Review TRENDS in Genetics Vol.22 No.3 March 2006

Full text provided by www.sciencedirect.com

Genomics of microRNA

V. Narry Kim[1] and Jin-Wu Nam[2]

[1]Department of Biological Sciences and Research Center for Functional Cellulomics, Seoul National University, Seoul, 151-742, Korea
[2]Graduate Program in Bioinformatics, Seoul National University, Seoul, 151-742, Korea

Discovered just over a decade ago, microRNA (miRNA) is now recognized as one of the major regulatory gene families in eukaryotic cells. Hundreds of miRNAs have been found in animals, plants and viruses, and there are certainly more to come. Through specific base-pairing with mRNAs, these tiny ~22-nt RNAs induce mRNA degradation or translational repression, or both. Because a miRNA can target numerous mRNAs, often in combination with other miRNAs, miRNAs operate highly complex regulatory networks. In this article, we summarize the current status of miRNA gene mining and miRNA expression profiling. We also review up-to-date knowledge of miRNA gene structure and the biogenesis mechanism. Our focus is on animal miRNAs.

Introduction

Since the discovery of RNA interference (RNAi), efforts to identify endogenous small RNAs have led to the discovery of hundreds of miRNAs in nematodes, fruit flies and humans [1–21]. More than 500 different miRNAs have been identified in animals and plants, where the number of miRNA genes is expected to increase to 500–1000 per species, which would comprise ~2–3% of protein-coding genes [22]. At the time of writing, the miRNA database (http://www.sanger.ac.uk/Software/Rfam/mirna/) contained 114 *Caenorhabditis elegans* miRNAs, 78 *Drosophila melanogaster* miRNAs, 369 *Danio rerio* miRNAs, 122 *Gallus gallus* (chicken) miRNAs, 326 *Homo sapiens* miRNAs and 117 *Arabidopsis thaliana* miRNAs [23]. Nearly all miRNAs are conserved in closely related species and many have homologs in distant species. At least a third of *C. elegans* miRNAs have homologs in humans [24], suggesting that their functions could also be conserved throughout the evolution of animal lineages.

Large DNA viruses have also been found to carry miRNA genes: five in Epstein–Barr virus, 12 in Kaposi sarcoma-associated herpesvirus, nine in mouse γ-herpesvirus 68, and nine in human cytomegalovirus.

In this review, we summarize the recent advances in miRNA gene identification, biogenesis and expression profiling, with a focus on mammalian miRNAs.

Identification of microRNA genes

Considerable effort has been devoted to the identification of miRNA genes since the discoveries of numerous miRNAs in 2001 following the earlier genetic identification of *lin-4* and *let-7* [1,25,26]. Three types of approaches have been used to identify miRNA genes. The first approach is through forward genetics, as is the case for the first two miRNAs. In *Drosophila*, the gene responsible for the bantam phenotype displaying smaller body size turned out to be a miRNA gene [27]. Also in *C. elegans*, *lsy-6* mutants with defects in neuronal left–right asymmetry had mutations in an miRNA gene [7]. Because genetic studies are based on clear phenotypes, the *in vivo* functions of genetically identified RNAs are well established.

A breakthrough in large-scale miRNA identification was made when directional cloning was used to construct a cDNA library for endogenous small RNAs [28,29]. Briefly, total RNA is separated in denaturing polyacrylamide gel from which small RNA of 19–25 nt is gel-purified. The size-fractionated RNA is then ligated to 5′ and 3′ adaptor molecules, and amplified by RT–PCR to construct the cDNA library for sequencing individual clones. Usually, more than half of the cloned sequences are from fragments of larger RNAs such as tRNA or rRNA. According to current convention, miRNA is defined as a single-stranded RNA, ~22 nt in length, that is generated by the RNase III-type enzymes from a local hairpin structure embedded in an endogenous transcript [30]. miRNAs differ from small interfering RNAs (siRNAs) in that siRNAs are processed from long double-stranded RNAs (dsRNAs) [31]. So, when a small RNA is discovered by cDNA cloning, it must meet the following criteria to be classified as a miRNA (for more detailed information see Ref. [32]). First, the small RNA sequence should be present in one arm of the hairpin structure, which lacks large internal loops or bulges. The hairpins are usually ~80 nt in animals, but the lengths are more variable in plants. The small RNA sequences should be phylogenetically conserved in closely related species. The sequence conservation is usually lower in the loop region than in the mature miRNA segment. Second, its expression should be confirmed by hybridization to a size-fractionated RNA sample, typically by northern blotting. The blot normally shows both the mature form (a ~22-nt band) and the hairpin precursor (a ~70-nt band). When the expression is too low for detection, if the putative precursor of the cloned small RNA has a conserved hairpin structure it can be annotated as a miRNA. Close homologs in other species can be annotated as miRNA orthologs without experimental validation, provided that they fulfill the conserved precursor criterion. Small RNAs that do not meet these requirements can be classified as either siRNAs or other provisional classes [31]. Hundreds of miRNAs have been

Corresponding author: Kim, V.N. (narrykim@snu.ac.kr).
Available online 30 January 2006

www.sciencedirect.com 0168-9525/$ - see front matter © 2006 Elsevier Ltd. All rights reserved. doi:10.1016/j.tig.2006.01.003

cloned from various cell lines, diverse tissues of mouse, fly and zebrafish, and a range of developmental stages in mouse, fly, worm, frog and zebrafish [2,5,6,10,14,20,27,29]. 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.

This problem can be overcome, at least in part, by bioinformatic approaches and examining genomic sequences. Computational identification is based largely on the phylogenetic conservation and the structural characteristics of miRNA precursors. Simple homology searches can reveal homologs of known miRNAs. Also useful is to look for stem-loops in the vicinity of known miRNA genes because more than half of known miRNA genes are present as tandem arrays within operon-like clusters [5,26,33]. More extensive gene mining at a genomic level typically begins with identifying conserved genomic segments in the intergenic area that potentially fold into a stem-loop structure. The first miRNA search algorithm was MiRscan, which successfully predicted miRNA genes that display close homology in two nematode worms, *C. elegans* and *C. briggsae* [24]. MiRscan was further improved by defining conserved sequence motifs found in the vicinity of nematode miRNA genes [34]. MiRscan is available on the web (http://genes.mit.edu/mirscan/). Another sensitive computational scoring tool is miRseeker, which has been applied to screen conserved stem-loops in insect genomes [8].

Further improvements have since been introduced, enabling impressive expansion of the list of miRNA genes in diverse species (Table 1). One of the particularly useful features used in recent studies was the strong conservation of the nucleotides in the 5′ region of mature miRNA. Seven nucleotides at 2–7 positions (relative to the 5′ end of miRNA), known as 'seed' sequences, are crucial in binding to the target mRNA. Through systematic comparative genomics analysis, Xie *et al.* revealed 106 conserved motifs in the 3′ untranslated region (UTR) of mRNA [19], many of which (72 motifs) turned out to be complementary to the 5′ ends of nearly half of the known miRNAs and were capable of forming 6–8 bp seed duplexes [19]. Encouraged by this result, Xie *et al.* used the entire conserved motifs to predict miRNA genes and obtained 129 novel miRNA candidates. When 12 candidates were tested for expression with an RNA sample pooled from just ten different tissues, six candidates were validated for expression. This indicated that many of the 129 predicted candidates are likely to be authentic miRNA genes.

Another recent prediction method took advantage of conservation of known miRNA genes [18]. The stems of miRNA hairpins are highly conserved whereas the loop region is poorly conserved. The degree of conservation rapidly declines for sequences immediately flanking the miRNA hairpins. Using this characteristic profile, 976 candidate miRNAs were predicted and 16 out of 69 representative candidates (23%) were confirmed for their expression by northern blotting.

More recently, Nam *et al.* introduced a probabilistic co-learning model (ProMiR), a variant of the pair hidden Markov model, to screen statistically both strongly and weakly conserved stem-loops in the human genome [35]. ProMiR successfully detects miRNA genes with no clear sequence similarity to known miRNA genes. Nine out of 23 representative candidate genes (39%) were validated by a RT-PCR-based method for expression in HeLa cells.

An intensive integrative approach was taken by Bentwich *et al.*, who combined bioinformatics predictions with microarray analysis and sequence-directed cloning [21]. Hairpins from the entire human genome were first screened. After excluding the hairpins that are located in repetitive elements and protein-coding regions, high-throughput validation was performed for ~5300 selected hairpins by microarray analysis on five different human

Table 1. miRNA gene prediction algorithms

Prediction target	Methods	Detection of non-conserved miRNA	Name of program	Species	Refs
Pre-miRNA	Comparative analysis, stem-loop conservation	−	miRscan	Nematode	[24]
Pre-miRNA	Sequential and structural properties	−	srnaloop	Nematode	[9]
Pre-miRNA	Comparative analysis, stem-loop conservation	−	miRseeker	Fly	[8]
Pre-miRNA	Sequential and structural properties, comparative analysis	−	–	*Arabidopsis*	[105]
Pre-miRNA	Sequence or structural alignment	−	ERPIN	Animal, plant	[106]
Pre- miRNA and mature miRNA	Seed match, comparative analysis	−	findMiRNA	*Arabidopsis*	[107]
Pre-miRNA	Phylogenetic shadowing profile	+	–	Human	[18]
Mature miRNA	Seed match	+	–	Human	[19]
Pre-miRNA	Sequence or structural alignment	−	miralign	Animal, plant	[108]
Pre- miRNA and mature miRNA	Probabilistic model of pairwise sequences	+	ProMiR	Human	[35]
Pre-miRNA	Sequential and structural properties	+	PalGrade	Human	[21]

tissues. Microarray yielded 886 candidates that gave strong signals. Out of these, 359 were further subjected to final confirmation by a sequence-directed cloning procedure to yield 89 novel human miRNAs. Interestingly, most nonconserved miRNAs found in this study are located in two large clusters and they seem to be poorly conserved outside primates. Taken together, these new results suggest the presence of a greater number of miRNA genes than previously estimated. New estimates are ~400–500 conserved miRNAs in humans [19,21]. Including less conserved miRNAs, the number could be at least 800 (2–3% of the total number of genes in humans).

Bioinformatically predicted miRNAs should be validated for their expression. Various methods have been developed for this purpose. Northern blotting is most widely used although it might not be sensitive enough to detect less-abundant miRNAs and it does not reveal the actual miRNA sequences (the exact boundaries of a miRNA). Primer extension analysis can be used to map the 5′ end so as to complement northern analysis. PCR-based amplification of adaptor-ligated cDNA has also been developed for validation of predicted miRNAs. cDNA clones can be amplified using the primer with partial coverage of the predicted miRNA sequences and the primer that matches to the adaptor [24]. The PCR product is then cloned and sequenced. This method is sensitive although it can be difficult in practice when the actual mature miRNA region is not known. A more informative but complicated method uses a biotin-labeled oligonucleotide to capture miRNA from a cDNA library, which is then amplified by PCR using adaptor primers [21].

Genomic distribution and gene structure

miRNA genes are scattered in all chromosomes in humans except for the Y chromosome. Approximately 50% of known miRNAs are found in clusters [1,3,26] and they are transcribed as polycistronic primary transcripts [36]. The miRNAs in a given cluster are often related to each other, suggesting that the gene cluster is a result of gene duplication. A miRNA gene cluster also often contains unrelated miRNAs. A plausible but yet-to-be validated possibility is that the clustered miRNAs are functionally related by virtue of targeting the same gene or different genes in the same pathway.

It was initially thought that most miRNA genes were located in intergenic regions [1,26]. However, recent analyses of miRNA gene locations showed that the majority (~70%) of mammalian miRNA genes (161 out of 232) are located in defined transcription units (TUs) [37]. By combining up-to-date genome assemblies and expressed sequence tag (EST) databases, Rodriguez *et al.* demonstrated that many miRNA genes (117 out of 161) were found in the introns in the sense orientation, which is more than previously expected (Figure 1) [37]. Of these 117 intronic miRNAs, 90 miRNAs are in the introns of protein-coding genes, whereas 27 miRNAs are in the introns of noncoding RNAs (ncRNAs). In some cases (14 miRNAs), miRNAs are present in either an exon or an intron ('mixed') depending on the alternative splicing pattern. So, miRNA genes can be categorized based on their genomic locations: intronic miRNA in protein coding TU; intronic miRNA in noncoding TU; and exonic miRNA in noncoding TU (Figure 1). 'Mixed' miRNA genes can be assigned to one of the above groups depending on the given splicing pattern. Unexpectedly, a large number of

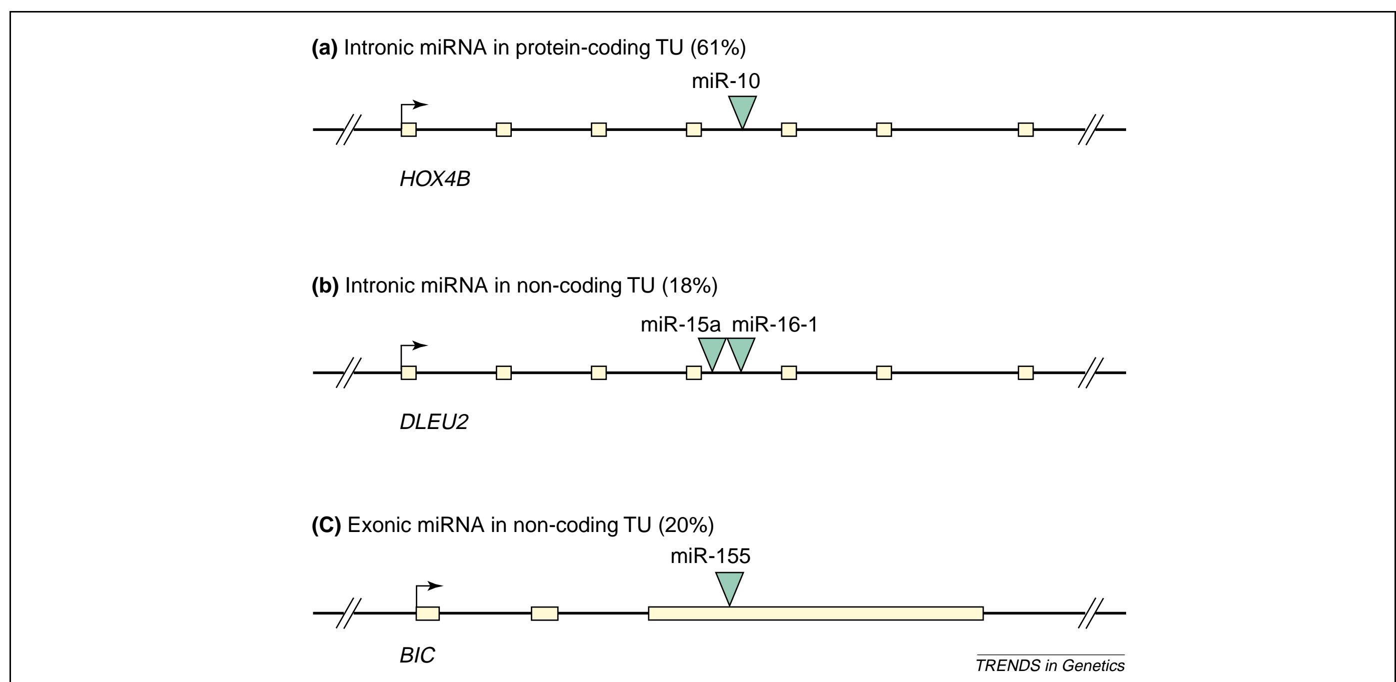

Figure 1. Genomic organization and structure of miRNA genes. **(a)** Intronic miRNA in a protein-coding transcriptional unit (TU). As an example, miR-10 in *HOX4B* gene is shown. The green triangle indicates the location of a miRNA stem-loop and the exons are shown in yellow. **(b)** Intronic miRNAs in a noncoding transcript. The miR-15a ~ 16-1 cluster is shown, which is found in the fourth intron of a previously defined noncoding RNA gene, *DLEU2*. **(c)** The structure of exonic miRNA in noncoding transcripts, such as miR-155. This figure is not drawn to scale.

miRNAs are found in introns of protein-coding genes (90 out of 161 miRNAs). This indicates that previous informatic searches confined to intergenic regions might have missed some miRNA genes. The location of some intronic miRNAs is well conserved among diverse species. For instance, miR-7 is found in the intron of hnRNP K in both insects and mammals. Another interesting example is the miR-106b∼25∼93 family that is found in the intron 13 of *MCM7* in both humans and mice.

As expected for genes sharing the same promoters, the 'host' transcript and miRNAs usually have similar expression profiles [1,26,37,38]. miRNA promoters have been identified experimentally in numerous studies [39–45]. Bioinformatic searches for miRNA-specific promoter elements upstream of miRNA sequences have not been successful. Instead, the characterized miRNA promoters contain general RNA polymerase II (Pol II) transcriptional regulatory elements previously found in protein-coding genes.

Biogenesis pathway

Transcription of the miRNA gene is mediated by RNA polymerase II (Pol II) [43,44] (Figure 2). The primary transcript contains the 7-methylguanosine cap and a poly(A) tail, which are unique for Pol II transcripts [43,44]. Further evidence for Pol II-dependent transcription was provided by demonstrating the α-amanitin sensitivity of miRNA transcription and the physical association of Pol II with the miRNA promoter [43]. Recent analysis of miRNA promoters has also revealed that a significant number of miRNA core promoters contain typical Pol II elements such as the TATA box [45,46]. However, one cannot formally exclude the possibility that a few miRNA genes might be transcribed by another type of RNA polymerase. Several miRNA genes in the mouse γ-herpesvirus 68 (MHV68) are apparently transcribed by Pol III, indicating that some viral miRNAs have evolved independently and such miRNAs could be generated through a distinct biogenetic pathway [47].

Pol II-dependent transcription enables temporal and positional control over miRNA expression so that a specific set of miRNAs can be expressed during development and under specific conditions and in certain cell types. For example, c-*myc* transactivates the transcription of the miR-17 cluster [48]. Muscle specific miR-1-1 and miR-1-2 are induced by serum response factors MyoD and Mef2 [49].

Transcription of miRNA genes yields long primary transcripts (pri-miRNAs) that contain a local foldback structure (Figure 2). The stem-loop structure is cleaved by the nuclear RNase III Drosha to release the precursor of miRNA (pre-miRNA) [50] (Figure 2). The remnants (the flanking fragments) are supposedly degraded in the nucleus, although the 3′ flanking fragment can often be found in EST databases [43]. Drosha is a large protein of ∼160 kDa, which is conserved in animals but not in plants [51–53]. Drosha requires a cofactor, the DiGeorge syndrome critical region gene 8 (DGCR8) protein in humans (also known as Pasha in *Drosophila* and *C. elegans*) [54–57]. Drosha and DGCR8 form a large ∼650 kDa complex in humans [55,56] and a ∼500 kDa complex in *Drosophila* [54], which is known as the microprocessor complex. Because DGCR8 (or Pasha) contains two double-stranded RNA-binding domains (dsRBDs), it is believed to assist Drosha in substrate recognition [54–57], although the precise biochemical role remains to be dissected.

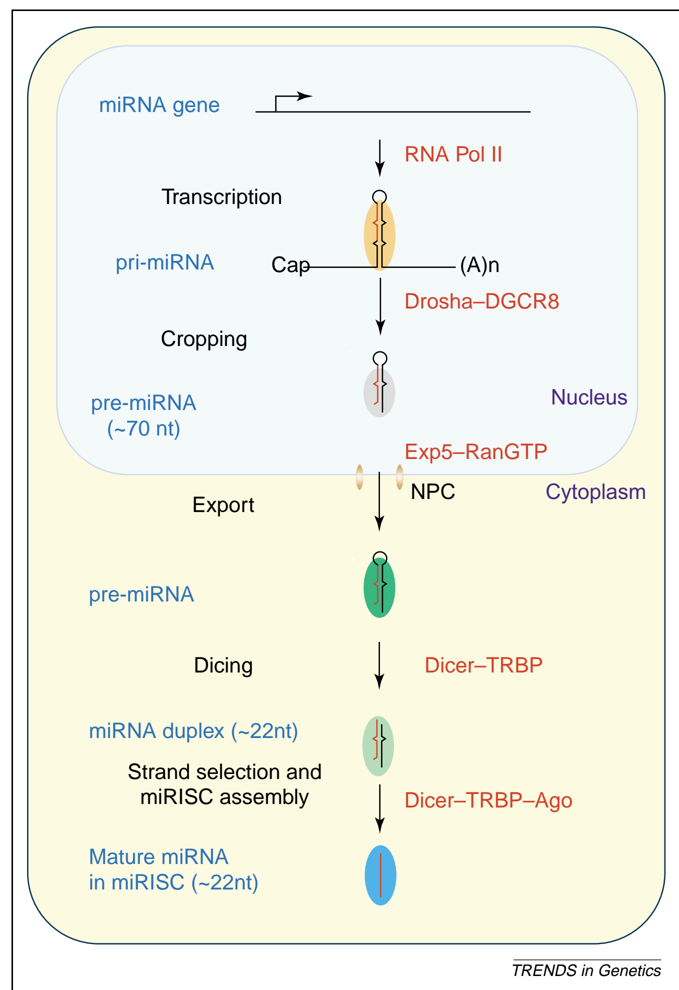

Figure 2. Model for miRNA biogenesis. miRNA genes are transcribed by an RNA polymerase II (Pol II) to generate the primary transcripts (pri-miRNAs). The initiation step (cropping) is mediated by the Drosha–DGCR8 complex (also known as the microprocessor complex). Drosha and DGCR8 are located mainly in the nucleus. The product of the nuclear processing is ∼70-nt pre-miRNA, which possesses a short stem plus ∼2-nt 3′ overhang. This structure can serve as a signature motif that is recognized by the nuclear export factor, Exportin-5 (Exp5). Pre-miRNA constitutes a transport complex together with Exp5 and its cofactor Ran (the GTP-bound form). Upon export, the cytoplasmic RNase III Dicer participates in the second processing step (dicing) to produce miRNA duplexes. The duplex is separated and usually one strand is selected as the mature miRNA, whereas the other strand is degraded. In *Drosophila*, R2D2 forms a heterodimeric complex with Dicer and binds to one end of the siRNA duplex, thereby selecting one strand of the duplex. It is not known if miRNAs use the same machinery for strand selection. It is also unclear whether an R2D2 homolog functions in animals other than *Drosophila*.

Because the next processing enzyme, Dicer, is confined to the cytoplasm, the Drosha product, pre-miRNA, needs to be exported to the cytoplasm. Export of pre-miRNA is mediated by one of the Ran-dependent nuclear transport receptors, exportin-5 (Exp5) [58–60]. When Exp5 is depleted by RNAi, mature miRNAs are reduced but pre-miRNA does not accumulate in the nucleus. This suggests that pre-miRNA could be relatively unstable and also that pre-miRNA might be stabilized through its interaction with Exp5 [59]. Exp5 recognizes the 'minihelix motif',

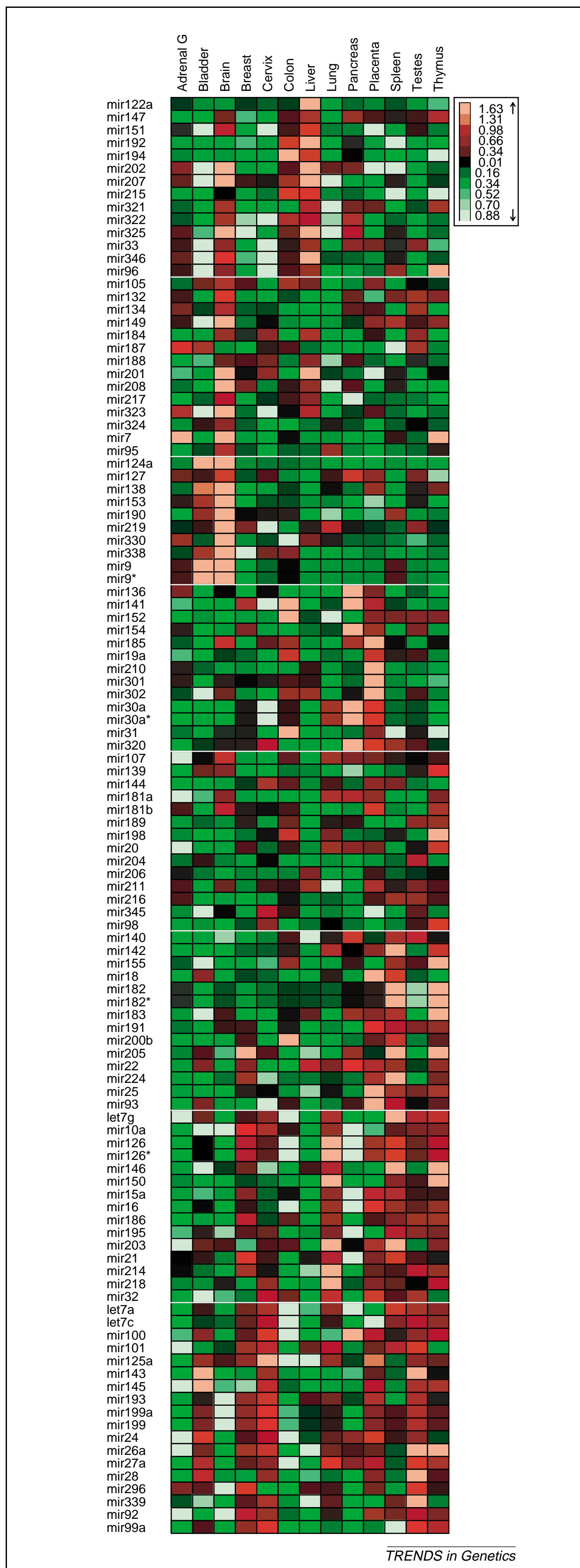

Figure 3. A summary of miRNA profiles described in recent studies. Data presented in three articles [38,85,94] were standardized with *z*-statistics to enable direct comparison of the data sets. To integrate three independent data sets, correlation analysis was performed for the values to be integrated. Compatible samples with correlation >0.25 were included and the values were averaged. Integrated data were then clustered using a self-organizing map (SOM) algorithm (3×3 grids, 50 000 iterations), yielding the homogeneity of 0.672. Light green represents no signal (–1), black is the lowest signal (0) and red is the greatest signal (+1); see the gradient indicated by the key in the figure.

which consists of a >14-bp stem and a short 3′ overhang [61,62].

On export, pre-miRNAs are processed into ~22-nt miRNA duplexes by the cytoplasmic RNase III Dicer [63–67]. Dicer is a highly conserved protein of ~200 kDa. Similar to Drosha, Dicer associates with a dsRBD-containing partner. For instance, Loquacious (also known as R3D1) interacts with Dicer-1 in *Drosophila* [68–70], and HIV-1 TAR RNA-binding protein (TRBP) binds to Dicer in humans [71]. Some of the Dicer cofactors do not seem to be required for the cleavage reaction itself because purified human Dicer and *Drosophila* Dicer-2 can catalyze the cleavage reaction [71–74]. The Dicer cofactors instead seem to have various roles in miRNA stability and effector complex formation.

Mature miRNAs are incorporated into the effector complexes, which are known as 'miRNP', 'mirgonaute' or, more generally, 'miRISC' (miRNA-containing RNA-induced silencing complex). The effector complex containing siRNA, in distinction, is referred to as 'RISC', 'sirgonaute' or 'siRISC'. During RISC assembly, the cleavage products (~22-nt miRNA duplexes) are rapidly converted into single strands. Usually one strand of this short-lived duplex disappears, whereas the other strand remains as a mature miRNA. The strand with relatively unstable base pairs at the 5′ end typically remains in siRISC [75,76]. R2D2, the fly cofactor for Dicer-2, binds to the more stable end of the siRNA duplex [77], thereby orienting the protein complex on the siRNA duplex. It remains unknown if similar machinery acts on the miRNA duplex and how conserved this machinery is.

Advances in miRNA expression profiling

miRNA expression can be regulated at multiple steps during RNA biogenesis, although it remains to be determined which step is controlled and how this control is achieved. Transcriptional regulation is likely to be the major control mechanism, although some miRNAs seem to be controlled at the post-transcriptional level. Expression profiling studies indicate that most miRNAs are under the control of developmental or tissue-specific signaling, or both [2,5,20,29,38,78–94]. Figure 3 presents a summary of miRNA profiles described in recent studies [38,85,94]. Three independent data sets determined from human tissues were normalized and the averaged values are presented as a heat map. miRNAs were sorted into eight clusters based on their similarity in tissue specificity.

A range of techniques are available for miRNA quantification, including northern blotting, dot blotting, RNase protection assay, primer extension analysis, Invader assay and quantitative PCR. Large-scale cDNA cloning can also provide information on the relative expression level of miRNAs in diverse samples. However, most of these techniques involve laborious procedures,

Table 2. miRNA expression profiling methods

Method	Label or dye	Detection target	St[a]	Sp[b]	Species	Samples	Refs
cDNA cloning		Mature miRNA	+	+++	*Xenopus*	Nine developmental stages	[20]
		Mature miRNA	+	+++	Zebrafish	Five developmental stages, two cell lines	[29]
Microarray	Biotin	Mature miRNA	+	+	Human and mouse	Six adult tissues	[87]
	Biotin	Pre-miRNA	++		Human and mouse	HeLa,18 human adult and two fetal tissues, mouse macrophage	[82]
	Cy3	Mature miRNA	+		Mouse	Five developmental stages of brain	[86]
	Cy3	Mature miRNA	++		Zebrafish	14 developmental stages (male and female)	[88]
	Cy3 and Cy5	Mature miRNA	++		Human	26 tissues	[94]
	Cy3 and Cy5	Mature miRNA	++		Mouse	Four developmental stages, seven tissues, ES cell	[93]
	Cy3 and Cy5	Pre-miRNA	++		Human	Five adult tissues, HeLa	[85]
	Cy5	Mature miRNA	++		Human	24 tissues	[38]
	Fluor	Mature miRNA	++		Mouse	14 tissues, three developmental stages	[84]
	Radioisotope	Mature miRNA	++		Mouse	Hematopoietic or non-hematopoietic cells	[91]
	RAKE assay	Mature miRNA	++	+++	Human	HeLa, Jurkat, malignant meningioma, oligodendroglioma	[92]
Quantitative RT-PCR		Pre-miRNA	+++	+++	Human	Six cancer cell lines	[83,89]
Bead-based flow cytometry		Mature miRNA	++	++	Human	218 tissues (normal and abnormal)	[90]

[a]The relative sensitivity of the detection methods is presented in three levels. The greatest sensitivity (+++) indicates that $<1\ \mu g$ of total RNA is sufficient for analysis. Medium sensitivity (++) indicates that $<10\ \mu g$ is required, whereas the least sensitive methods (+) necessitate $>10\ \mu g$ of total RNA sample.
[b]Relative specificity is also divided into three levels, according to the capability to distinguish related miRNAs.

making it difficult to determine the level of all known miRNAs. Recent development of easy quantification methods has enabled large-scale expression profiling of miRNAs (Table 2). Currently, the most widely used method is based on microarrays [38,78,82,84–88,91–94]. Although microarray is a powerful method for high-throughput analysis, the small size of miRNAs poses a challenge for conventional microarray techniques because it is difficult to create a single hybridization condition suitable for all miRNAs on the chip. Thus, some of the microarrays employ probes that are complementary to pre-miRNAs rather than mature miRNAs [82,85]. However, because maturation of miRNA is often regulated, the level of pre-miRNA does not always correlate with that of mature miRNA. Recently developed microarrays detect mature miRNA by employing antisense oligonucleotides that specifically bind to the mature miRNA sequence [38,84,86–88,91–94]. However, the problem of potential cross-hybridization of related miRNAs still remains unresolved. In addition, systematic bias could be introduced during reverse transcription, PCR amplification, enzymatic labeling or fluorescence tag ligation. These problems were successfully avoided by developing a new procedure called the RNA-primed array-based Klenow enzyme (RAKE) assay [92]. The DNA oligonucleotide probe on the slide contains the sequences antisense to miRNA and the universal spacer sequences. When miRNA binds to the probe, miRNA serves as a primer for extension upon the addition of the Klenow enzyme, and generates a double-stranded fragment with incorporated tagged nucleotides, which is easily detected. This method is particularly useful when closely related miRNAs need to be separately analyzed, because miRNAs with mismatches at the 3′ end cannot be extended. However, it should be mentioned that microarrays in most cases are not as quantitative as northern blotting so that it is difficult to determine precisely the relative abundance of miRNAs. Therefore, the problem of a narrow dynamic range remains to be overcome.

RT-PCR is unarguably the most sensitive method but could be difficult to use in high-throughput analysis if the number of miRNAs exceeds 300. The technical difficulty also stems from the short length of miRNA (~22 nt). Current methods are based on the detection of miRNA precursors rather than the mature miRNAs [83]. It is necessary to modify the methods to detect mature miRNA specifically without introducing experimental bias. A recent study showed that RT primers containing a partial stem-loop structure can enhance specificity for mature miRNAs [89].

The most recent innovation in miRNA detection involves the bead-based flow cytometric method [90]. Each individual bead is marked with fluorescence tags (which can yield up to 100 colors, each representing a single miRNA) and coupled to probes that are complementary to miRNAs of interest. miRNAs are ligated to the 5′ and 3′ adaptors, reverse-transcribed, amplified by PCR using a common biotinylated primer, hybridized to the capture beads, and stained with streptavidin–phycoerythrin. The beads are then analyzed using a flow cytometer

capable of measuring bead color (denoting miRNA identity) and phycoerythrin intensity (denoting miRNA abundance). Because hybridization takes place in solution, this method offers more specific detection of closely related miRNAs compared with conventional glass-slide microarrays. The complicated procedure, however, needs to be improved. Importantly, this study analyzed 217 known human miRNAs in 218 samples from normal and tumor tissues, demonstrating a surprisingly accurate correlation of the miRNAs with the development and differentiation of tumors. This offers the prospect of using miRNA expression profiles to help in the diagnosis of cancer.

Concluding remarks

To date 326 human miRNA genes have been discovered, and there are certainly more to come. A single miRNA can bind to and regulate many different mRNA targets and, conversely, several different miRNAs can bind to and cooperatively control a single mRNA target [95]. Two recent independent studies predicted that 20–30% of human genes could be controlled by miRNAs [19,96]. Thus, hundreds of RNAs and thousands of the targets appear to compose remarkably complex regulatory networks, thereby mediating many facets of eukaryotic cell function [97,98]. As Bartel and Chen proposed, the unique combination of miRNAs expressed in each cell type might affect or 'dampen' the use of thousands of target mRNAs [99]. Supporting this hypothesis, the experimental introduction of muscle- or brain-specific miRNA into HeLa cells shifted the overall mRNA profile to that of muscle or brain, respectively [100].

Tantalizing correlations between miRNA expression and human diseases have recently been demonstrated. Certain miRNAs are modulated in tumors and might be directly involved in the pathology (reviewed by Croce and Calin [101]). For instance, Let-7 targets the oncogenic *RAS* gene and *let-7* is often downregulated in lung cancer [102]. Deletion or downregulation of the miR-15∼16-1 cluster is highly correlated with chronic lymphocytic leukemia cases [101]. In fact, miR-16-1 was shown to suppress the antiapoptotic *Bcl-2* gene in humans [103]. Some other miRNAs, such as the miR-17∼92 cluster, have oncogenic potential [48,104], as demonstrated by a transgenic study where enforced expression of the miR-17∼92 cluster induced B-cell lymphoma when coexpressed with c-*myc*. Understanding the miRNA-guided network will provide a new window for diagnostics and therapy of many human diseases.

Acknowledgements

This work was supported by grants from the National Research Laboratory Programs (M1050000010905J000010910 and M10412000095-04J0000-03610) and the SRC program of KOSEF (R11–2005–009–01003–0).

References

1 Lagos-Quintana, M. *et al*. (2001) Identification of novel genes coding for small expressed RNAs. *Science* 294, 853–858
2 Lagos-Quintana, M. *et al*. (2002) Identification of tissue-specific microRNAs from mouse. *Curr. Biol.* 12, 735–739
3 Mourelatos, Z. *et al*. (2002) miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev.* 16, 720–728
4 Calin, G.A. *et al*. (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
5 Aravin, A.A. *et al*. (2003) The small RNA profile during *Drosophila melanogaster* development. *Dev. Cell* 5, 337–350
6 Houbaviy, H.B. *et al*. (2003) Embryonic stem cell-specific microRNAs. *Dev. Cell* 5, 351–358
7 Johnston, R.J. and Hobert, O. (2003) A microRNA controlling left–right neuronal asymmetry in *Caenorhabditis elegans*. *Nature* 426, 845–849
8 Lai, E.C. *et al*. (2003) Computational identification of *Drosophila* microRNA genes. *Genome Biol.* 4, R42
9 Grad, Y. *et al*. (2003) Computational and experimental identification of *C. elegans* microRNAs. *Mol. Cell* 11, 1253–1263
10 Dostie, J. *et al*. (2003) Numerous microRNPs in neuronal cells containing novel microRNAs. *RNA* 9, 180–186
11 Lim, L.P. *et al*. (2003) Vertebrate microRNA genes. *Science* 299, 1540
12 Lagos-Quintana, M. *et al*. (2003) New microRNAs from mouse and human. *RNA* 9, 175–179
13 Kim, J. *et al*. (2004) Identification of many microRNAs that copurify with polyribosomes in mammalian neurons. *Proc. Natl. Acad. Sci. U. S. A.* 101, 360–365
14 Kasashima, K. *et al*. (2004) Altered expression profiles of microRNAs during TPA-induced differentiation of HL-60 cells. *Biochem. Biophys. Res. Commun.* 322, 403–410
15 Suh, M.R. *et al*. (2004) Human embryonic stem cells express a unique set of microRNAs. *Dev. Biol.* 270, 488–498
16 Poy, M.N. *et al*. (2004) A pancreatic islet-specific microRNA regulates insulin secretion. *Nature* 432, 226–230
17 Weber, M.J. (2005) New human and mouse microRNA genes found by homology search. *FEBS J* 272, 59–73
18 Berezikov, E. *et al*. (2005) Phylogenetic shadowing and computational identification of human microRNA genes. *Cell* 120, 21–24
19 Xie, X. *et al*. (2005) Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. *Nature* 434, 338–345
20 Watanabe, T. *et al*. (2005) Stage-specific expression of microRNAs during *Xenopus* development. *FEBS Lett.* 579, 318–324
21 Bentwich, I. *et al*. (2005) Identification of hundreds of conserved and nonconserved human microRNAs. *Nat. Genet.* 37, 766–770
22 Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297
23 Griffiths-Jones, S. (2004) The microRNA registry. *Nucleic Acids Res.* 32, D109–D111
24 Lim, L.P. *et al*. (2003) The microRNAs of *Caenorhabditis elegans*. *Genes Dev.* 2, 2
25 Lee, R.C. and Ambros, V. (2001) An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294, 862–864
26 Lau, N.C. *et al*. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858–862
27 Brennecke, J. *et al*. (2003) *bantam* encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila*. *Cell* 113, 25–36
28 Ambros, V. and Lee, R.C. (2004) Identification of microRNAs and other tiny noncoding RNAs by cDNA cloning. *Methods Mol. Biol.* 265, 131–158
29 Chen, P.Y. *et al*. (2005) The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes Dev.* 19, 1288–1293
30 Ambros, V. *et al*. (2003) MicroRNAs and other tiny endogenous RNAs in *C. elegans*. *Curr. Biol.* 13, 807–818
31 Kim, V.N. (2005) Small RNAs: classification, biogenesis, and function. *Mol. Cells* 19, 1–15
32 Ambros, V. *et al*. (2003) A uniform system for microRNA annotation. *RNA* 9, 277–279
33 Seitz, H. *et al*. (2003) Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene. *Nat. Genet.* 34, 261–262
34 Ohler, U. *et al*. (2004) Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification. *RNA* 10, 1309–1322

35 Nam, J.W. *et al*. (2005) Human microRNA prediction through a probabilistic co-learning model of sequence and structure. *Nucleic Acids Res.* 33, 3570–3581
36 Lee, Y. *et al*. (2002) MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.* 21, 4663–4670
37 Rodriguez, A. *et al*. (2004) Identification of mammalian microRNA host genes and transcription units. *Genome Res.* 14, 1902–1910
38 Baskerville, S. and Bartel, D.P. (2005) Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* 11, 241–247
39 Tam, W. (2001) Identification and characterization of human *BIC*, a gene on chromosome 21 that encodes a noncoding RNA. *Gene* 274, 157–167
40 Aukerman, M.J. and Sakai, H. (2003) Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes. *Plant Cell* 15, 2730–2741
41 Johnson, S.M. *et al*. (2003) The time of appearance of the *C. elegans* let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. *Dev. Biol.* 259, 364–379
42 Bracht, J. *et al*. (2004) Trans-splicing and polyadenylation of let-7 microRNA primary transcripts. *RNA* 10, 1586–1594
43 Lee, Y. *et al*. (2004) MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* 23, 4051–4060
44 Cai, X. *et al*. (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* 10, 1957–1966
45 Xie, Z. *et al*. (2005) Expression of *Arabidopsis* MIRNA genes. *Plant Physiol.* 138, 2145–2152
46 Houbaviy, H.B. *et al*. (2005) Characterization of a highly variable eutherian microRNA gene. *RNA* 11, 1245–1257
47 Pfeffer, S. *et al*. (2005) Identification of microRNAs of the herpesvirus family. *Nat. Methods* 2, 269–276
48 O'Donnell, K.A. *et al*. (2005) c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* 435, 839–843
49 Zhao, Y. *et al*. (2005) Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* 436, 214–220
50 Lee, Y. *et al*. (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415–419
51 Filippov, V. *et al*. (2000) A novel type of RNase III family proteins in eukaryotes. *Gene* 245, 213–221
52 Fortin, K.R. *et al.* (2002) Mouse ribonuclease III. cDNA structure, expression analysis, and chromosomal location. *BMC Genomics* DOI:10.1186/1471-2164-3-26 (http://www.biomedcentral.com/1471-2164/3/26)
53 Wu, H. *et al*. (2000) Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. *J. Biol. Chem.* 275, 36957–36965
54 Denli, A.M. *et al*. (2004) Processing of primary microRNAs by the Microprocessor complex. *Nature* 432, 231–235
55 Gregory, R.I. *et al*. (2004) The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432, 235–240
56 Han, J. *et al*. (2004) The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.* 18, 3016–3027
57 Landthaler, M. *et al*. (2004) The human DiGeorge syndrome critical region gene 8 and its *D. melanogaster* homolog are required for miRNA biogenesis. *Curr. Biol.* 14, 2162–2167
58 Lund, E. *et al*. (2004) Nuclear export of microRNA precursors. *Science* 303, 95–98
59 Yi, R. *et al*. (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* 17, 3011–3016
60 Bohnsack, M.T. *et al*. (2004) Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* 10, 185–191
61 Gwizdek, C. *et al*. (2003) Exportin-5 mediates nuclear export of minihelix-containing RNAs. *J. Biol. Chem.* 278, 5505–5508
62 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
63 Bernstein, E. *et al*. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363–366
64 Grishok, A. *et al*. (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
65 Hutvagner, G. *et al*. (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293, 834–838
66 Ketting, R.F. *et al*. (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* 15, 2654–2659
67 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
68 Leuschner, P.J. *et al*. (2005) MicroRNAs: Loquacious speaks out. *Curr. Biol.* 15, R603–R605
69 Forstemann, K. *et al.* (2005) Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biol.* 3, e236 DOI:10.1371/journal.pbio.0030236 (http://biology.plosjournals.org)
70 Saito, K. *et al.* (2005) Processing of pre-microRNAs by the Dicer-1-Loquacious complex in *Drosophila* cells. *PLoS Biol.* 3, e235 DOI: 10.1371/journal.pbio.0030235 (http://biology.plosjournals.org)
71 Chendrimada, T.P. *et al*. (2005) TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* 436, 740–744
72 Zhang, H. *et al*. (2002) Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO J.* 21, 5875–5885
73 Zhang, H. *et al*. (2004) Single processing center models for human Dicer and bacterial RNase III. *Cell* 118, 57–68
74 Liu, Q. *et al*. (2003) R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway. *Science* 301, 1921–1925
75 Schwarz, D.S. *et al*. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115, 199–208
76 Khvorova, A. *et al*. (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115, 209–216
77 Tomari, Y. *et al*. (2004) A protein sensor for siRNA asymmetry. *Science* 306, 1377–1380
78 Krichevsky, A.M. *et al*. (2003) A microRNA array reveals extensive regulation of microRNAs during brain development. *RNA* 9, 1274–1281
79 Sempere, L.F. *et al*. (2003) Temporal regulation of microRNA expression in *Drosophila melanogaster* mediated by hormonal signals and broad-complex gene activity. *Dev. Biol.* 259, 9–18
80 Sempere, L.F. *et al*. (2004) Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. *Genome Biol.* 5, R13
81 Calin, G.A. *et al*. (2004) MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc. Natl. Acad. Sci. U. S. A.* 101, 11755–11760
82 Liu, C.G. *et al*. (2004) An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. *Proc. Natl. Acad. Sci. U. S. A.* 101, 9740–9744
83 Schmittgen, T.D. *et al*. (2004) A high-throughput method to monitor the expression of microRNA precursors. *Nucleic Acids Res.* 32, e43
84 Babak, T. *et al*. (2004) Probing microRNAs with microarrays: tissue specificity and functional inference. *RNA* 10, 1813–1819
85 Barad, O. *et al*. (2004) MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. *Genome Res.* 14, 2486–2494
86 Miska, E.A. *et al*. (2004) Microarray analysis of microRNA expression in the developing mammalian brain. *Genome Biol.* 5, R68
87 Sun, Y. *et al*. (2004) Development of a micro-array to detect human and mouse microRNAs and characterization of expression in human organs. *Nucleic Acids Res.* 32, e188
88 Wienholds, E. *et al*. (2005) MicroRNA expression in zebrafish embryonic development. *Science* 309, 310–311
89 Chen, C. *et al.* (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucelic Acids Res*. 33, DOI:10.1093/nar/gni178 (http://nar.oxfordjournals.org)
90 Lu, J. *et al*. (2005) MicroRNA expression profiles classify human cancers. *Nature* 435, 834–838
91 Monticelli, S. *et al*. (2005) MicroRNA profiling of the murine hematopoietic system. *Genome Biol.* 6, R71
92 Nelson, P.T. *et al*. (2004) Microarray-based, high-throughput gene expression profiling of microRNAs. *Nat. Methods* 1, 155–161

93 Thomson, J.M. *et al*. (2004) A custom microarray platform for analysis of microRNA gene expression. *Nat. Methods* 1, 47–53
94 Shingara, J. *et al*. (2005) An optimized isolation and labeling platform for accurate microRNA expression profiling. *RNA* 11, 1461–1470
95 Lewis, B.P. *et al*. (2003) Prediction of mammalian microRNA targets. *Cell* 115, 787–798
96 Lewis, B.P. *et al*. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120, 15–20
97 Wienholds, E. and Plasterk, R.H. (2005) MicroRNA function in animal development. *FEBS Lett.* 579, 5911–5922
98 Kidner, C.A. and Martienssen, R.A. (2005) The developmental role of microRNA in plants. *Curr. Opin. Plant Biol.* 8, 38–44
99 Bartel, D.P. and Chen, C.Z. (2004) Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat. Rev. Genet.* 5, 396–400
100 Lim, L.P. *et al*. (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 433, 769–773
101 Croce, C.M. and Calin, G.A. (2005) miRNAs, cancer, and stem cell division. *Cell* 122, 6–7
102 Johnson, S.M. *et al*. (2005) RAS is regulated by the let-7 microRNA family. *Cell* 120, 635–647
103 Cimmino, A. *et al*. (2005) miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc. Natl. Acad. Sci. U. S. A.* 102, 13944–13949
104 He, L. *et al*. (2005) A microRNA polycistron as a potential human oncogene. *Nature* 435, 828–833
105 Wang, X.J. *et al*. (2004) Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. *Genome Biol.* 5, R65
106 Legendre, M. *et al*. (2005) Profile-based detection of microRNA precursors in animal genomes. *Bioinformatics* 21, 841–845
107 Adai, A. *et al*. (2005) Computational prediction of miRNAs in *Arabidopsis thaliana*. *Genome Res.* 15, 78–91
108 Wang, X. *et al*. (2005) MicroRNA identification based on sequence and structure alignment. *Bioinformatics* 21, 3610–3614

AGORA initiative provides free agriculture journals to developing countries

The *Health Internetwork Access to Research Initiative* (HINARI) of the WHO has launched a new community scheme with the UN Food and Agriculture Organization.

As part of this enterprise, Elsevier has given 185 journals to *Access to Global Online Research in Agriculture* (AGORA). More than 100 institutions are now registered for the scheme, which aims to provide developing countries with free access to vital research that will ultimately help increase crop yields and encourage agricultural self-sufficiency.

According to the Africa University in Zimbabwe, AGORA has been welcomed by both students and staff. 'It has brought a wealth of information to our fingertips' says Vimbai Hungwe. 'The information made available goes a long way in helping the learning, teaching and research activities within the University. Given the economic hardships we are going through, it couldn't have come at a better time.'

For more information visit:
http://www.healthinternetwork.net