Available online at www.sciencedirect.com
ScienceDirect
Biochemical and Biophysical Research Communications 363 (2007) 6–11
BBRC
www.elsevier.com/locate/ybbrc

Mini Review

miRNomics—The bioinformatics of microRNA genes

Zhumur Ghosh [a,1], Jayprokas Chakrabarti [a,b], Bibekanand Mallick [a,*,1]

[a] Computational Biology Group, Indian Association for the Cultivation of Science, 2A & 2B, Raja S.C. Mullick Road, Jadavpur, Calcutta 700032, India
[b] Gyanxet, BF-286, Salt Lake, Calcutta 700064, India

Received 27 July 2007
Available online 13 August 2007

Abstract

MicroRNAs (miRNAs) are tiny genetic rheostats in plants, animals, and viruses, regulating the expression of messenger RNAs by targeting transcripts for cleavage or translational repression. Their regulatory impact is even more pervasive as a potential therapeutic tool. Since inception, computational methods have been an invaluable tool complementing experimental approaches. Here, we outline miRNA-bioinformatics highlighting the biological and therapeutic repertoire of miRNAs, in silico prediction of miRNA genes and their targets, along with a glimpse of the bioinformatic challenges that lie ahead.

Keywords: miRNA; Bioinformatics; MicroRNA; Cancer; Noncoding RNA; Stem cell; Apoptosis; miRNA target

One of the important recent advancements in biology research has been the discovery of the versatile RNA molecules that regulate expression of genes. For years, RNAs were thought to have just two major functions in cells. The coding RNAs are essential intermediaries in gene expression and noncoding RNAs (ribosomal and transfer RNAs) have structural, catalytic, and information-decoding roles in protein synthesis. The path-breaking discovery by Andrew Fire, Craig Mello, and their colleagues in 1998 obligated people to think otherwise with their discovery of RNA interference (RNAi)—the silencing of genes by noncoding RNAs (ncRNAs) [1]. Although miniscule molecules, 19–25 nucleotides (nts) in length, miRNAs have large-scale effects due to their regulatory functions. These are encoded in a wide variety of organisms from viruses and unicellular eukaryotes to multicellular eukaryotes [2,3] and comprise 1–3% of the genome [4].

lin-4 is the first founding member of the miRNA family which was discovered in 1993 by Lee et al. in *Caenorhabditis elegans* [5]. Downregulation of the protein lin-14 by lin-4 regulated the timings of larval development in *C. elegans*. let-7 (lethal-7), the second miRNA, was also discovered in the same worm seven years later by Reinhart et al. [6]. let-7 mediates translational repression of lin-41, the genes involved in larval development of *C. elegans*. With the discovery of the homologs of *let-7* in flies and humans [7], miRNAs were considered to represent conserved novel gene regulators. Thereafter this new regulatory RNA field began to garner attention. Nowadays, these genes form the pervading and widespread feature in plant and animal genomes [8–10]. Hundreds of miRNAs have been cloned in several organisms [11–13]. Computational prediction too has been an efficient strategy to find low expression or tissue specific miRNA genes [14]. The first miRNAs detected in a viral genome were reported in Science 2004 by Pfeffer and colleagues [2] in Epstein–Barr virus (EBV). The discovery of miRNAs in EBVs by Pfeffer et al. [2] encouraged the researchers a lot to look for miRNAs in other viruses too.

Identification and validation of novel miRNAs have become one of the most challenging problems in understanding miRNA mediated gene regulation. The initial approach of random cloning and sequencing leads towards the identification of few hundreds of miRNAs. This directed to a premature conclusion that the repertoire of such miRNAs is limited. Recently, deep sequencing approaches

* Corresponding author. Fax: +91 33 24732805.
E-mail addresses: tpbm@mahendra.iacs.res.in, vivekm@iitian.iitkgp.ernet.in (B. Mallick).
[1] These authors contributed equally to this work.

0006-291X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2007.08.030

have not only revealed that there are lot more miRNAs, but have also rendered the inefficiency and diminishing returns of such approaches to their discovery. Thus, an integrative approach combining computational/bioinformatic predictions with experimental techniques gives a complete view of the miRNA biology [15,16]. In this review, we focus mainly on the computational approaches towards getting an insight into miRNA gene identification, target prediction. Further we present the therapeutic potentials of these RNAs along with a perspective of the hibernating challenges in miRNA-bioinformatics.

miRNA biogenesis and function

miRNA genes are first transcribed from different genomic locations (within introns of genes or outside genes or in polycistronic clusters or exist individually) [17] as primary transcripts (pri-miRNAs) (refer to Fig. 1) by RNA polymerase II. miRNAs residing in the intron of a host gene are processed by sharing the same promoter and other regulatory elements of the host gene. Further, clustered miRNA genes in polycistronic transcripts are likely to be coordinately regulated [14]. The pri-miRNAs are processed to precursor miRNAs (pre-miRNAs) by the RNase endonuclease Drosha in the nucleus [18]. These pre-miRNAs are ~60 to ~100 nts with a stem-loop secondary structure. Specific RNA cleavage by Drosha predetermines the mature miRNA sequence and provides the substrates for subsequent processing steps. Recently, it has been shown that miRNAs can also be produced by Drosha independent pathway from excised introns [19]. Pre-miRNAs get imported to the cytoplasm from the nucleus by Exportin-5 and are cleaved by the cytoplasmic ribonuclease dicer to form double stranded mature miRNAs in animals. On the contrary, plant miRNAs are cleaved into miRNA:miRNA* duplex possibly by dicer-like enzyme 1 (DCL1) in the nucleus rather than in the cytoplasm [14,18], then the duplex is translocated into the cytoplasm by HASTY, the plant ortholog of exportin 5 [14]. The strands of this duplex separate and release mature miRNA of 19–25 nt in length [14,18]. Plant miRNAs undergo further modification by methylation at the 3′ end by HEN1 [20]. One strand of the duplex becomes a mature miRNA. Mature miRNAs gets integrated into a protein–RNA complex (miRNP). miRNAs within the complex interact with their target mRNA by base pairing. miRNAs can work by two modes-mRNA cleavage or translational repression without RNA cleavage. The degree of complementarity determines whether the target mRNA is cleaved or not. In plants, miRNAs base pair with their mRNA targets by precise or nearly precise complementarity. In contrast, most animal miRNAs bear imprecise complementarity to their mRNA targets [21]. A single miRNA has the potential to recognize several target sequences in the 3′ UTR and cause translation inhibition of many different genes in animals [21].

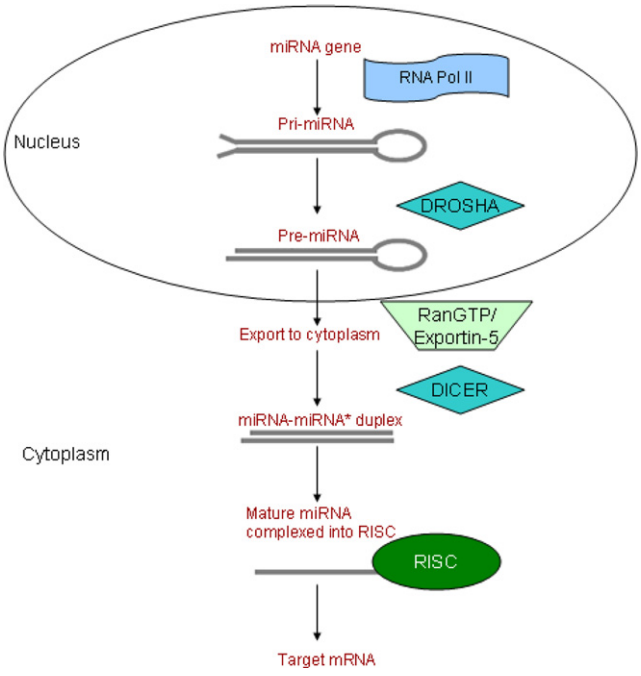

Fig. 1. The steps involved in biogenesis of miRNAs in animals.

MicroRNAs perform many cellular processes in animals such as developmental timing, cell death, hematopoiesis, and patterning of the nervous system [22]. Lin-4 and let-7 of *C. elegans* play essential roles in controlling timing events during larval development. The tissue restricted expression of many miRNAs reveals their probable role in the determination or maintenance of cell lineage [23]. MicroRNA miR-196 regulates the homeobox transcription factor HoxB8 which indicates its role in development [23]. Furthermore miR-1 plays a crucial role in the development of the heart and skeletal muscle. miRNAs are also linked to the regulation of insulin secretion. All these examples imply the importance of miRNA in diverse cellular processes. Consequently, dysregulation of miRNA function might lead to various diseases. Tourette's syndrome is a neurological disorder whose probable cause lies in mistargeting of a miRNA, if one of its targets has been mutated [24]. A single nucleotide polymorphism (SNP) that correlates with this disease is located in the 3′ untranslated region (3′ UTR) of the gene SLITRK1 (SLIT and TRK-like 1). This SNP is located within the binding site for miR-189, rendering SLITRK1 a better target of the miRNA.

Cancer is a consequence of disordered genome function. More than 200 miRNA sequences discovered in the human genome contribute to the development of cancer [25]. Levels of some miRNAs are altered in cancer. Let-7 is less expressed in human lung cancers [25]. Lower survival rate is observed in patients with diminished expression of let-7. Overexpression of let-7 inhibits growth of lung cancer cells in vitro [26]. Role of miRNA in leukemia is even more significant. A majority of DNA alterations occur in a region on chromosome 13 that are associated with mantle cell lymphoma and B cell chronic lymphocytic leukemia

(B-CLL) [27]. Interestingly, miRNAs hsa-miR-15a and hsamiR-16-1 reside in this region of human chromosome 13, which is now limited to about 30 kb, and lies between exons 2 and 5 of the LEU2 gene. LEU2 has been excluded as the tumor suppressor gene. Further, hsa-miR-15a and hsamiR-16-1 are found to be downregulated or deleted in 70% of tumor cells from patients with B-CLL. A role of these miRNAs in oncogenesis has been proposed [27]. Studies examining the specific role for miRNAs in regulating genes in mammalian cells that may be involved in cancer associated process have only just been initiated. A recent report shows that disrupting miRNA regulation of Hmga2 gene in human enhances oncogenic transformation [28]. Previous studies suggest that hsa-miR-21 found on chr. 17 are overexpressed in glioblastoma and have an anti-apoptotic function [29]. miRNAs are not only implicated in disease but are also associated with disease states. Human disease fragile X syndrome is caused by the loss of the fragile X metal retardation protein (FMRP). This is a common form of mental retardation.

The process of cell proliferation and suppression of apoptosis are found to be controlled by miRNAs in *Drosophila melanogaster* [30]. The post-transcriptional control of Notch signaling pathway in *D. melanogaster*, an evolutionary conserved signal transduction cascade that is required for patterning and normal development, is performed by miRNAs. miRNA genes are preferentially expressed in stem cells and may be involved in maintaining the pluripotent state. Three miRNAs are identified by Chen et al. [31] which are specifically expressed in hematopoietic cells. miR-296 is expressed in mouse embryonic stem cells, but not in differentiated cells and adult tissues. On the contrary, some other miRNAs such as miR-15a, miR-16, and miR-19b are found in both embryonic stem cells and adult tissues. Several of the novel miRNA genes identified are homologous to those cloned from mouse ES cells. These miRNA specific expression profiles suggest that miRNAs may have important roles in maintaining mammalian [32] stem cell pluripotency and it will be interesting to see if these stem cell-specific genes are involved in this process.

About two-thirds of the known plant miRNAs control the expression of translation factors that regulate crucial steps during plant development. Plant miRNAs have developmental role such as in phase changes, leaf morphogenesis, etc. [33].

Sources of miRNA data

The miRBase::Sequence database [34] (http://microrna.sanger.ac.uk/sequences) is the most popular miRNA database maintained by the Sanger Institute. This database contains all published mature miRNA sequences, together with their predicted hairpin precursors and annotation related to their discovery, structure, and function. The total miRNA entries in this database are 5071 as of August, 2007 (Release 10.0). It has three sections based on their function. (a) *miRBase Sequence Database* is a searchable database of published miRNA sequences and annotation. (b) *miRBase registry* provides unique names for novel miRNA genes prior to publication of results. The names are assigned by the Registry as sequential numerical identifiers. (c) *miRBase targets database*, a database of computationally predicted miRNA targets in animals. This database use a novel fully automated pipeline to predict miRNA targets. Other sources of miRNA include TarBase, Argonaute, and miRGen (refer Table 2b provided as Supplementary material). TarBase offers a manually curated and comprehensive set of experimentally supported targets in different species. Argonaute is a comprehensive database of mammalian miRNAs along with their origin and regulated target genes. It collects latest information from both literature and other databases. miRGen is a database for the study of animal microRNA genomic organization and function.

Structural features of miRNA genes—basis of computational prediction

The precursor miRNAs have a well-predicted stable extended stem-loop hairpin structure with continuous helical pairing and a few internal bulges. The length of hairpin stem-loops of pre-miRNAs is generally longer in plants than that of animals. In animals these precursors are 60–80 nucleotides in length [7] whereas in plants their length ranges from 60 to more than 400 nucleotides [32]. The lengths of viral pre-miRNAs vary from 60 to 119 with an average of 79 nucleotides [35]. The molecular features of plant pre-miRNAs are also different from those of animal miRNA precursors, having more variable predicted secondary structures. To identify a precursor structure, RNA folding program such as the Vienna package, Mfold or RNA analyzer is used. As mentioned earlier, precursor should have a low free energy ($\leqslant -20$ kcal/mol) and the precursor should be unbranched. Zheng et al. [36] proposed 43 features, 11 global features, and 32 local features of pre-miRNAs based upon the nucleotide sequences and secondary structure features. The global features are symmetric difference, number of base pairs, GC content, length base pair ratio (length of the sequence/the number of base pairs), sequence length, length of central loop, free energy per nucleotide, bulge size, bulge number, tail length, and the number of tail(s). The local features include local structural triplet elements proposed by Xue et al. [37].

Mature miRNAs are usually highly conserved among the genomes of related species. Viral miRNAs are less conserved than plant and animal miRNAs, suggesting that viral miRNAs may evolve rapidly. Different animal miRNAs are often clustered within a single precursor RNA, whereas plant miRNAs derive from individual precursors. These are usually 20–22 nucleotides in animals and 20–24 nucleotides in plants [14] derived from pre-miRNAs (shown below in Fig. 2). Mature miRNAs are present in either arm of the hairpin structure and lack large internal

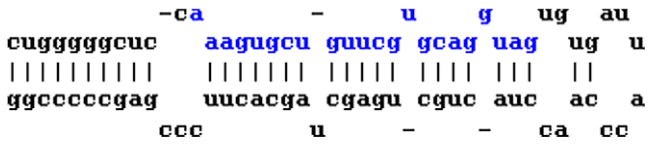

Fig. 2. Sequence and structure of a pre-miRNA encoding a mature miRNA (depicted in blue letters). Here, the pre-miRNA is hsa-mir-93 (Accession No. MI0000095, http://microrna.sanger.ac.uk/cgi-bin/sequences/mirna_entry.pl?acc=MI0000095), located on the complementary strand of chromosome 7 of *Homo sapiens* at 99529327–99529406. This precursor encodes the mature miRNA (hsa-miR-93) which is located between 12 and 33 bases on the upper arm of hsa-mir-93. (For interpretation of color mentioned in this figure the reader is referred to the web version of the article.)

loops or bulges. These are never located on the terminal loop of the hairpin structure [38]. Profiles of mature miRNAs of the 5′ and 3′ stem-loop precursor arms have the following features:

- In vertebrate organisms, 5′ arm motif has a conserved U at the beginning of the sequence and a GU rich region in the 3′ end (of the 5′ arm) around position 18–25. On the contrary, the 3′ arm motif essentially only shows U at the beginning of the mature sequence. Further, the 5′ arm motif also contains an upstream semi-conserved G at position 13. For invertebrates 3′ arm motif is more characteristic showing a highly conserved U at the mature start [39]. In case of plants [39] both the 5′ and the 3′ arm motifs show characteristic, but different, motifs. The 5′ arm motif having a strongly conserved U and the 3′ arm a conserved C at the mature end.
- In animals, mature miRNA shows near complementarity to the 3′ untranslated (3′ UTR) regions of target mRNAs [24]. Moreover, in plants miRNA–target interaction is more precise showing a perfect complementarity to the target mRNA sequence.

None of these properties if used separately is specific enough to accurately predict miRNA genes. However, if they are used together, dependable predictions can be made.

These structure based information help in the identification of unknown miRNA genes or prediction of its structure with the help of various folding algorithms-mfold [40], RNAfold [41], and RNA Analyzer [42]. Apart from secondary structure information, many methods rely on phylogenetic conservation of both sequence and structure, thermodynamic stability of hairpins and sequence, and structural similarity to known miRNAs, or use information on genomic location relative to known miRNAs. The different computational techniques used to find miRNA genes in viruses, plants, and animals are discussed in Table 1 (Supplementary material).

Computational prediction of miRNA targets

The identification of miRNA targets is an essential step towards understanding their regulatory function. Biological functions have only been assigned to few miRNAs, mostly due to the difficulty of miRNA target identification. Computational target prediction in plants is comparatively straightforward and powerful because miRNA and target mRNA are nearly perfectly complementary [14,36]. On the contrary, bioinformatic prediction of miRNA targets in animals is a difficult task as functional duplexes in animals can be more variable in structure: they contain only short complementary sequence stretches, interrupted by gaps, and mismatches. To date, specific rules for functional miRNA–target pairing that capture all known functional targets have not been devised. This has created problems for search strategies, which apply different assumptions about how to best identify functional sites. Still, different computational approaches have been developed and used to unveil such animal miRNA targets in genomes [43]. The large majority of these methods are based on the nature of the pairing between the miRNA and the target gene in animals and plants. Lists of computational methods and databases of miRNA targets have been enlisted in Table 2 (Supplementary material).

Target prediction in plants

Because of the near perfect complementarity of miRNAs with their targets, the prediction in plants is easier and straightforward. In plants, miRNA targets are mainly transcription factors. These miRNA-regulated transcription factors regulate developmental patterning, cell proliferation, environmental and hormonal responses [44].

Target prediction in animals and viruses

Animal miRNA targets appear to be more functionally diverse than their plant counterparts [14]. Criteria for developing these computational pipelines include the following:

- The miRNA sequence is complementary to the 3′ UTR sequence of potential target mRNAs. The complementarity is in the form of a "seed" of 7 or more nt at 5′ end of the miRNA to bind to the 3′ UTR sequence. This is essential to confer regulation of miRNA target. This seed has been described as crucial to target repression [38,45]. A requirement for matches outside of the seed has been required as well [46].
- Sites with weaker 5′ complementarity require compensatory pairing to the 3′ end of miRNA in order to confer regulation [47] whereas extensive pairing to the 3′ end of miRNA is not sufficient to confer regulation on its own without a minimal element of 5′ complementarity.
- The number and location of G:U wobbles has been used by target prediction programs to estimate the stability of a sequence, and therefore its likelihood. A G:U wobble occurs when guanine and uracil base-pair and form a hydrogen bond with each other. Such a wobble creates a local distortion in RNA shape and can reduce binding

and therefore the effectiveness of the silencing mechanism if inside the 50 seed region on a miRNA [45,47]. While G–U wobbles are reported to have no effect outside of a seed, Vella et al. [46] report that a G:U wobble is critical for downregulation of lin-41. G:U wobbles also provide recognition sites for protein and RNA enzymes and RNA-binding proteins. Therefore, a small number of G:U wobbles in the seed are expected of an miRNA, but at some point a larger number of wobbles reduce the likelihood of a sequence belonging to an miRNA.

- Cooperative translational control is observed where more than one miRNA typically regulates one message, whereas one miRNA may have several target genes, reflecting target multiplicity. This implies combinatorial control of a single target by multiple miRNAs. This can be an important feature of miRNA targeting, very similar to the mode of transcription factor control of genes [45], and multiple binding sites for an miRNA on the 3′ UTR can increase the efficiency of RNA silencing.
- miRNAs could also target other miRNAs for silencing.
- The kinetics and thermodynamics of the association between miRNA and its target can be determined by RNA folding programs [48].

Integrative approach combining computational searches for known pairings and experimental conformations is commonly used to measure the validity of the approaches used for elucidating the targets. Fundamental differences exist in the methods used to measure conservation and prediction of single or multiple binding sites in miRNA targets [49,50], and also in the statistical approaches chosen.

Although animal miRNAs target similar genes, they further control over a large spectrum of diverse biological functions which includes protein metabolism, intracellular signaling cascades, etc. [49–51]. Some of these targets are often relevant in a disease context [51]. Table 2 (Supplementary material) gives a listing of target prediction algorithms and their basic working principle. As a precautionary note we must mention that the predictions of these algorithms vary significantly with the results showing an overlap of 10–20% between them. Computational prediction of targets is bound to improve in due course of time through refinement of the training sets, rules for complementarity and statistical approaches. Along with the *in silico* methods, experimental validation will be essential to benchmark any new predictive methods.

Conclusion

With the discovery of the regulatory role of miRNA in cancers, lots of fervor have been created among the scientists. It is probably tip of the iceberg. Its immense therapeutic potential needs to be harnessed further. The convergence of genomic technologies and the development of miRNA drugs designed against specific molecular targets will provide many opportunities for using bioinformatics to bridge the gap between biological knowledge and clinical therapy. It has marked the beginning of unveiling the versatility of these mini-molecules. Although a large number of miRNA genes have been predicted additional forays into the genome reveal scores of new genes. Additionally, genes can be linked to cancer via recurrent genomic or genetic abnormalities. Further, there lies a difficult problem in identifying miRNA targets as well as deciphering the exact mechanism of translational repression of mRNAs by these snippets. miRNA-bioinformatics can lead us towards quicker solution. Integrating large and disparate datasets, gene-level distinctions can be made between the different biological states represented by the data. Better understanding of *in vivo* mechanisms of miRNAs will lead us towards more precise *in silico* modeling of repression pathways and identification of miRNA genes and their targets. Biological validation of computational results is absolutely essential. Conversely, such biological validation often depends on computational techniques [7,14]. Hence, it is a symbiotic process. An iterative integration of both experimental and bioinformatic approaches can provide better insight into biogenesis, mechanisms, and functions of miRNAs.

Bioinformatic approaches and their associated methodologies can be applied across a range of technologies, facilitating rapid identification of new target leads for further experimental validation. In other words, bioinformatics can accelerate this silent race from RNA laboratory to clinic.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2007.08.030.

References

[1] S. Fire, S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver, C.C. Mello, Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*, Nature 391 (1998) 806–811.
[2] S. Pfeffer, M. Zavolan, F.A. Grasser, M. Chien, J.J. Russo, J. Ju, B. John, A.J. Enright, D. Marks, C. Sander, T. Tuschl, Identification of virus-encoded microRNAs, Science 304 (2004) 734–736.
[3] M.W. Jones-Rhoades, D.P. Bartel, B. Bartel, MicroRNAs and their regulatory roles in plants, Annu. Rev. Plant. Biol. 57 (2006) 19–53.
[4] Y. Zhao, D. Srivastava, A developmental view of microRNA function, Trends Biochem. Sci. 32 (2007) 189–197.
[5] R.C. Lee, R.L. Feinbaum, V. Ambros, The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14, Cell 75 (1993) 843–854.
[6] B.J. Reinhart, F.J. Slack, M. Basson, A.E. Pasquinelli, J.C. Bettinger, A.E. Rougvie, H.R. Horvitz, G. Ruvkun, The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*, Nature 403 (2000) 901–906.
[7] V. Ambros, B. Bartel, D.P. Bartel, C.B. Burge, J.C. Carrington, X. Chen, G. Dreyfuss, S.R. Eddy, S. Griffiths-Jones, M. Marshall, M. Matzke, G. Ruvkun, T. Tuschl, A uniform system for microRNA annotation, RNA 9 (2003) 277–279.
[8] Z. Mourelatos, J. Dostie, S. Paushkin, A. Sharma, B. Charroux, L. Abel, J. Rappsilber, M. Mann, G. Dreyfuss, miRNPs: a novel class of

Ribonucleo proteins containing numerous microRNAs, Genes Dev. 16 (2002) 720–728.
[9] C. Llave, K.D. Kasschau, M.A. Rector, J.C. Carrington, Endogenous and silencing-associated small RNAs in plants, Plant Cell 14 (2002) 1605–1619.
[10] M.Y. Park, G. Wu, A. Gonzalez-Sulser, H. Vaucheret, R.S. Poethig, Nuclear processing and export of microRNAs in Arabidopsis, Proc. Natl. Acad. Sci. USA 102 (2005) 3691–3696.
[11] M. Lagos-Quintana, R. Rauhut, W. Lendeckel, T. Tuschl, Identification of novel genes coding for small expressed RNAs, Science 294 (2001) 853–858.
[12] R.C. Lee, V. Ambros, An extensive class of small RNAs in *Caenorhabditis elegans*, Science 294 (2001) 862–864.
[13] N. Lau, L. Lim, E. Weinstein, D.P. Bartel, An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*, Science 294 (2001) 858–862.
[14] D.P. Bartel, MicroRNAs: genomics, biogenesis, mechanism, and function, Cell 116 (2004) 281–297.
[15] I. Bentwich, A. Avniel, Y. Karov, R. Aharonov, S. Gilad, O. Barad, A. Barzilai, P. Einat, U. Einav, E. Meiri, E. Sharon, Y. Spector, Z. Bentwich, Identification of hundreds of conserved and nonconserved human microRNAs, Nat. Genet. 37 (2005) 766–770.
[16] R. Sachidanandam, RNAi as a bioinformatics consumer, Brief. Bioinform. 6 (2005) 146–162.
[17] S.Y. Ying, S.L. Lin, Intronic microRNAs, Biochem. Biophys. Res. Commun. 326 (2005) 515–520.
[18] Y. Lee, K. Jeon, J.T. Lee, S. Kim, V.N. Kim, MicroRNA maturation: stepwise processing and subcellular localization, EMBO J. 21 (2002) 4663–4670.
[19] J.G. Ruby, C. Jan, D.P. Bartel, Intronic microRNA precursors that bypass Drosha processing, Nature 448 (2007) 83–86.
[20] B. Yu, Z. Yang, J. Li, S. Minakhina, M. Yang, R.W. Padgett, R. Steward, X. Chen, Methylation as a crucial step in plant microRNA biogenesis, Science 307 (2005) 932–935.
[21] B. Wang, T.M. Love, M.E. Call, J.G. Doench, C.D. Novina, Recapitulation of short RNA-directed translational gene silencing in vitro, Mol. Cell. 22 (2006) 553–560.
[22] K.S. Kosik, The neuronal microrna system, Nat. Rev. Neurosci. 7 (2006) 911–920.
[23] E. Wienholds, R.H. Plasterk, MicroRNA function in animal development, FEBS Lett. 579 (2005) 5911–5922.
[24] J.F. Abelson, K.Y. Kwan, B.J. O'Roak, D.Y. Baek, A.A. Stillman, T.M. Morgan, C.A. Mathews, D.L. Pauls, M.R. Rasin, M. Gunel, N.R. Davis, A.G. Ercan-Sencicek, D.H. Guez, J.A. Spertus, J.F. Leckman, L.S. Dure, R. Kurlan, H.S. Singer, D.L. Gilbert, A. Farhi, A. Louvi, R.P. Lifton, N. Sestan, M.W. State, Sequence variants in SLITRK1 are associated with Tourette's syndrome, Science 310 (2005) 317–320.
[25] G.A. Calin, C.M. Croce, MicroRNA-Cancer connection: the beginning of a new tale, Cancer Res. 66 (2006) 7390–7394.
[26] A.E. Pasquinelli, B.J. Reinhart, F. Slack, M.Q. Martindale, M.I. Kuroda, B. Maller, D.C. Hayward, EE. Ball, B. Degnan, P. Muller, J. Spring, A. Srinivasan, M. Fishman, J. Finnerty, J. Corbo, M. Levine, P. Leahy, E. Davidson, G. Ruvkun, Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA, Nature 408 (2000) 86–89.
[27] G.A. Calin, C.G. Liu, C. Sevignani, M. Ferracin, N. Felli, C.D. Dumitru, M. Shimizu, A. Cimmino, S. Zupo, M. Dono, M.L. Dell'Aquila, H. Alder, L. Rassenti, T.J. Kipps, F. Bullrich, M. Negrini, C.M. Croce, MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias, Proc. Natl. Acad. Sci. USA 101 (2004) 11755–11760.
[28] C. Mayr, M.T. Hemann, D.P. Bartel, Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation, Science 315 (2007) 1576–1579.
[29] J.A. Chan, A.M. Krichevsky, K.S. Kosik, MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells, Cancer Res. 65 (2005) 6029–6033.
[30] M. Jovanovic, M.O. Hengartner, miRNAs and apoptosis: RNAs to die for, Oncogene 25 (2006) 6176–6187.
[31] C.-Z. Chen, L. Li, H.F. Lodish, D.P. Bartel, MicroRNAs modulate hematopoietic lineage differentiation, Science 303 (2004) 83–86.
[32] N.R. Smalheiser, V.I. Torvik, Mammalian microRNAs derived from genomic repeats, Trends Genet. 21 (2005) 322–326.
[33] P.X. Zhang, G.P. Cobb, T.A. Anderson, Plant microRNA: a small regulatory molecule with big impact, Dev. Biol. 289 (2006) 3–16.
[34] S. Griffiths-Jones, R.J. Grocock, S. van Dongen, A. Bateman, A.J. Enright, miRBase: microRNA sequences, targets and gene nomenclature, Nucleic Acids Res. 34 (2006) D140–D144.
[35] X. Pan, B. Zhang, M. San Francisco, G.P. Cobb, Characterizing viral microRNAs and its application on identifying new microRNAs in viruses, J. Cell. Physiol. 211 (2007) 10–18.
[36] Y. Zheng, W. Hsu, M-Li. Lee, L. Wong, Exploring essential attributes for detecting MicroRNA precursors from background sequences, VDMB (2006) 131–145.
[37] C. Xue, F. Li, T. He, G. Liu, Y. Li, X. Zhang, Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine, BMC Bioinformatics 6 (2005) 310.
[38] S. Kim, J. Nam, J. Rhee, W. Lee, B. Zhang, miTarget: microRNA target gene prediction using a support vector machine, BMC Bioinformatics 7 (2006) 411.
[39] J. Gorodkin, J.H. Havgaard, M. Ensterö, M. Sawera, P. Jensen, M. Ohman, M. Fredholm, MicroRNA sequence motifs reveal asymmetry between the stem arms, Comput. Biol. Chem. 30 (2006) 249–254.
[40] M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction, Nucleic Acids Res. 31 (2003) 3406–3415.
[41] I.L. Hofacker, M. Fekete, P.F. Stadler, Secondary structure prediction for aligned RNA sequences, J. Mol. Biol. 319 (2002) 1059–1066.
[42] P. Bengert, T. Dandekar, A software tool-box for analysis of regulatory RNA elements, Nucleic Acids Res. 31 (2003) 3441–3445.
[43] H. Iwama, T. Masaki, S. Kuriyama, Abundance of microRNA target motifs in the 3/-UTRs of 20527 human genes, FEBS Lett. 581 (2007) 1805–1810.
[44] C.A. Kidner, R.A. Martienssen, The developmental role of microRNA in plants, Curr. Opin. Plant Biol. 8 (2005) 38–44.
[45] J.G. Doench, P.A. Sharp, Specificity of microRNA target selection in translational repression, Genes Dev. 18 (2004) 504–511.
[46] M.C. Vella, E.Y. Choi, S.Y. Lin, K. Reinert, F.J. Slack, The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3′ UTR, Genes Dev. 18 (2004) 132–137.
[47] J. Brennecke, A. Stark, R.B. Russell, S.M. Cohen, Principles of microRNA-target recognition, PLoS Biol. 3 (2005) e85.
[48] S. Wuchty, W. Fontana, I.L. Hofacker, P. Schuster, Complete suboptimal folding of RNA and the stability of secondary structures, Biopolymers 49 (1999) 45–165.
[49] M. Rehmsmeier, P. Steffen, M. Höchsmann, R. Giegerich, Fast and effective prediction of microRNA/target duplexes, RNA 10 (2004) 1507–1517.
[50] A.J. Enright, B. John, U. Gaul, T. Tuschl, C. Sander, D.S. Marks, MicroRNA targets in Drosophila, Genome Biol. 5 (2003).
[51] B.P. Lewis, C.B. Burge, D.P. Bartel, Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets, Cell 120 (2005) 15–20.