on experimental data. This high accuracy is important for the identification of novel miRNAs, a process which usually results in high false positive rates. MiRanalyzer predicts new miRNAs basing on a random forest machine learning algorithm and good performances are obtained (area under the curve [AUC] of 97.9%). Furthermore, the tool detects matches of the reads against other libraries of transcribed sequences [46].

In practice, all these prediction methods require a sufficient number of characterized miRNAs as training samples and rely on genome annotation to reduce the number of predicted putative miRNAs. Although each of these methods has their own unique advantages, they have not been perfected yet. No consensus has been reached on the problem of verification. Each miRNA prediction program has sought to verify its computational conclusions by differently methods of experimental detection.

Experimental detection of miRNAs

Experimental detection of miRNAs is technically challenging because of their small size, sequence similarity among various members, low level, and tissue-specific or developmental stage-specific expression. Each technique has its advantages along with some limitations over its application at the moment. A suitable method should be chosen based on the requirements of the investigation and

the experimental conditions in order to get accurate information on miRNA expression. An increase in the amount of detection methods available have been developed and applied in miRNAs exploration. The main conventional miRNA detection and new methods, have been summarized blow.

Northern blotting detection

Northern blotting is a technique for locating specific RNA sequences in a solution of mixed RNAs and a robust technique that can provide information on the size and expression of predicted miRNAs and precursor miRNAs. It is also frequently used to assess accumulation levels of miRNAs of interest from size-fractionated cDNA libraries [58].The sample containing miRNA is run on an electrophoresis gel. Next, the miRNA is transferred to a nitrocellulose membrane, followed by soaking in a solution containing a fluorescent or radiolabeled oligonucleotide probe which is complementary to the target miRNA for hybridization to occur. After unhybridized DNA has been removed, the miRNA target can be detected [59]. Although this method offers confirmation of the highest quality, however, it is time-consuming, low sensitivity and a large amount of total RNA per sample which is required. Locked nucleic acid (LNA)-modified oligonucleotide probes were used to enhance the efficiency of hybridization for solving the sensitivity problem of miRNA Northern blotting technology [60]. The technique can be extremely beneficial when small amounts of samples are available, expression levels of target miRNAs are low, or subtle discrimination of related miRNAs (differing only by a few nucleotides) is necessary. LNA probes exhibit unprecedented thermal stability and show improved hybridization properties against complementary RNA targets.

Microarray detection

miRNA microarray technology is actually based on nucleic acid hybridization between target molecules and their corresponding complementary probes. Liu et al. [61] published the first report of genome wide miRNA expression profiling by a microarray in human and mouse tissues. The microarray detection was probed with a radioisotope-labeled low molecular weight fraction (<60 nt) of total RNA. MiRNA oligonucleotide probes that usually have amine-modified 5′ termini are immobilized onto glass slides through covalent crosslinking between the amino groups and the SAM (self-assembling monolayer), forming a ready-to-use miRNA microarray. The isolated miRNAs are labeled with fluorescent dye and then hybridized with the miRNA microarray, resulting in specific binding of the labeled miRNAs to the corresponding probes. The fluorescence emission from labeled miRNAs bound at different positions on the slides can be detected. Consequently, the kinds of miRNAs and their relative quantities in the studied sample can be evaluated by analyzing the fluorescence signal data. However, small size of miRNAs poses difficulties for this technique, as it is difficult to create a single hybridization condition suitable for all miRNAs because of having positive non-specific hybridization. The recent developments in microarray analysis is involves more sensitive and specific methods [62]. The microarray platform miChip is based on LNA-modified, Tm-normalized capture probes spotted onto NHS (*N*-hydroxysuccinamide)-coated glass slides and can discriminate between miRNAs with single nucleotide differences [62].

In situ hybridization detection

In situ hybridization (IHS) for miRNA detection has been developed recently [63]. It may be used to locate the cellular and subcellular distribution and determine the spatio-temporal expression patterns of candidate miRNAs [63]. However, the normal DNA or RNA probes may not work well in this method owing to their poor binding affinity to target miRNA. To increase the affinity, Kloosterman et al. [64] introduced LNA into the ISH probes and successfully observed the expression of 115 conserved vertebrate miRNAs in zebrafish embryos by ISH. They found that most miRNAs were expressed in a tissue-specific manner during segmentation and the later stages, but not early in development, which demonstrated that their function was not involved in tissue fate establishment but in differentiation or maintenance of tissue identity [64]. Use of the probes successfully detected conserved vertebrate miRNAs in zebrafish, mice, and frog embryos [65]. Nelson et al. demonstrated coordinated miRNA expression on archival formalin-fixed paraffin-embedded (FFPE) human brains and oligodendroglial tumors by using RAKE (RNA-primed, array-based, Klenow Enzyme) miRNA microarray platform in conjunction with LNA-based in situ hybridization [66]. Deo et al. [67] have recently improved the specificity by using RNA oligonucleotide probes linked to a fluorescein hapten and highly specific washing conditions with tetra-methyl ammonium chloride (TMAC). The method could directly detect mature miRNAs in tissue sections from developing mouse embryos, adult brain, and the eye. ISH can precisely locate a specific miRNA within tissue, but it is not suitable for high-throughput profiling [68].

Invader miRNA assay

A called the invader miRNA assay method was developed for the sensitive and specific detection and quantitation of

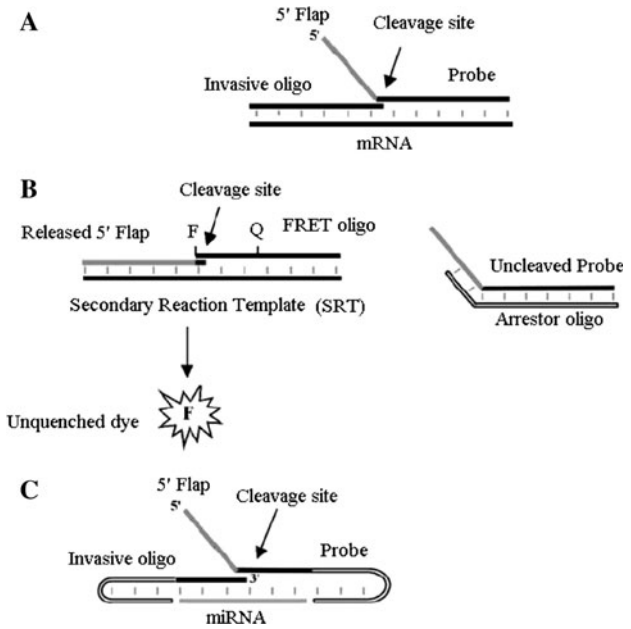

Fig. 3 Schematic representation of the invader miRNA assay

miRNAs or siRNAs [69]. The invader assay is based on enzymatic cleavage by a structure-specific 5′nuclease (Cleavase) of a synthetic oligonucleotide (the probe) that is in an appropriate overlap-flap structure. This method is shown in Fig. 3. The firstly, primary reaction of the invader mRNA assay: Annealing of the invasive and probe oligonucleotides to an RNA target forms an overlap-flap structure that is a substrate for the structure-specific 5′ nuclease, Cleavase (Fig. 3a). Secondary reaction to generate quantifiable signals: a secondary overlap-flap structure is formed by hybridizing both the 5′ flap that had been released in the primary reaction and a FRET (fluorescence resonance energy transfer) oligonucleotide to a secondary reaction template (SRT). The FRET oligonucleotide is labeled with a fluorophore (F) and a quencher (Q) so cleavage between them generates the fluorescences signal (Fig. 3b). Invader miRNA assay primary reaction: The overall structure of the substrate resembles that shown in A, except that the short size of the miRNA target requires the inclusion of extra structures derived from the invasive and probe oligonucleotides, forming a dumbbell-like structure. The noncomplementary 5′-flap is hatched and the arrow indicates the site of cleavage, which releases the 5′ flap (Fig. 3c). The Invader miRNA assay has the ability to detect and quantitate as few as 20,000 molecules of an individual miRNA. Its specificity allowed for discrimination between miRNAs differing by a single nucleotide, and between precursor and mature miRNAs. This assay was used successfully in the analysis of several miRNAs, using as little as 50 ∼ 100 ng of total cellular RNA or as few as 1,000 lysed cells [69].

Real-time RT-PCR

Currently, there are three methods for quantitative PCR of miRNAs. One of the advantages of three type of detection is that it can verify miRNAs that are expressed at low levels, but it is limited by high cost. The first method is that primer-extension (PE), real-time PCR method [70]. The PE-qPCR assay involved two steps (Fig. 4a). In the first step, a tailed, gene-specific primer (GS) was used (1) to convert the RNA template into cDNA; (2) to introduce a

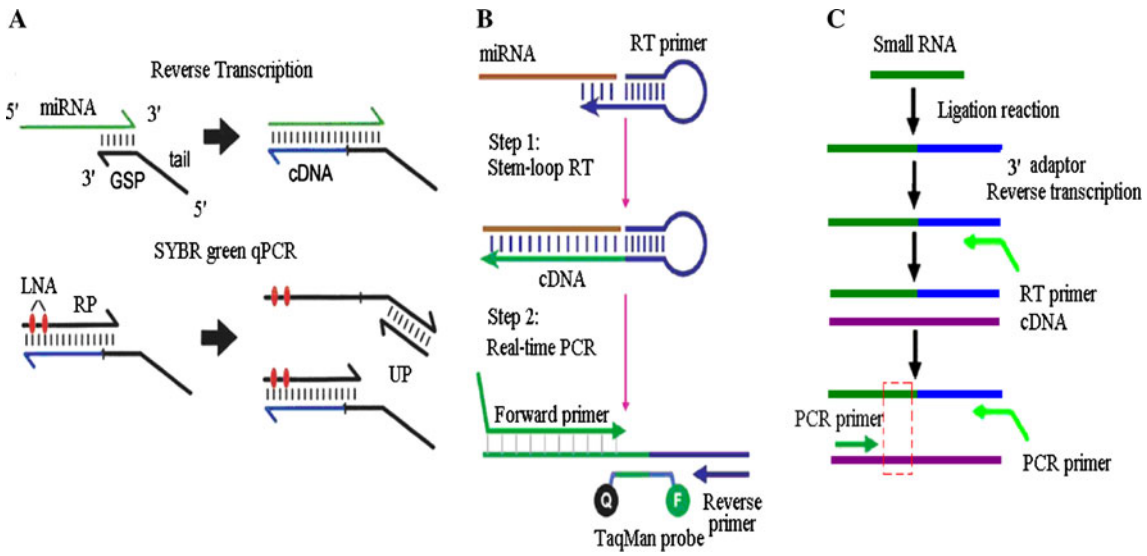

Fig. 4 Schemes of two kinds of RT-PCR methods for assay of miRNAs. **a** Primer-extension **b** Cloning miRNAs with stem–loop RT primers. c Cloning miRNAs with 3′ adaptor

''universal'' PCR binding site to one end of the cDNA molecule; and (3) to extend the length of the cDNA to facilitate subsequent monitoring by qPCR. In the second step, the resulting primer-extended, full-length cDNA was quantified by real-time PCR using a combination of an LNA-containing, miRNA/siRNA-specific ''reverse'' primer (LNA-R primer) and a generic universal primer common to all assays. Amplification of this chimeric cDNA was monitored in the real-time PCR reaction by SYBR green fluorescence. The assay has a high dynamic range and provides linear readout over differences in miRNA concentrations that span 6–7 orders of magnitude. It is capable of discriminating between related miRNA family members that differ by subtle sequence differences. The author used the method for quantitative analysis of six miRNAs across 12 tissue samples. The data confirm striking variation in the patterns of expression of these noncoding regulatory RNAs [70].

The second method is to use real-time PCR by specific stem–loop primers for quantifying both the precursor and mature miRNAs [71]. It included two steps: RT and real time PCR. First, the stem–loop RT primer is hybridized to a miRNA molecule and then reverse transcribed with a multiscribe reverse transcriptase (Fig. 4b). Next, the RT products are quantified using conventional TaqMan PCR. The assays method was specific for mature miRNAs and discriminated among related miRNAs that differ by as little as one nucleotide. Furthermore, they are not affected by genomic DNA contamination. The assay is achieved routinely with as little as 25 pg of total RNA for precise quantification of most miRNAs [71].

The third methods is that total RNA was polyadenylated by poly (A) polymerase and then cDNA was synthesized by an RT primer and reverse transcriptase using the 3′-tailed total RNA as templates [72]. The cDNAs of miRNAs for real-time PCR, with miRNA-specific forward primer and reverse primer complementary to 3′ adapter (Fig. 4c). It was demonstrated that can readily discriminate the expression of miRNAs having as few as one nucleotide sequence difference using as little as 100 pg total RNA for miRNAs sequence [72].

Rolling circle amplification detection

Jonstrup et al. [73] presented a simple miRNA detection protocol based on padlock probes and rolling circle amplification (RCA). Padlock probes are linear DNA probes where the terminal sequences are designed to hybridize to two adjacent target sequences. Under right conditions, DNA ligase will ligate the termini of the padlock probe on a perfectly matching RNA template, accurately distinguishing matched and mismatched substrates. The miRNA, used as a template, can subsequently be used as primer for rolling circle amplification, thereby linearly amplifying the target sequence in a quantitative manner. It can be performed without specialized equipment and is capable of measuring the content of specific miRNAs in a few nanograms of total RNA [73].

Radiolabeled splinted ligation detection

In 2007, Maroney et al. [74, 75] developed the splinted-ligation technique for detection of miRNAs. It is a nucleic acid hybridization assay that uses a bridge oligonucleotide with perfect base pairing complementarity to a target small RNA and a 5′-end-radiolabeled ligation oligonucleotide (Fig. 5). Simultaneous base pairing between both the small RNA and ligation oligonucleotide to the bridge oligonucleotide yields a double-stranded structure with a nick on one strand, which can be ligated with T4DNA ligase, thus labeling the target small RNA. In addition, because the labeled phosphate provided by the ligation oligonucleotide is rendered insensitive to phosphatase activity, the label present on the unligated oligonucleotide can be removed by incubation with phosphatase after the ligation step. Following the splinted-ligation reaction, labeled small RNAs carrying a nucleotide extension equivalent to the length of the ligation oligonucleotide and any residual-labeled ligation oligonucleotides can then be separated by denaturing gel electrophoresis and visualized by autoradiography or phosphorimaging. This method has been successfully used to study expression of various classes of biological small RNAs from nanogram to microgram amounts of total RNA without an amplification step. It is significantly simpler to perform and more sensitive than either northern blotting or ribonuclease protection assays [74].

Concluding remarks

Since the first lin-4 miRNA of C. elegans was discovered in 1993, the increased interest in miRNAs, and concerns that miRNAs are known to play important roles in living organisms, have attracted many molecular researchers to study miRNAs [10]. Bioinformatics approaches and detection methods of miRNAs contributed greatly to understanding these tiny molecules. Technologies such as NGS will boost the discovery of many expressed small RNAs and will undoubtedly result in the identification of more candidate miRNAs. We believe that the growth of the quantity and quality of experimentally determined miRNA genes will be the driving force for next generation of computational miRNA tools. Although the detection methods of miRNA mentioned above have their own advantages and disadvantages, with the availability of

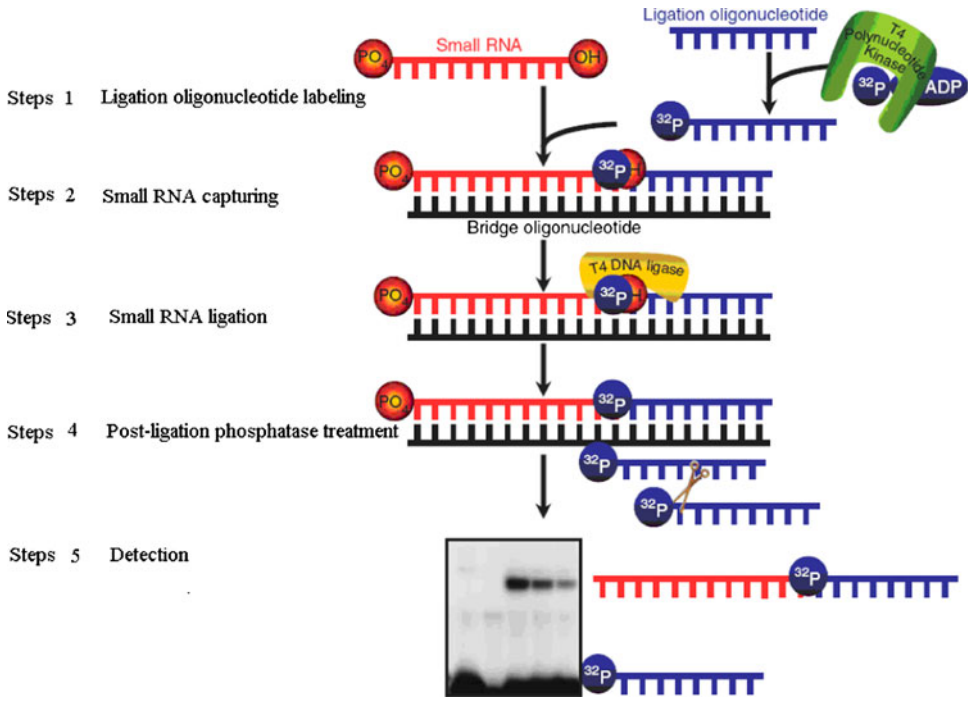

Fig. 5 Flowchart depicting each step of the small RNA detection using splinted-ligation method

reliable, quantitative, and sensitive methods of miRNA measurement, the analysis of differential expression of miRNAs will be able to improve and provide. An integrated detection approach, which combines computational prediction together with high-throughput biological detection, has been most effective in discovery of miRNAs [76]. Interested readers can find out more about them from the corresponding literature.

Acknowledgments This research was partly supported by research grants from the National Basic Research 973 Program (2005CB121004); the National High Tech Development Project of China, the 863 Program (2006AA10A119); Innovation Foundation for Graduate Students of Jiangsu Province and the Natural Science Foundation of China (61001013).

References

1. Huang JC, Babak T, Corson TW, Chua G, Khan S, Gallie BL, Hughes TR, Blencowe BJ, Frey BJ, Morris QD (2007) Using expression profiling data to identify human microRNA targets. Nat Methods 4:1045–1049
2. Carrington JC, Ambros V (2003) Role of microRNAs in plant and animal development. Science 301:336–338
3. Pasquinelli AE, Hunter S, Bracht J (2005) MicroRNAs: a developing story. Curr Opin Genet Dev 15:200–205
4. Williams AE (2008) Functional aspects of animal microRNAs. Cell Mol Life Sci 65:545–562
5. Huang Y, Shen XJ, Zou Q, Wang SP, Tang SM, Zhang GZ (2010) Biological functions of microRNAs: a review. J Physiol Biochem. doi:10.1007/s13105-010-0050-6
6. Kozomara A, Griffiths-Jones S (2010) miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. doi:10.1093/nar/gkq1027
7. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS (2003) MicroRNA targets in Drosophila. Genome Biol 5:R1
8. Liu B, Li J, Tsykin A (2009) Discovery of functional miRNA-mRNA regulatory modules with computational methods. J Biomed Inform 42:685–691
9. Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75:843–854
10. Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. Nature 403:901–906
11. Lin SY, Johnson SM, Abraham M, Vella MC, Pasquinelli A, Gamberi C, Gottlieb E, Slack FJ (2003) The *C elegans* hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. Dev Cell 4:639–650
12. Abrahante JE, Daul AL, Li M, Volk ML, Tennessen JM, Miller EA, Rougvie AE (2003) The *Caenorhabditis elegans* hunchback-like gene lin-57/hbl-1 controls developmental time and is regulated by microRNAs. Dev Cell 4:625–637
13. Vella MC, Choi EY, Lin SY, Reinert K, Slack FJ (2004) The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3′UTR. Genes Dev 18:132–137
14. Slack FJ, Basson M, Liu Z, Ambros V, Horvitz HR, 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
15. Peters JL, Cnudde F, Gerats T (2003) Forward genetics and map-based cloning approaches. Trends Plant Sci 8:484–491
16. Ambros V, Lee RC (2004) Identification of microRNAs and other tiny noncoding RNAs by cDNA cloning. Methods Mol Biol 265:131–138
17. Xu G, Zhang Y, Jia H, Li J, Liu X, Engelhardt JF, Wang Y (2009) Cloning and identification of microRNAs in bovine alveolar macrophages. Mol Cell Biochem 332:9–16
18. Long JE, Chen HX (2009) Identification and characteristics of cattle microRNAs by homology searching and small RNA cloning. Biochem Genet 47:329–343

19. He X, Zhang Q, Liu Y, Pan X (2007) Cloning and identification of novel microRNAs from rat hippocampus. Acta Biochim Biophys Sin (Shanghai) 39:708–714
20. Pfeffer S, Zavolan M, Grasser FA, Chien M, Russo JJ, Ju J, John B, Enright AJ, Marks D, Sander C, Tuschl T (2004) Identification of virus-encoded microRNAs. Science 304:734–736
21. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T (2002) Identification of tissue-specific microRNAs from mouse. Curr Biol 12:735–739
22. Bentwich I, Avniel A, Karov Y, Aharonov R, Gilad S, Barad O, Barzilai A, Einat P, Einav U, Meiri E, Sharon E, Spector Y, Bentwich Z (2005) Identification of hundreds of conserved and nonconserved human microRNAs. Nat Genet 37:766–770
23. Yu X, Zhou Q, Li SC, Luo Q, Cai Y, Lin WC, Chen H, Yang Y, Hu S, Yu J (2008) The silkworm (Bombyx mori) microRNAs and their expressions in multiple developmental stages. PLoS One 3:e2997
24. Pfeffer S, Lagos-Quintana M, Tuschl T (2005) Cloning of small RNA molecules. Curr Protoc Mol Biol, Chapter 26: Unit 26.4
25. Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, Rajewsky N (2008) Discovering microRNAs from deep sequencing data using miRDeep. Nat Biotechnol 26: 407–415
26. Git A, Dvinge H, Salmon-Divon M, Osborne M, Kutter C, Hadfield J, Bertone P, Caldas C (2010) Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. RNA 16:991–1006
27. Morozova O, Marra MA (2008) Applications of next-generation sequencing technologies in functional genomics. Genomics 92:255–264
28. Olson AJ, Brennecke J, Aravin AA, Hannon GJ, Sachidanandam R (2008) Analysis of large-scale sequencing of small RNAs. Pac Symp Biocomput 13:126–136
29. Metzker ML (2010) Sequencing technologies—the next generation. Nat Rev Genet 11:31–46
30. Cahill MJ, Koser CU, Ross NE, Archer JA (2010) Read length and repeat resolution: exploring prokaryote genomes using next-generation sequencing technologies. PLoS One 5:e11518
31. Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M, Eaves CJ, Marra MA (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Res 18:610–621
32. Kawaji H, Hayashizaki Y (2008) Exploration of small RNAs. PLoS Genet 4:e22
33. Stark A, Kheradpour P, Parts L, Brennecke J, Hodges E, Hannon GJ, Kellis M (2007) Systematic discovery and characterization of fly microRNAs using 12 Drosophila genomes. Genome Res 17: 1865–1879
34. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456:53–59
35. Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, Zeng K, Malek JA, Costa G, McKernan K, Sidow A, Fire A, Johnson SM (2008) A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. Genome Res 18:1051–1063
36. Petrosino JF, Highlander S, Luna RA, Gibbs RA, Versalovic J (2009) Metagenomic pyrosequencing and microbial identification. Clin Chem 55:856–866
37. Doran J, Strauss WM (2007) Bio-informatic trends for the determination of miRNA-target interactions in mammals. DNA Cell Biol 26:353–360
38. Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP (2003) Vertebrate microRNA genes. Science 299:1540
39. Schuster P, Fontana W, Stadler PF, Hofacker IL (1994) From sequences to shapes and back: a case study in RNA secondary structures. Proc Biol Sci 255:279–284
40. Ohler U, Yekta S, Lim LP, Bartel DP, Burge CB (2004) Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification. RNA 10:1309–1322
41. Mendes ND, Freitas AT, Sagot MF (2009) Current tools for the identification of miRNA genes and their targets. Nucleic Acids Res 37:2419–2433
42. Lai EC, Tomancak P, Williams RW, Rubin GM (2003) Computational identification of Drosophila microRNA genes. Genome Biol 4:R42
43. Ruby JG, Stark A, Johnston WK, Kellis M, Bartel DP, Lai EC (2007) Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. Genome Res 17:1850–1864
44. Berezikov E, Guryev V, van de Belt J, Wienholds E, Plasterk RH, Cuppen E (2005) Phylogenetic shadowing and computational identification of human microRNA genes. Cell 120:21–24
45. Xu Y, Zhou X, Zhang W (2008) MicroRNA prediction with a novel ranking algorithm based on random walks. Bioinformatics 24:i50–i58
46. Hackenberg M, Sturm M, Langenberger D, Falcon-Perez JM, Aransay AM (2009) miRanalyzer: a microRNA detection and analysis tool for next-generation sequencing experiments. Nucleic Acids Res 37:W68–W76
47. Nam JW, Shin KR, Han J, Lee Y, Kim VN, Zhang BT (2005) Human microRNA prediction through a probabilistic co-learning model of sequence and structure. Nucleic Acids Res 33:3570–3581
48. Sewer A, Paul N, Landgraf P, Aravin A, Pfeffer S, Brownstein MJ, Tuschl T, van Nimwegen E, Zavolan M (2005) Identification of clustered microRNAs using an ab initio prediction method. BMC Bioinformatics 6:267
49. Xue C, Li F, He T, Liu GP, Li Y, Zhang X (2005) Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine. BMC Bioinformatics 6:310
50. Grundhoff A, Sullivan CS, Ganem D (2006) A combined computational and microarray-based approach identifies novel microRNAs encoded by human gamma-herpesviruses. RNA 12: 733–750
51. Hertel J, Stadler PF (2006) Hairpins in a Haystack: recognizing microRNA precursors in comparative genomics data. Bioinformatics 22:e197–e202
52. Sheng Y, Engstrom PG, Lenhard B (2007) Mammalian microRNA prediction through a support vector machine model of sequence and structure. PLoS One 2:e946
53. Yousef M, Nebozhyn M, Shatkay H, Kanterakis S, Showe LC, Showe MK (2006) Combining multi-species genomic data for microRNA identification using a Naive Bayes classifier. Bioinformatics 22:1325–1334
54. Yousef M, Jung S, Showe LC, Showe MK (2008) Learning from positive examples when the negative class is undetermined–microRNA gene identification. Algorithms Mol Biol 3:2
55. Jiang P, Wu H, Wang W, Ma W, Sun X, Lu Z (2007) MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. Nucleic Acids Res 35:W339–W344
56. Helvik SA, Snove O Jr, Saetrom P (2007) Reliable prediction of Drosha processing sites improves microRNA gene prediction. Bioinformatics 23:142–149
57. Adai A, Johnson C, Mlotshwa S, Archer-Evans S, Manocha V, Vance V, Sundaresan V (2005) Computational prediction of miRNAs in *Arabidopsis thaliana*. Genome Res 15:78–91

58. Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science 294:858–862
59. Pall GS, Codony-Servat C, Byrne J, Ritchie L, Hamilton A (2007) Carbodiimide-mediated cross-linking of RNA to nylon membranes improves the detection of siRNA, miRNA and piRNA by northern blot. Nucleic Acids Res 35:e60
60. Varallyay E, Burgyan J, Havelda Z (2007) Detection of microRNAs by Northern blot analyses using LNA probes. Methods 43:140–145
61. Liu CG, Calin GA, Meloon B, Gamliel N, Sevignani C, Ferracin M, Dumitru CD, Shimizu M, Zupo S, Dono M, Alder H, Bullrich F, Negrini M, Croce CM (2004) An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. Proc Natl Acad Sci USA 101:9740–9774
62. Wang X (2006) Systematic identification of microRNA functions by combining target prediction and expression profiling. Nucleic Acids Res 34:1646–1652
63. Wienholds E, Kloosterman WP, Miska E, Alvarez-Saavedra E, Berezikov E, de Bruijn E, Horvitz HR, Kauppinen S, Plasterk RH (2005) MicroRNA expression in zebrafish embryonic development. Science 309:310–311
64. Kloosterman WP, Wienholds E, de Bruijn E, Kauppinen S, Plasterk RH (2006) In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. Nat Methods 3:27–29
65. Kloosterman WP, Steiner FA, Berezikov E, de Bruijn E, van de Belt J, Verheul M, Cuppen E, Plasterk RH (2006) Cloning and expression of new microRNAs from zebrafish. Nucleic Acids Res 34:2558–2569
66. Nelson PT, Baldwin DA, Kloosterman WP, Kauppinen S, Plasterk RH, Mourelatos Z (2006) RAKE and LNA-ISH reveal microRNA expression and localization in archival human brain. RNA 12:187–191
67. Deo M, Yu JY, Chung KH, Tippens M, Turner DL (2006) Detection of mammalian microRNA expression by in situ hybridization with RNA oligonucleotides. Dev Dyn 235: 2538–2548
68. Berezikov E, Cuppen E, Plasterk RH (2006) Approaches to microRNA discovery. Nat Genet 38:S2–S7
69. Allawi HT, Dahlberg JE, Olson S, Lund E, Olson M, Ma WP, Takova T, Neri BP, Lyamichev VI (2004) Quantitation of microRNAs using a modified Invader assay. RNA 10:1153–1161
70. Raymond CK, Roberts BS, Garrett-Engele P, Lim LP, Johnson JM (2005) Simple, quantitative primer-extension PCR assay for direct monitoring of microRNAs and short-interfering RNAs. RNA 11:1737–1744
71. Schmittgen TD, Lee EJ, Jiang J, Sarkar A, Yang L, Elton TS, Chen C (2008) Real-time PCR quantification of precursor and mature microRNA. Methods 44:31–38
72. Shi R, Chiang VL (2005) Facile means for quantifying microRNA expression by real-time PCR. Biotechniques 39:519–525
73. Jonstrup SP, Koch J, Kjems J (2006) A microRNA detection system based on padlock probes and rolling circle amplification. RNA 12:1747–1752
74. Maroney PA, Chamnongpol S, Souret F, Nilsen TW (2008) Direct detection of small RNAs using splinted ligation. Nat Protoc 3:279–287
75. Maroney PA, Chamnongpol S, Souret F, Nilsen TW (2007) A rapid, quantitative assay for direct detection of microRNAs and other small RNAs using splinted ligation. RNA 13:930–936
76. Li W, Ruan K (2009) MicroRNA detection by microarray. Anal Bioanal Chem 394:1117–1124