CHAPTER 4

MicroRNAs AS POST-TRANSCRIPTIONAL MACHINES AND THEIR INTERPLAY WITH CELLULAR NETWORKS

Sarath Chandra Janga* and Swathi Vallabhaneni

Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA
**Corresponding Author: Sarath Chandra Janga—Email: sarath@igb.uiuc.edu*

Abstract: Gene expression is a highly controlled process which is known to occur at several levels in eukaryotic organisms. Although RNAs have been traditionally viewed as passive molecules in the pathway from transcription to translation, there is increasing evidence that their metabolism is controlled by a class of small noncoding RNAs called MicroRNAs (miRNAs). MicroRNAs (miRNAs) control essential gene regulatory pathways in both plants and animals however our understanding about their function, evolution and interplay with other cellular components is only beginning to be elucidated. In this chapter, we provide an overview of the recent developments in our understanding of this class of RNAs from diverse perspectives including biogenesis, mechanism of their function, evolution of their clusters, and discuss the approaches currently available for the construction of post-transcriptional networks governed by them. We also present our current understanding on these post-transcriptional networks in the context other cellular networks. We finally argue that such developments would not only allow us to gain a deeper understanding of regulation at a level that has been under-appreciated over the past decades, but would also allow us to use the newly developed high-throughput approaches to interrogate the prevalence of these phenomena in different states, and thereby exploit the functions of these RNA molecules for therapeutic advantage in higher eukaryotes.

RNA Infrastructure and Networks, edited by Lesley J. Collins.

INTRODUCTION

For a long time, RNA was thought to have a supplementary role in genome function—an intermediate in the flow of biological information from gene to protein. In the past decade, however, the centrality of RNA as a key player in gene expression has become clear; it is far from being merely an intermediate between gene and protein. Instead, new insights into RNA structure and function have revolutionized our view of molecular biology, and the role of RNA in the evolution of life itself. According to the central dogma of molecular biology, RNAs are passive messengers and only take charge of transferring genetic information or carrying out DNA instructions, or code, for protein production in cells. However, this central dogma is getting challenged by the findings that tiny fragments of noncoding RNA typically ~22 nucleotides in length, namely microRNA (miRNA), are able to negatively regulate protein-coding genes by interfering with mRNA's original instructions. Recent studies indicate that miRNAs have emerged as central post-transcriptional repressors of gene expression. MicroRNAs suppress gene expression via imperfect base pairing to the 3′ untranslated region (3′ UTR) of target mRNAs leading to repression of protein production or mRNA degradation.[1-4] These noncoding regulatory RNA molecules have been found in diverse plants, animals, some viruses and even algae, and it now seems likely that all multicellular eukaryotes and perhaps some unicellular eukaryotes utilize them to regulate gene expression. It is currently estimated that miRNAs account for about ~1% of the predicted genes in higher eukaryotes and up to 10-30% of their genes might be regulated by miRNAs.[1]

MicroRNAs regulate many key biological processes, including cell growth, development, proliferation, differentiation, apoptosis and metabolism by controlling the expression of the genes they bind to at the post-transcriptional level. Indeed, animals lacking some mature miRNAs are unable to survive or reproduce.[5-7] For instance, a single malfunctioning microRNA can be sufficient to cause cancer in mice.[8] These discoveries in the recent past have offered new insights into another layer of gene regulation and at the same time underscore the important role noncoding mRNAs play in the cells. MicroRNAs have biological functions in many aspects, with targets ranging from signaling proteins, metabolic enzymes to transcription factors. This diversity and abundance in their target repertoire offers an enormous level of combinatorial possibilities, and suggest that miRNAs and their targets can be densely interconnected with other cellular networks. Thus, it is reasonable to think that miRNAs exert their function by regulating and integrating cellular networks at different levels in the cell.

In this chapter, we provide a comprehensive overview of this quickly developing area of post-transcriptional regulatory networks formed by miRNAs. We organize it into the following major sections: (1) biogenesis, genomic organization and evolution of miRNAs; (2) methods used for the identification of miRNAs and their targets; (3) interplay of miRNAs with other cellular networks and finally; (4) expression dynamics of miRNAs in the context of other regulatory factors. With increasing amount of evidence in support of the roles of miRNAs in diverse cellular processes in a variety of model systems, it is possible to envisage that miRNAs can form promising therapeutic targets in years to come.

BIOGENESIS, GENOMIC ORGANIZATION AND EVOLUTION OF miRNAs

RNA molecules have been thought to be important players in regulating gene expression in metazoans since a vast class of endogenously expressed small RNAs, termed microRNAs, were identified. The first miRNAs were discovered via genetic screening of the nematode *Caenorhabditis elegans*.[9,10] Genes encoding miRNAs contain many similarities and distinctions compared to protein-coding genes. Typically, one miRNA gene can encode only one particular mRNA; however, there are frequent exceptions where the coding sequences of some groups of miRNAs are clustered and expressed as a single polycistronic transcript.[11,12] MicroRNAs in the same cluster may not necessarily show sequence similarities or regulate the identical target genes, but they are often co-expressed.[13,14] Due to their co-expression, they function collaboratively to orchestrate the appropriate dosage of multiple gene products. In addition, if a new miRNA sequence should appear within an existing transcription unit, it immediately expresses its new product without invention or duplication of enhancers and promoters. This enables new miRNA genes to more easily appear without complete gene/transcript duplication and may account for the abundance of miRNA gene clusters. Indeed, many new animal miRNAs are thought to arise from accumulation of nucleotide sequence changes and not from gene duplication supporting this notion.[15] Most miRNAs are located in either the intergenic regions or the intron regions of genes encoding proteins or other noncoding RNAs in the antisense orientation, and expressed as independent transcription units.[11,12] Some intronic miRNA genes, however, are arranged in the sense orientation with respect to the protein or RNA-coding genes, and are cotranscribed with their precursor mRNAs.[14] This type of transcript is organized such that the miRNA sequence is located within an intron.

MicroRNAs are transcribed by RNA polymerase II as long precursor transcripts, which are called primary miRNAs (pri-miRNAs). The pri-miRNAs are capped and polyadenylated, and can reach several kilobases in length.[16] A single pri-miRNA might contain one, or up to several, miRNAs. Several sequential steps of transcript processing are required to produce mature miRNAs from pri-miRNAs. In the nucleus, there is a microprocessor complex in which the major components are the RNase-III enzyme Drosha and its partner DGCR8/Pasha,[17,18] which initially recognize pri-miRNAs and then excise the precursor miRNA (pre-miRNA), a 60-100 nucleotide intermediate that makes imperfect stem-loop hairpin structure containing the mature miRNA. Exportin-5, a nuclear export factor, recognizes and transports the pre-miRNAs to cytoplasm.[19,20] In the cytoplasm, Dicer, a second RNase-III enzyme, cleaves the pre-miRNAs to generate double-stranded 18-24 nucleotide long RNA molecules—mature miRNAs.[16,21,22] RNA-induced silencing complex (RISC), the core component of which is the argonaute protein,[16] incorporates one of these two strands—the guide strand of miRNAs. Finally, the miRNA guides the RISC complex to the target mRNA to suppress gene expression via imperfect base pairing to the 3′ UTR of target mRNAs, leading to repression of protein production or mRNA degradation.[1,4] Targeting by miRNAs is known to affect an mRNA's activity via different mechanisms. The earliest model claimed that an RNAi-like degradation mechanism occurs when the complementarity within miRNA-target duplex is nearly perfect, whereas imperfect pairing results in the suppression of translation possibly via the disruption of elongation[23] or 5′ capping of the mRNA.[24,25] This dichotomy has been severely challenged by miRNA-induced destabilization of mRNA half-life where the miRNA-mRNA target interaction is imperfect. These imperfect miRNA-target effects have indeed been documented in both humans and *C. elegans*.[26-29] However, a recent

study showed that changes in mRNA levels closely reflect the impact of miRNAs on gene expression and indicate that destabilization of target mRNAs is the predominant reason for reduced protein output in humans, with 84% of the decreased protein production accounted by the lowered mRNA levels.[30]

Evaluating conservation among species is a robust approach to find critical features within biological sequences and to understand the contribution they make to the existence of an organism. In this context, sequence homology can be found with very high frequency in most miRNAs across closely related species, and some miRNAs are ubiquitous in all the metazoans.[31] In addition, most miRNAs share almost identical biosynthetic pathways, precursor structure and effector mechanism. It is due to this high degree of similarity from biosynthesis to mechanism of action, conservation of miRNAs and their gene clusters have been studied from diverse perspectives across organisms. Members of a given miRNA family can be recovered from genome sequences due to the high sequence conservation of the mature miRNA and the characteristic stable hairpin structure of the precursor. Such a systematic analysis can be used for not only identifying miRNA homologs but to construct phylogenies of the family to understand the likely time of evolutionary origin. Several studies have used the continuously increasing repertoire of miRNAs in eukaryotes, documented in databases,[32] to show that miRNA evolution correlates with major body plan innovations in animal phylogeny.[33-35] These studies also propose that the diversity of the microRNA repertoire, complexity of their expression patterns and the diversity of the miRNA targets are correlated with an animal's morphological complexity. It is possible to imagine that miRNA's can act as important players in controlling large fractions of gene networks in a co-ordinated manner as they are known to influence almost all cellular processes. It is also increasingly becoming clear that lineage-specific microRNAs may account for phenotypic variation in closely related species.[33,36,37] For instance, in a large scale study comparing the miRNA repertoires of human and chimpanzee brains, Plasterk and colleagues[37] have shown that many of the then identified novel miRNAs are not conserved beyond primates, indicating their recent origin. They also found a significant fraction to be species-specific, whereas others are expanded in one of the species through duplication events. These studies support the notion that evolution of miRNAs is an ongoing process and that along with a small fraction of ancient, highly conserved miRNAs, there are a number of emerging miRNAs in the current genome sequences.[15]

The evolution of microRNAs is characterized not only by the continuing innovation of novel families but also by the diversification of established families spawning additional paralogous family members. Animal miRNAs are often organized in genomic clusters, usually indicating a single polycistronic primary precursor transcript, which may carry members of several distinct miRNA families. A second type of cluster organization comprising of many miRNAs which are transcribed independently or possibly in small groups has also been found. An example of this type is the C19MC cluster,[38] whose members are independently transcribed by pol-III. It is now becoming clear that miRNA families evolve through gene duplication and gene loss, among other mechanisms such as accumulation of mutations in the neighborhood of established miRNA gene clusters (discussed above).[15,39-41] Two distinct modes of duplication events have been proposed: (1) local duplications leading to additional copies on the same primary transcript, and (2) nonlocal duplications which eventually place the paralogs under different transcriptional control. Evidently, the cause for nonlocal duplications is mostly the whole-genome duplication events in early vertebrate evolution.[41] Mature miRNA paralogs usually acquire

minimal substitutions, suggesting that the differences in the functions among members of a miRNA family might be associated with the differences in their expression and/or downstream processing in order to affect their targets. In contrast to the small-sized but abundant number of miRNA families in the animal genomes, plants have been found to have fewer but larger miRNA families.[42] In addition, plants have also been found to harbor a large number of species-specific miRNAs, which often outnumber the conserved fraction.[41,43-45] Many of these species-specific miRNAs are single copy genes and show significant sequence similarity with their putative targets supporting the view that these miRNAs are indeed evolutionarily recent.

METHODS TO IDENTIFY miRNAs AND THEIR TARGETS

Computational methods have played an important role in the prediction of miRNAs from the very beginning.[11] Traditionally, some major features such as the hairpin-shaped stem loop structure, minimal folding free-energy and high evolutionary conservation have been used in the computational identification of miRNAs. In general, approaches for the identification of miRNAs can be broadly classified into three major classes: Approaches based on the features in the sequence and their conservation, approaches based on thermodynamic stability of the miRNAs and experimental data-driven approaches.

Most simple strategies employed for the identification of miRNAs fall into the first category and use sequence homology to experimentally known miRNAs as well as the characteristic hairpin structure of the pre-miRNA.[11] However, conservation is high at both sides of the stem region and is decreasing towards the unpaired region of the apical loop. If only one mature miRNA is produced from the precursor, the region encoding the mature sequences is best conserved. In some cases, both sides of the hairpin produce mature sequences in which case both are conserved equally well. Several software tools have been designed to take into account these features of miRNA sequences and their conservation (see Table 1). These approaches are typically used for the identification of miRNAs in a newly sequenced genome. For closely related species, phylogenetic shadowing has also been proposed to be a powerful approach to identify regions that are under stabilizing selection and hence exhibit the characteristic variations in sequence conservation between stems, loop and mature miRNA (see Chapter 2, Figure 3, page 24 for an example of an RNA comparative alignment).[46] Some approaches such as Mirscan-II have also exploited the genomic context of a miRNA under the notion that miRNAs usually appear in genomic clusters.[47,48] More complicated approaches in this category employ a number of different features contained in bona fide miRNAs, in order to predict novel ones.[11] MicroRNA detection without the aid of comparative sequence analysis is a very hard task but unavoidable when species-specific miRNAs are of prime interest. This is because a typical idea of the above approaches using comparative genomics is to filter out hairpins that are not evolutionary conserved in related species. Such a filtering step makes the methods unable to identify novel miRNAs when there is no known close homology either because of the limitation of the current data or due to the rapid evolution of miRNAs. The mir-abela approach, first searches for hairpins that are robust against changes in the folding windows and are also thermodynamically stable, and then uses a support vector machine to identify miRNAs among these filtered candidates.[49] A similar approach has been proposed by Xue et al.[50] These approaches fall into the second group of methods for predicting miRNAs. Some approaches use both

Table 1. Different resources and methods available for finding miRNA genes

Algorithm	Description of Features	Data Sets	Download/Web Server	References
Mir-abela	A support vector machine approach which computes 16 different statistics from the entire hair pin structure	Human	http:www.mirz.unibas.ch/cgi/pred_miRNA_genes.cgi	49
miPred	This approach uses 32 global and intrinsic hairpin folding attributes based on sequence, structural and folding thermodynamics	Human	http://bioinfo.au.tsinghua.edu.cn/mirnasvm/	93
RNAmicro	12 features based on structure, sequence, composition, thermodynamic stability and structural conservation are included in this approach	Animal	http://www.bioinf.uni-leipzig.de/~jana/index.php/jana-hertel-software/65-jana-hertel-rnamicro	94
Micro-ProcessorSVM	This approach includes 686 features from structure and sequence; 7 additional Drosha processing site features are also included	Human	http://demo1.interogan.com/miRNA/	95
BayesmiRNAfind	Secondary structure and sequence features are included in this approach	Mouse	https://bioinfo.wistar.upenn.edu/miRNA/miRNA/login.php	96
miPred	This approach uses structure—sequence composition and free energy of the secondary structure	Human	http://www.bioinfo.seu.edu.cn/miRNA	51
MiRFinder	This approach uses 18 parameters including secondary structure, differences of the stem region of miRNA and secondary structures of pre-miRNAs	Human	http://www.bioinfomatics.org/mirfinder/	97

sequence and thermodynamic aspects of miRNA to build integrated models to predict miRNAs.[51,52] Finally, experimental driven-approaches can be considered to be the result of the recent explosion in the deep-sequencing techniques, which allow the unambiguous identification of miRNAs in different populations. Since these technologies result in the production of large amount of sequence data, bioinformatics tools are essential to analyze and understand the resulting sequences. For instance, Mirdeep uses a probabilistic model of miRNA biogenesis to score the compatibility of the position and frequency of sequenced RNA with the secondary structure of the miRNA precursor. The main idea being to detect miRNAs by analyzing how sequenced RNAs are compatible with miRNA precursors processed in the cell, by comparing with published datasets.[53] Other approaches such as Miranalyzer fall into this category[54] (Table 2). More recent studies also report the accumulation of the processed next generation sequencing datasets into unified databases for efficient downstream analysis of post-transcriptional regulation in model organisms.[55,56]

MicroRNAs act as guide molecules that program the RISC complex to recognize a target miRNA, so it is essential to understand the mechanism by which miRNAs recognize their targets in order to predict target mRNAs for a given miRNA sequence. Currently, the number of miRNA-mRNA interactions is still small which makes it difficult for computational approaches to accurately predict the targets. For instance, Tarbase,[57] a database dedicated to the collection of experimentally identified miRNA-mRNA interactions currently reports 1333 mRNAs targeted by ~150 animal miRNAs. These known interactions have been used by a number of groups to develop rules for the prediction of miRNA targets in the form of a variety of computational methods (see Table 2). In plants, many targets can be predicted with confidence simply by searching for the messages with extensive complementarity to the miRNAs.[58] However, in animals, extensive complementarity is not a common phenomenon.[59]

Current prediction methods are diverse in both the adopted approach and performance (see Table 2). However, some major guiding principles have emerged over the years. These include (1) absence of a perfect complementarity between miRNA and mRNA; (2) miRNA:mRNA duplexes are asymmetric with the 5′ end of the miRNA binding more strongly than the 3′ end of the miRNA; (3) the region comprising positions 2-7 on the miRNA often exhibit perfect complementarity and is therefore commonly referred to as the seed region of miRNA; (4) target sites with evolutionary conserved seed regions tend to be more likely true sites and show stronger regulatory impact than nonconserved ones, and (5) highly conserved miRNAs have many conserved targets (see Fig. 1).[60-66]

MicroRNA INTERPLAY WITH OTHER CELLULAR NETWORKS

MicroRNAs have been shown to have biological functions in many aspects. Their targets range from signaling proteins, enzymes to transcription factors and RNA binding proteins. The diversity and abundance of miRNA targets offer an enormous level of combinatorial possibilities and suggest that miRNAs with their targets might form a rather complex regulatory network intertwined with other cellular networks such as signal transduction, metabolic, gene regulatory and protein interaction networks. Therefore, in order to understand the global principles of miRNA regulation, it is imperative to understand how they take part in different cellular processes at a systems level. Four

Table 2. Different methods and software implementations available for predicting miRNA targets from genomic sequences

Tool	Supported Organisms	Website	References
TargetScan, TargetScanS	Vertebrates, flies and worms	http://hollywood.mit.edu/burgelab/software	65
miRanda	Flies, Vertebrates	http://www.microrna.org/	98
DIANA-microT	Vertebrates	http://diana.cslab.ece.ntua.gr	99
RNA hybrid	Flies	http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/	100
PicTar	Nematodes, flies and vertebrates	http://pictar.mdc-berlin.de/	101
Rna22	Nematodes, flies and vertebrates	http://cbcsrv.watson.ibm.com/rna22.html	102
EIMMo	Vertebrates, flies and worms	http://www.mirz.unibas.ch./EIMMo2	103
PITA Top	Vertebrates, flies and worms	http://genie.weizmannac.il/pubs/mir07/mir07_data.html	104
TargetBoost	Worm and fruit fly	http://demo1.interagon.com/demo	105
mirWIP	*C. elegans*	http://www.psi.toronto.edu/genmir	106
miRDeep	Vertebrates, flies and worms	http://www.mdc-berlin.de/en/research/research_teams/systems_biology_of_gene_regulatory_elements/projects/miRDeep/index.html	53
miRanalyzer	Vertebrates, flies and worms	http://web.bioinformatics.cicbiogune.es/microRNA/	54

different kinds of networks are commonly studied within the context of a cell, namely gene regulatory, metabolic, signaling and protein interaction networks. In what follows, we will summarize our current understanding of the influence of miRNAs on these cellular networks.

Signaling network of a cell can be considered as a complex system responding to stimuli, signals and messages from other cells and environment. Once a cell receives signals, it processes the information and the signals reach transcription factors, thereby triggering responses of gene regulatory networks. In other words, the signaling network integrates the extra- and intra-cellular signals for the proper functioning of a cell. Signaling networks are typically represented as graphs containing both directed and undirected edges, with the nodes representing proteins. In such a representation directed links represent activation or inactivation relationships between proteins, while the undirected links represent physical interactions between proteins. In a large-scale analyses performed by Cui et al,[67] the authors collected publicly available dataset on signaling interactions in human[68] and extended it by manually curating additional signaling interactions to construct the largest known signaling interactome. By employing already available miRNA target predictions in human, Cui and coworkers

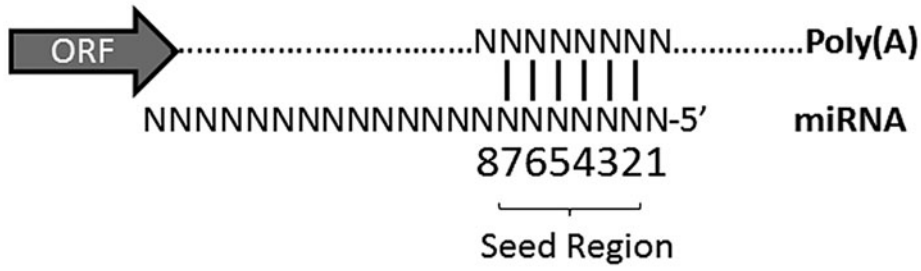

Figure 1. Representation of a miRNA-mRNA duplex with the seed region labeled. Nucleotide stretch numbered from 2 to 7 on the 5′ end of the miRNA is considered as its seed region and is often found to be completely complementary to the 3′ untranslated region (UTR) on the target transcript, where the miRNA binds and represses its activity by translational inhibition or by destabilization. Base complementarity is also found between position 1 and 8 on the miRNA and the target transcript in addition to the match on the seed region. The vertical dashes represent the match (complementarity) between the seed region and mRNA transcript using Watson-Crick base pairing and is commonly referred to as the canonical site. Other supplementary sites (in conjunction with the match on the seed region) on the 3′ region of the miRNA have also been found where there is complementarity between the transcripts however these are far less common than the canonical sites alone.[60] Multiple miRNAs can bind to the same UTR of the mRNA with members of the same miRNA family binding to identical sites, to work in a combinatorial fashion.

showed that the fraction of miRNA targets increases with the signal information flow from the upstream to the downstream in this network e.g., from ligands, cell surface receptors, intracellular signaling proteins to nuclear proteins. In other words, only 9.1% of the ligands were miRNA targets while half of the nuclear proteins, most of which were transcription factors were found to be miRNA targets. In addition, they noted that highly linked scaffold proteins such as CRK (sarcoma virus CT10 oncogene homolog) and SNAP25 (synaptosome-associated protein) have higher probability to be targeted by miRNAs. More generally, this study[67] showed that miRNAs avoid targeting common components of cellular pathways, possibly to avoid disturbing basic cellular processes because these common proteins are highly shared by basic cellular machines and should be frequently used in various cellular conditions. A network analysis of the signaling interactions allowed them to infer that miRNAs frequently target positively linked network motifs compared to negatively linked ones.[67] By selectively regulating positive regulatory motifs, scaffolding proteins and networks' downstream components, miRNAs may provide a mechanism to terminate the pre-existing messages and facilitate quick and robust transitions for responses to new signals.

Gene regulatory networks comprise of the regulatory interconnections between transcription factors (TFs), RNA binding proteins (RBPs), noncoding RNAs and their target genes or transcripts. In a simpler version of these networks only one type of regulatory molecule (either TF or RBP or RNA) is considered to target genes or transcripts to construct transcriptional or post-transcriptional networks, while more complex representations involve the integration of multiple layers of regulation. In one of the preliminary studies on the impact of miRNAs on the transcriptional regulatory network of a eukaryote, the authors accumulated publicly available TF binding data for three embryonic stem cell specific TFs—NANOG, SOX2 and OCT4,[69] to show that miRNA targets are significantly enriched in the genes that are regulated by many

transcription factors. They extended these observations to the complete genome by using computationally predicted binding sites in human to show that a gene that is regulated by a large number of TFs is also more likely to be regulated by many miRNAs.[70] This observation suggests that genes which are more complexly regulated at transcriptional level are required to be turned on more frequently and hence are more likely to be controlled under different conditions. Therefore, they are also more likely to be regulated by many miRNAs. Another area of intense research is to dissect the integrated network of transcriptional and post-transcriptional regulatory interactions, to unravel patterns of interconnections which appear more often than expected by chance—often referred to network motifs analogous to sequence motifs in genomes.[71,72] For instance, a commonly occurring network motif which was identified in a number of biological networks including transcriptional regulatory networks, where in TFs form the regulatory nodes and the genes transcriptionally controlled by them form the target nodes, is the Feed Forward Loop (FFL).[72,73] In this motif, a TF, X, regulates another TF, Y and a target gene, Z while the second TF, Y also regulates the activity of Z forming a loop. Initial efforts to study miRNA network motifs have focused on analyzing the co-operation between TFs and miRNAs by using computational approaches.[74] These computational techniques involved the identification of pairs of TF-miRNAs, which show high co-occurrence of their sequence motifs in a number of target genes and resulted in our understanding that miRNAs and TFs work together to form network motifs.[74-76] These studies have shown that two kinds of FFLs are prevalent in the integrated miRNA-TF regulatory networks: (1) Type I comprising of a TF activating/repressing both the miRNA and the target gene, while the miRNA represses the target transcript; and (2) Type II comprising of a TF activating the miRNA but repressing the expression of the target gene (or the other way around) with the miRNA also repressing the expression of the target. One example has been characterized experimentally for type I FFL—MicroRNA, miR-17-5p represses the target gene, E2F1 and both the miRNA and E2F1 are transcriptionally activated by c-Myc in human cells.[77] This type of motif has been proposed to prevent proliferation of noise-driven transitions.[77] In addition to these motif types, superposed or composite FFL motifs, where in the interacting miRNA-TF pairs have a bi-directional relationship between them and indirect feed-forward motifs have also been observed (Fig. 2). In the former, miRNA represses a TF and a target gene while the TF activates both the miRNA and the target. The later motif constitutes a TF activating another TF and a target gene; with the second TF activating a miRNA, which itself represses the target gene. However, the functions of these motifs have not been studied experimentally.

Protein interaction networks provide a snapshot of the proteome's functional organization by elucidating the links between pairs of proteins in a cell. Large-scale determination of these maps during the last decade in various model organisms has enabled us to exploit them to understand the cell's functional architecture and their interplay with the rest of the cellular constituents. In the context of miRNAs, one study showed that analogous to the transcriptional networks discussed above, proteins with more number of interacting partners are preferentially targeted by miRNAs.[78] In other words, protein connectivity in the human protein interaction network was shown to be positively correlated with the number of miRNA target-site types. This observation indicates that when a protein has more number of interacting partners, it will be regulated by many regulatory molecules. The study also showed that inter-modular hubs i.e., proteins which

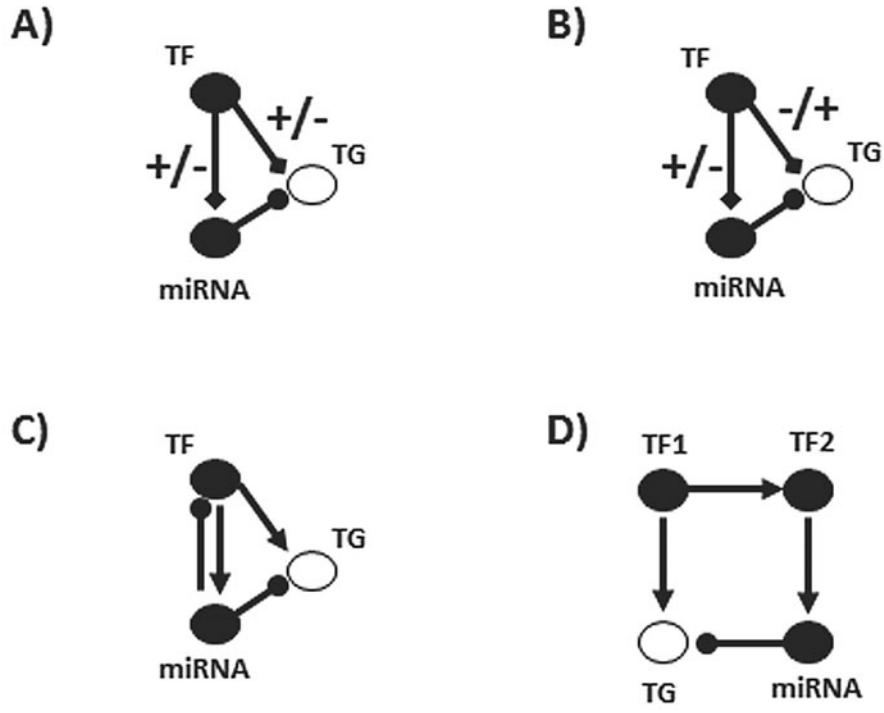

Figure 2. Different types of network motifs identified in the integrated miRNA-TF regulatory networks. Darker nodes correspond to regulators, either Transcription Factors (TFs) or miRNAs, while light colored nodes correspond to the target genes controlled by them. Oval arrows always correspond to the repression mediated by miRNAs while diamond arrows shown for TFs correspond to both activation and repression. When only an arrow is shown for a TF, it corresponds to activation mediated by the TF. A) In a type 1 Feed Forward Loop (FFL), TF sends either an activation or inhibition signal, at the same time to both miRNA and the Target Gene (TG), thereby causing an incoherent effect on target genes' expression. B) In type 2 FFL, likewise there are two possibilities, one where TF activates miRNA and inhibits TG at the same time and the other where TF inhibits miRNA and activates TG. This results in a net coherent effect on the TG. C) In superposed/composite FFL, TF activates miRNA and TG while the miRNA represses both of them. D) In an indirect feed forward loop, feedforward effect is achieved when TF1 activates a TG and TF2, which in succession activates miRNA. MicroRNA then represses TG to bring in a cascading effect.

link different modules/complexes in the protein interaction network, are more likely to be regulated by miRNAs.

EXPRESSION DYNAMICS AND CROSSTALK OF miRNAs WITH OTHER REGULATORY FACTORS

Analogous to transcriptional regulatory networks, wherein TFs regulate their target genes in a combinatorial manner,[79] most miRNAs fine tune the expression of hundreds of genes in a combinatorial manner at the post-transcriptional level.[1] This combinatorial regulation among miRNAs is believed to result in the following outcomes on their targets:

(1) several miRNAs have been found to regulate a single mRNA target by binding its transcript preferentially in the 3′ UTR region although regulation by multiple miRNAs in coding sequences has also been reported;[80,81] (2) MicroRNAs which are clustered and are often co-expressed could regulate functionally related proteins resulting in the targeting of same or functionally related pathways.[82-85] These computational and experimental studies suggest that co-ordinate regulation by miRNAs is a flexible and efficient strategy to regulate cellular processes in a condition or tissue-specific manner. However, whether the combinatorial interplay is conserved across species or there would have been a conserved set of coregulating pairs which maintain this phenomenon in specific tissues or conditions in related organisms is unclear. MicroRNAs, in addition to their interplay with other miRNAs, are also known to combinatorially control genes with the help of TFs (discussed above), RNA-binding proteins and chromatin associated factors to control important functions during development and differentiation.[86-90] Precise roles and mechanisms of most of which remain unclear to date.

In addition to the key roles played by miRNAs in the context of other regulatory proteins in the cell, they are also known to work as regulators for buffering the protein expression noise. Since miRNAs can tune a protein's expression level more rapidly by targeting the encoded transcript at the post-transcriptional level compared to transcription factors at the transcriptional level, they have been proposed to significantly shorten the response delay and in turn provide more effective noise buffering. For example, in flies miR-9a is suggested to set up a threshold for signals in a positive feedback loop so that it can filter out noise.[91] During fly sensory organ development, a fly gene, *senseless's* expression is activated by proneuronal proteins and feeds back positively to reinforce proneural gene expression. If *senseless,* the target of miR-9a is highly expressed, flies produce extra sense organs. MicroRNA miR-9a has been suggested to set a threshold that *senseless* expression has to overcome to induce the normal developmental program. These observations are strongly supported by the fact that most of the developmental genes are highly conserved targets of miRNAs, indicating the importance of preserving their regulatory circuitry across organisms with similar developmental programs. It is also evident from genome-wide studies that miRNAs might act as genetic buffers to constrain gene expression divergence between organisms, as was shown in a study comparing the expression divergences of miRNA targets across diverse organisms.[92] It is possible that the abundance of miRNAs in higher eukaryotes minimizes the expression divergence in a population by decreasing the detrimental effects of errors in gene regulation. Such a buffering phenomenon might also provide the cell with a way to silence-accumulating mutations without being subjected to extreme selective forces and hence contribute to the evolvability of an organism.

CONCLUSION

The emerging picture of miRNA regulation is getting far richer and increasingly complex than the simple linear pathways that one would have imagined these micromanagers to be controlling a decade ago. It is now becoming clear that miRNAs not only control a significant fraction of the protein coding genes in higher eukaryotes, but also function as key regulators in development, proliferation, differentiation, apoptosis, signalling and metabolism. In addition to their core roles in controlling cellular mRNA pool by managing the protein output, miRNAs have also been shown to work in mediating cross-talk between

different cellular networks by integrating various signals. Given the enormous ability these tiny RNAs have in managing the interactome of higher eukaryotes, the stage is set to manipulate their precise roles in regulating RNAs for therapeutic benefit in years to come.

ACKNOWLEDGEMENTS

SCJ acknowledges financial support from Institute for Genomic Biology, University of Illinois at Urbana-Champaign. We would also like to thank Magdy Alabady and Nitish Mittal for providing helpful comments on a previous version of this manuscript.

REFERENCES

1. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004; 116(2):281-297.
2. Carthew RW. Gene regulation by microRNAs. Curr Opin Genet Dev 2006; 16(2):203-208.
3. Carthew RW, Sontheimer EJ. Origins and Mechanisms of miRNAs and siRNAs. Cell 2009; 136(4):642-655.
4. Valencia-Sanchez MA, Liu J, Hannon GJ et al. Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev 2006; 20(5):515-524.
5. Cao X, Yeo G, Muotri AR et al. Noncoding RNAs in the mammalian central nervous system. Annu Rev Neurosci 2006; 29:77-103.
6. Plasterk RH. Micro RNAs in animal development. Cell 2006; 124(5):877-881.
7. Wienholds E, Koudijs MJ, van Eeden FJ et al. The microRNA-producing enzyme Dicer1 is essential for zebrafish development. Nat Genet 2003; 35(3):217-218.
8. Costinean S, Zanesi N, Pekarsky Y et al. Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice. Proc Natl Acad Sci U S A 2006; 103(18):7024-7029.
9. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 1993; 75(5):843-854.
10. Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell 1993; 75(5):855-862.
11. Lee RC, Ambros V. An extensive class of small RNAs in Caenorhabditis elegans. Science 2001; 294(5543):862-864.
12. Lagos-Quintana M, Rauhut R, Lendeckel W et al. Identification of novel genes coding for small expressed RNAs. Science 2001; 294(5543):853-858.
13. Baskerville S, Bartel DP. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA 2005; 11(3):241-247.
14. Rodriguez A, Griffiths-Jones S, Ashurst JL et al. Identification of mammalian microRNA host genes and transcription units. Genome Res 2004; 14(10A):1902-1910.
15. Lu J, Shen Y, Wu Q et al. The birth and death of microRNA genes in Drosophila. Nat Genet 2008; 40(3): 351-355.
16. Kim VN. MicroRNA biogenesis: coordinated cropping and dicing. Nat Rev Mol Cell Biol 2005; 6(5):376-385.
17. Denli AM, Tops BB, Plasterk RH et al. Processing of primary microRNAs by the Microprocessor complex. Nature 2004; 432(7014):231-235.
18. Gregory RI, Yan KP, Amuthan G et al. The Microprocessor complex mediates the genesis of microRNAs. Nature 2004; 432(7014):235-240.
19. Bohnsack MT, Czaplinski K, Gorlich D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA 2004; 10(2):185-191.
20. Lund E, Guttinger S, Calado A et al. Nuclear export of microRNA precursors. Science 2004; 303(5654):95-98.
21. Bernstein E, Kim SY, Carmell MA et al. Dicer is essential for mouse development. Nat Genet 2003; 35(3):215-217.
22. Hutvagner G, McLachlan J, Pasquinelli AE et al. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science 2001; 293(5531):834-838.
23. Olsen PH, Ambros V. The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. Dev Biol 1999; 216(2):671-680.
24. Humphreys DT, Westman BJ, Martin DI et al. MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. Proc Natl Acad Sci U S A 2005; 102(47):16961-16966.

25. Pillai RS, Bhattacharyya SN, Artus CG et al. Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science 2005; 309(5740):1573-1576.
26. Jing Q, Huang S, Guth S et al. Involvement of microRNA in AU-rich element-mediated mRNA instability. Cell 2005; 120(5):623-634.
27. Lim LP, Lau NC, Garrett-Engele P et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 2005; 433(7027):769-773.
28. Sood P, Krek A, Zavolan M et al. Cell-type-specific signatures of microRNAs on target mRNA expression. Proc Natl Acad Sci U S A 2006; 103(8):2746-2751.
29. Bagga S, Bracht J, Hunter S et al. Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. Cell 2005; 122(4):553-563.
30. Guo H, Ingolia NT, Weissman JS et al. Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 2010; 466(7308):835-840.
31. Lim LP, Lau NC, Weinstein EG et al. The microRNAs of Caenorhabditis elegans. Genes Dev 2003; 17(8): 991-1008.
32. Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 2011; 39(Database issue):D152-D157.
33. Lee CT, Risom T, Strauss WM. Evolutionary conservation of microRNA regulatory circuits: an examination of microRNA gene complexity and conserved microRNA-target interactions through metazoan phylogeny. DNA Cell Biol 2007; 26(4):209-218.
34. Heimberg AM, Sempere LF, Moy VN et al. MicroRNAs and the advent of vertebrate morphological complexity. Proc Natl Acad Sci U S A 2008; 105(8):2946-2950.
35. Prochnik SE, Rokhsar DS, Aboobaker AA. Evidence for a microRNA expansion in the bilaterian ancestor. Dev Genes Evol 2007; 217(1):73-77.
36. Bentwich I, Avniel A, Karov Y et al. Identification of hundreds of conserved and nonconserved human microRNAs. Nat Genet 2005; 37(7):766-770.
37. Berezikov E, Thuemmler F, van Laake LW et al. Diversity of microRNAs in human and chimpanzee brain. Nat Genet 2006; 38(12):1375-1377.
38. Borchert GM, Lanier W, Davidson BL. RNA polymerase III transcribes human microRNAs. Nat Struct Mol Biol 2006; 13(12):1097-1101.
39. Tanzer A, Amemiya CT, Kim CB et al. Evolution of microRNAs located within Hox gene clusters. J Exp Zool B Mol Dev Evol 2005; 304(1):75-85.
40. Tanzer A, Stadler PF. Evolution of microRNAs. Methods Mol Biol 2006; 342:335-350.
41. Fahlgren N, Howell MD, Kasschau KD et al. High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. PLoS One 2007; 2(2):e219.
42. Li A, Mao L. Evolution of plant microRNA gene families. Cell Res 2007; 17(3):212-218.
43. Lindow M, Krogh A. Computational evidence for hundreds of nonconserved plant microRNAs. BMC Genomics 2005; 6:119.
44. Barakat A, Wall PK, Diloreto S et al. Conservation and divergence of microRNAs in Populus. BMC Genomics 2007; 8:481.
45. Fattash I, Voss B, Reski R et al. Evidence for the rapid expansion of microRNA-mediated regulation in early land plant evolution. BMC Plant Biol 2007; 7:13.
46. Berezikov E, Guryev V, van de Belt J et al. Phylogenetic shadowing and computational identification of human microRNA genes. Cell 2005; 120(1):21-24.
47. Ohler U, Yekta S, Lim LP et al. Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification. RNA 2004; 10(9):1309-1322.
48. Altuvia Y, Landgraf P, Lithwick G et al. Clustering and conservation patterns of human microRNAs. Nucleic Acids Res 2005; 33(8):2697-2706.
49. Sewer A, Paul N, Landgraf P et al. Identification of clustered microRNAs using an ab initio prediction method. BMC Bioinformatics 2005; 6:267.
50. Xue C, Li F, He T et al. Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine. BMC Bioinformatics 2005; 6:310.
51. Jiang P, Wu H, Wang W et al. MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. Nucleic Acids Res 2007; 35(Web Server issue):W339-W344.
52. Ding J, Zhou S, Guan J. MiRenSVM: towards better prediction of microRNA precursors using an ensemble SVM classifier with multi-loop features. BMC Bioinformatics 2010; 11 Suppl 11:S11.
53. Friedlander MR, Chen W, Adamidi C et al. Discovering microRNAs from deep sequencing data using miRDeep. Nat Biotechnol 2008; 26(4):407-415.
54. Hackenberg M, Sturm M, Langenberger D et al. miRanalyzer: a microRNA detection and analysis tool for next-generation sequencing experiments. Nucleic Acids Res 2009; 37(Web Server issue):W68-W76.
55. Yang JH, Li JH, Shao P et al. starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. Nucleic Acids Res 2011; 39(Database issue):D202-D209.

56. Yang JH, Shao P, Zhou H et al. deepBase: a database for deeply annotating and mining deep sequencing data. Nucleic Acids Res 2010; 38(Database issue):D123-130.
57. Papadopoulos GL, Reczko M, Simossis VA et al. The database of experimentally supported targets: a functional update of TarBase. Nucleic Acids Res 2009; 37(Database issue):D155-158.
58. Rhoades MW, Reinhart BJ, Lim LP et al. Prediction of plant microRNA targets. Cell 2002; 110(4):513-520.
59. Davis E, Caiment F, Tordoir X et al. RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. Curr Biol 2005; 15(8):743-749.
60. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell 2009; 136(2):215-233.
61. Baek D, Villen J, Shin C et al. The impact of microRNAs on protein output. Nature 2008; 455(7209):64-71.
62. Selbach M, Schwanhausser B, Thierfelder N et al. Widespread changes in protein synthesis induced by microRNAs. Nature 2008; 455(7209):58-63.
63. Brennecke J, Stark A, Russell RB et al. Principles of microRNA-target recognition. PLoS Biol 2005; 3(3):e85.
64. Krek A, Grun D, Poy MN et al. Combinatorial microRNA target predictions. Nat Genet 2005; 37(5):495-500.
65. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 2005; 120(1):15-20.
66. Xie X, Lu J, Kulbokas EJ et al. Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. Nature 2005; 434(7031):338-345.
67. Cui Q, Yu Z, Purisima EO et al. Principles of microRNA regulation of a human cellular signaling network. Mol Syst Biol 2006; 2:46.
68. Ma'ayan A, Jenkins SL, Neves S et al. Formation of regulatory patterns during signal propagation in a Mammalian cellular network. Science 2005; 309(5737):1078-1083.
69. Lee TI, Jenner RG, Boyer LA et al. Control of developmental regulators by Polycomb in human embryonic stem cells. Cell 2006; 125(2):301-313.
70. Cui Q, Yu Z, Pan Y et al. MicroRNAs preferentially target the genes with high transcriptional regulation complexity. Biochem Biophys Res Commun 2007; 352(3):733-738.
71. Shen-Orr SS, Milo R, Mangan S et al. Network motifs in the transcriptional regulation network of Escherichia coli. Nat Genet 2002; 31(1):64-68.
72. Alon U. Network motifs: theory and experimental approaches. Nat Rev Genet 2007; 8(6):450-461.
73. Janga SC, Collado-Vides J. Structure and evolution of gene regulatory networks in microbial genomes. Res Microbiol 2007; 158(10):787-794.
74. Shalgi R, Lieber D, Oren M et al. Global and local architecture of the mammalian microRNA-transcription factor regulatory network. PLoS Comput Biol 2007; 3(7):e131.
75. Zhou Y, Ferguson J, Chang JT et al. Inter- and intra-combinatorial regulation by transcription factors and microRNAs. BMC Genomics 2007; 8:396.
76. Wang J, Lu M, Qiu C et al. TransmiR: a transcription factor-microRNA regulation database. Nucleic Acids Res 2010; 38(Database issue):D119-D122.
77. O'Donnell KA, Wentzel EA, Zeller KI et al. c-Myc-regulated microRNAs modulate E2F1 expression. Nature 2005; 435(7043):839-843.
78. Liang H, Li WH. MicroRNA regulation of human protein protein interaction network. RNA 2007; 13(9):1402-1408.
79. Balaji S, Babu MM, Iyer LM et al. Comprehensive analysis of combinatorial regulation using the transcriptional regulatory network of yeast. J Mol Biol 2006; 360(1):213-227.
80. Fontana L, Pelosi E, Greco P et al. MicroRNAs 17-5p-20a-106a control monocytopoiesis through AML1 targeting and M-CSF receptor upregulation. Nat Cell Biol 2007; 9(7):775-787.
81. Tay Y, Zhang J, Thomson AM et al. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. Nature 2008; 455(7216):1124-1128.
82. Dews M, Homayouni A, Yu D et al. Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster. Nat Genet 2006; 38(9):1060-1065.
83. Ivanovska I, Cleary MA. Combinatorial microRNAs: working together to make a difference. Cell Cycle 2008; 7(20):3137-3142.
84. Kim YK, Yu J, Han TS et al. Functional links between clustered microRNAs: suppression of cell-cycle inhibitors by microRNA clusters in gastric cancer. Nucleic Acids Res 2009; 37(5):1672-1681.
85. Yuan X, Liu C, Yang P et al. Clustered microRNAs' coordination in regulating protein-protein interaction network. BMC Syst Biol 2009; 3:65.
86. Hafner M, Landthaler M, Burger L et al. PAR-CliP—a method to identify transcriptome-wide the binding sites of RNA binding proteins. J Vis Exp 2010(41).
87. Khalil AM, Guttman M, Huarte M et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A 2009; 106(28):11667-11672.
88. Alvarez-Saavedra M, Antoun G, Yanagiya A et al. miRNA-132 orchestrates chromatin remodeling and translational control of the circadian clock. Hum Mol Genet 2010.

89. van Wolfswinkel JC, Ketting RF. The role of small non-coding RNAs in genome stability and chromatin organization. J Cell Sci 2010; 123(Pt 11):1825-1839.
90. Laporte P, Merchan F, Amor BB et al. Riboregulators in plant development. Biochem Soc Trans 2007; 35(Pt 6):1638-1642.
91. Li Y, Wang F, Lee JA et al. MicroRNA-9a ensures the precise specification of sensory organ precursors in Drosophila. Genes Dev 2006; 20(20):2793-2805.
92. Cui Q, Yu Z, Purisima EO et al. MicroRNA regulation and interspecific variation of gene expression. Trends Genet 2007; 23(8):372-375.
93. Ng KL, Mishra SK. De novo SVM classification of precursor microRNAs from genomic pseudo hairpins using global and intrinsic folding measures. Bioinformatics 2007; 23(11):1321-1330.
94. Hertel J, Stadler PF. Hairpins in a Haystack: recognizing microRNA precursors in comparative genomics data. Bioinformatics 2006; 22(14):e197-e202.
95. Helvik SA, Snove O Jr., Saetrom P. Reliable prediction of Drosha processing sites improves microRNA gene prediction. Bioinformatics 2007; 23(2):142-149.
96. Yousef M, Jung S, Kossenkov AV et al. Naive Bayes for microRNA target predictions—machine learning for microRNA targets. Bioinformatics 2007; 23(22):2987-2992.
97. Huang JC, Babak T, Corson TW et al. Using expression profiling data to identify human microRNA targets. Nat Methods 2007; 4(12):1045-1049.
98. Enright AJ, John B, Gaul U et al. MicroRNA targets in Drosophila. Genome Biol 2003; 5(1):R1.
99. Kiriakidou M, Nelson PT, Kouranov A et al. A combined computational-experimental approach predicts human microRNA targets. Genes Dev 2004; 18(10):1165-1178.
100. Rehmsmeier M, Steffen P, Hochsmann M et al. Fast and effective prediction of microRNA/target duplexes. RNA 2004; 10(10):1507-1517.
101. Grun D, Wang YL, Langenberger D et al. microRNA target predictions across seven Drosophila species and comparison to mammalian targets. PLoS Comput Biol 2005; 1(1):e13.
102. Miranda KC, Huynh T, Tay Y et al. A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell 2006; 126(6):1203-1217.
103. Gaidatzis D, van Nimwegen E, Hausser J et al. Inference of miRNA targets using evolutionary conservation and pathway analysis. BMC Bioinformatics 2007; 8:69.
104. Kertesz M, Iovino N, Unnerstall U et al. The role of site accessibility in microRNA target recognition. Nat Genet 2007; 39(10):1278-1284.
105. Saetrom O, Snove O Jr., Saetrom P. Weighted sequence motifs as an improved seeding step in microRNA target prediction algorithms. RNA 2005; 11(7):995-1003.
106. Hammell M, Long D, Zhang L et al. mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein-enriched transcripts. Nat Methods 2008; 5(9):813-819.