Chapter 7

Computational Prediction of Functional MicroRNA–mRNA Interactions

Müşerref Duygu Saçar Demirci, Malik Yousef, and Jens Allmer

Abstract

Proteins have a strong influence on the phenotype and their aberrant expression leads to diseases. MicroRNAs (miRNAs) are short RNA sequences which posttranscriptionally regulate protein expression. This regulation is driven by miRNAs acting as recognition sequences for their target mRNAs within a larger regulatory machinery. A miRNA can have many target mRNAs and an mRNA can be targeted by many miRNAs which makes it difficult to experimentally discover all miRNA–mRNA interactions. Therefore, computational methods have been developed for miRNA detection and miRNA target prediction. An abundance of available computational tools makes selection difficult. Additionally, interactions are not currently the focus of investigation although they more accurately define the regulation than pre-miRNA detection or target prediction could perform alone. We define an interaction including the miRNA source and the mRNA target. We present computational methods allowing the investigation of these interactions as well as how they can be used to extend regulatory pathways. Finally, we present a list of points that should be taken into account when investigating miRNA–mRNA interactions. In the future, this may lead to better understanding of functional interactions which may pave the way for disease marker discovery and design of miRNA-based drugs.

Key words MicroRNA, Target, Regulation, Posttranscriptional regulation, Pathway extension, MiRNA–mRNA interaction

1 Introduction

The central dogma in biology describes a programmed flow of information from the genome to the phenotype [1]. However, as many dogmas before (e.g., the earth being the center of the universe), this viewpoint is challenged and a better explanation for reality can be gained by embracing relativity theory in biology [2]. This highlights the importance of regulation for the overall genetic programming which consists of many scopes including epigenetic, transcriptomic, and proteomic. Life, manifested at all scales, from metabolites to complete organisms, interplays and thereby implements the genetic program with the genome providing a parts list [2]. Posttranscriptional regulation via

Xin Lai et al. (eds.), *Computational Biology of Non-Coding RNA: Methods and Protocols*, Methods in Molecular Biology, vol. 1912, https://doi.org/10.1007/978-1-4939-8982-9_7,

noncoding RNAs modifies protein abundance. Proteins can be transcription factors or parts of biological machines and pathways with direct involvement in transcription, closing the regulatory loop. Currently, evidence is accumulating supporting transcriptionally active miRNAs which are reimported into the nucleus [3], providing a direct involvement of microRNAs (miRNAs) in transcriptional control. Mature miRNAs are small (18–24 nt) noncoding RNAs derived from hairpins involving a complex molecular pathway [4]. These mature miRNAs act as recognition keys for their target genes within silencing complexes such as RISC. Either targeted mRNAs are silenced through endonucleolytic cleavage [5] or the translation process is modulated by either increasing [6] or decreasing protein abundance.

Regulation of gene expression is important and dysregulation often leads to disease. MiRNA dysregulation has been implicated in cancer, amyotrophic lateral sclerosis, and many other diseases. MiRNAs are also potentially instrumentalized by viruses [7, 8] and other organisms such as *Toxoplasma gondii* [9, 10] to modulate the host environment to their advantage. This may also work vice versa and there are accounts of human miRNAs shaping the gut microbiome [11] but in general the miRNA-based communication between host and microbiome is not yet well understood [12]. MiRNAs can be detected experimentally using a number of approaches such as degradome sequencing (Table 1).

Such approaches are time consuming and cannot investigate all possible states of a biological system which is why top-down computational ab initio prediction of miRNAs is important [16]. The same holds true for miRNA targets which can be experimentally detected using, for example, HITS-CLIP [17]. The multiplicity of miRNA targets further increases the challenge so that computational methods are prerogative [18]. MiRNAs and their target mRNAs can form a regulative network based on their source transcripts. It is, therefore, important to combine transcriptional and posttranscriptional regulation when investigating miRNAs and their targets.

The further text is structured into three parts. First we discuss miRNA detection including experimental and computational methods. Information on miRNA functionality follows with a focus on targeting. Finally, regulatory networks resulting from miRNAs and their targets are introduced.

2 MicroRNA Detection

2.1 Experimental miRNA Detection

Since the first discovery of miRNAs, many experimental and computational approaches have been developed for their detection and analysis. In earlier studies, such as the identification of let-7 in *Caenorhabditis elegans*, forward genetic approaches were applied

Table 1
Comparison of miRNA and miRNA target detection strategies

Method	Advantage	Disadvantage
Northern Blotting	New and known miRNAs Validation	Need lots of RNA MiRNAs with low abundance Low throughput Low sensitivity Time consuming
qRT-PCR	Highest dynamic range and accuracy Absolute quantification	Throughput problems Normalization and specificity
Microarray	Cheaper Screening tool	Not quantitative Lower sensitivity and dynamic range
NGS	High sensitivity Detection of sequence variation New miRNAs	Artifacts/contamination Time consuming Complex bioinformatics
HITS-CLIP	Stringent isolation of the miRNA-mRNA-Ago Analyzing miRNA–mRNA interactome	Inefficient UV cross-linking Requirement of large number of cells for library preparation
CLASH	High-throughput identification Independent of bioinformatic predictions	Stringent purification Low efficiency of RNA-RNA ligation

Adapted and updated from [13–15]

[19]. However, such methods are limited due to some characteristics of miRNAs, for example, their small size [20]. Today, Northern blotting, microarray, NGS, and qRT-PCR are popular experimental techniques used for detection and/or validation of miRNAs, each of these methods having some inherent strengths and weaknesses (Table 1).

Experiments will remain the golden standard for confirmation of miRNAs, but it has become clear that it is not possible to capture all miRNAs experimentally since they may only be expressed in low quantity, at specific developmental times, or in specific tissues. Therefore, computational detection of miRNAs has become important.

2.2 Computational miRNA Detection

Various tools have been designed for the in silico prediction of miRNAs (Table 2). Although they may vary in numerous parts, most of these programs rely on the secondary structure of the miRNA precursor [45]. Algorithms such as RNAfold [46] not only perform RNA secondary structure prediction but also calculate the thermodynamic stability of the proposed RNA hairpin structures. Available methods generally include (1) genome-wide

Table 2
Selected computational tools for miRNA prediction

Tool	Year	Conservation	Structure	Sequence	Machine learning	NGS application
miRscan [21]	2003	+	+			
miRAlign [22]	2005		+	+		
ProMiR [23]	2005	+	+	+	+	
Triplet-SVM [24]	2005		+	+	+	
miR-abela [25]	2005		+	+	+	
RNAmicro [26]	2006		+	+	+	
BayesMiRNAfind [27]	2006	+	+	+		
miRFinder [28]	2007	+	+		+	
miPred [29]	2007		+		+	
MiRRim [30]	2007	+	+		+	
miRDeep [31]	2008		+			+
miRanalyzer [32]	2009				+	+
SSCprofiler [33]	2009	+	+	+	+	+
HHMMiR [34]	2009		+	+	+	
MIReNA [35]	2010	+	+	+		+
miRPara [36]	2011		+	+	+	
miRNAFold [37]	2012		+	+		
miREval 2.0 [38]	2013	+	+	+	+	
miR-PREFeR [39]	2014		+			+
miRBoost [40]	2015		+	+	+	
iMiRNA-SSF [41]	2016		+	+	+	
izMiR [42]	2017		+	+	+	+
miRge2.0 [43]	2018		+	+	+	+

The list is updated version from [44]. The list is sorted by publication year

prediction of hairpin structures, (2) filtering and/or scoring of those hairpins, and (3) experimental confirmation [45].

MiRNA gene prediction algorithms can be divided into several categories. Usually, either homology modeling or ab initio methods are applied to extract possible miRNAs from a genome [16]. Homology modeling is based on the idea that if a miRNA is identified in one organism then it is possible that its homologs

might be found in closely related species. Considering miRNA families, using homology-based approach in the same organism can be beneficial too. Since conservation is usually associated with function, most of the predictions made by tools using homology-based methods tend to be miRNAs. It has been shown that a miRNA prediction approach developed by using available genome sequences in castor bean (*Ricinus communis*) was able to detect 86.6% of miRNAs in Arabidopsis [47]. However, it is important to note that lack of conservation in miRNAs does not imply absence of function and that fast evolution of some miRNAs has been observed [48]. Although homology-based methods are quite helpful for initial screening of candidate miRNAs, they have several essential drawbacks. Most importantly, it is not possible to find novel and/or species-specific miRNAs. Therefore, various ab initio-based tools are developed for prediction of novel miRNAs.

Even though the ab initio approach does not directly rely on conservation, it still uses information obtained from known miRNAs. The majority of such approaches utilizes machine learning (ML) (Table 2) and specifically two-class classification. There are many factors that affect the performance of the prediction scheme but the most important one seems to be data quality [49]. Various tools have been designed for miRNA prediction. Some of them are constructed in a similar manner while others are using different approaches to achieve the task (Table 2). When comparing available tools and selecting one of them for analysis, the most important criteria should be accuracy of their results. Although some measurements like accuracy, sensitivity, and specificity are usually reported in papers about such tools, it is not possible to compare such values directly [42]. Based on a comparison of 15 tools in terms of sensitivity (min: 55, max: 98) and specificity (min: 40, max: 98), none of them provides sufficient confidence for experimental testing of all estimated 60 million miRNA-like structures found in the human genome [44]. Therefore, we recently suggested using consensus approaches for more reliable predictions rather than depending on any one tool [42].

Computational detection of miRNAs even in large eukaryotic genomes is now possible which can be used to answer questions such as about their locations within a genome and their multiplicity. Some miRNAs have large number of copies spread throughout the genome as for example for let-7. MiRNA copy number variation may lead to disease [50] which adds value to answering such questions. MiRNAs may vary in respect to their abundance with which they are encoded in a genome but their number of targets varies even more widely.

3 MicroRNA Targeting

MiRNAs follow a biogenesis pathway leading from their transcription to their incorporation into RISC [51]. Within RISC they help recognize their targets via sequence complementarity. The consensus is that RISC binding to their target mRNAs primes them for degradation or for translational repression when binding within the 3′-UTR region of the target mRNA. There have been accounts of translational activation when binding in the 5′-UTR and reimport of mature miRNAs into the nucleus leading to transcriptional control but this is not discussed here. Experimental approaches like PAR-CLIP [52], HITS-CLIP [53], and CLASH [54] are currently employed for the investigation of miRNA-based targeting (Table 1). Such techniques are limited to the availability of bound miRNA–mRNA in large enough quantity for detection. Competition among miRNA with multiple targets, miRNAs, or targets only expressed under specific conditions, and miRNAs or targets only expressed in low quantities, adds to this problem. MiRNA targets are stored in a number of databases such as miRTarBase [55] and TarBase [56]. Some databases contain computational predictions and others focus on experimentally validated targets (Table 3).

The experimental evidence points to the complementary binding of miRNA-loaded RISC to its targets as the most important factor for miRNA function. This is important for the design of computational methods for target prediction (Fig. 1). Investigation of the binding potential of miRNAs leads to a separation of the mature miRNA into seed region (nucleotides (1/2)-8 at 5′) and out region (3′ portion of the miRNA). The seed region generally forms perfect Watson-Crick complementarity with the target mRNA.

Computational approaches reflect the importance of complementary binding and heavily rely on it for detection of miRNA targets [18]. There are many means to classify computational tools for miRNA prediction and a recent survey grouped them into Web-based services, downloaded software, and R packages [73]. They conclude that Web-based tools are the most frequently used platform to predict miRNA–mRNA interactions and the top three tools from this category are TargetScan [74], miRanda [75], and DIANA Tools [76]. There has not been an independent assessment of prediction accuracy of the multitude of miRNA target prediction tools. Therefore, it cannot be judged whether preferential usage of these tools over standalone platforms and R scripts presents a problem. According to Riffo-Campos et al. TargetScan seems to be the most robust tool since its predictions have a higher probability of being biologically validated due to usage statistics [73].

Table 3
Databases containing miRNA targets and software for the prediction of miRNA targets

Name	Year	Type	Link
RNAhybrid [57]	2004	Webserver, predictions	https://bibiserv.cebitec.uni-bielefeld.de/rnahybrid;jsessionid=791293b42354681cb4afa2201b63
PicTar [58]	2005	Database, webserver, predictions	https://web.archive.org/web/20080724163022/http://pictar.bio.nyu.edu/
TargetScan [59]	2005	Database, webserver	http://www.targetscan.org/vert_72/
RNA22 [60]	2006	Webserver, predictions	https://cm.jefferson.edu/rna22/
TarBase [61]	2006	Database	http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=tarbase/index
NBmiRtar [62]	2007	Webserver, predictions	http://wotan.wistar.upenn.edu/NBmiRTar/
PITA [63]	2007	Webserver, predictions	https://genie.weizmann.ac.il/pubs/mir07/mir07_data.html
Diana-microT [64]	2009	Webserver	http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=tarbase/index
miRecords [65]	2009	Database	http://c1.accurascience.com/miRecords/
miRTarBase [66]	2011	Database	http://mirtarbase.mbc.nctu.edu.tw/php/index.php
miRwalk [67]	2011	Database, webserver	http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk2/
RepTar [68]	2011	Database	http://bioinformatics.ekmd.huji.ac.il/reptar/
StarBase [69]	2014	Database	https://web.archive.org/web/20110222111721/http://starbase.sysu.edu.cn/
Cupid [70]	2015	Matlab script	http://cupidtool.sourceforge.net/
MBSTAR [71]	2015	Webserver, predictions	https://www.isical.ac.in/~bioinfo_miu/MBStar30.htm
StarScan [72]	2015	Web-based software	http://bioinformatics.psb.ugent.be/webtools/startscan/

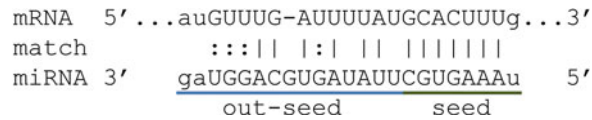

Fig. 1 Example duplex structure of a miRNA and its target mRNA. Seed sequence (green) and out seed (blue) are indicated

Unfortunately, the molecular biology of target binding is not fully understood. Therefore, machine learning tools are used to automatically learn them from known examples. To perform machine learning, parameterization of the miRNA–mRNA duplex (Fig. 1) is an important step. Peterson et al. reviewed the features used by different computational tools for the prediction of miRNA targets [77]. They found that the four main aspects of the miRNA–mRNA target interaction modeled in the tools reviewed are seed match, evolutionary conservation, free energy, and site accessibility. We recently investigated feature selection for miRNA target prediction using machine learning and found large differences depending on the parameters used [78]. Among the first computational tools for miRNA target prediction, Diana-microT determined interaction rules using bioinformatics with coupled experimental validation and was able to predict all known *C. elegans* miRNA targets. Many more tools have been developed subsequently (Table 3). Another aspect of target site detection involves the fast and accurate detection of approximately complementary matches. The miRanda algorithm for example employs dynamic programming to optimally align miRNAs with their targets [79] but other approaches employ BLAST [80] or use different heuristics for sequence alignment. RNAhybrid further includes hybridization energy in its search for target sites [81]. Lai observed very little overlap among the predicted targets identified by several miRNA target prediction tools [82]. This could be explained by the utilization of different feature sets which capture distinct target sites. Sethupathy and colleagues [61] also compared miRNA target prediction tools and found that about 30% of experimentally validated target sites are non-conserved which can also partially explain the difference among tools found by Lai [82]. A large part of features used model sequence conservation. Such features may not generalize well since we were able to differentiate among target sites from different species using sequence-based features [83]. Yousef et al. developed a target-prediction method NBmiRTar using machine learning with a naïve Bayes classifier which does not incorporate sequence conservation but generates a model from sequence and miRNA–mRNA duplex information [62]. Training and testing examples were derived from validated target sequences and artificially generated negative data. NBmiRTar incorporates information from the seed and the "out-seed" segments of the miRNA–mRNA duplex (Fig. 1) and thereby produces fewer false-positive predictions and fewer target candidates to be tested than other methods.

Most recently, Riffo-Campos et al. reviewed miRNA target tools and described the fundamental biology which these prediction tools are based on. They also characterized the main sequence-based algorithms, and offered some insights into their uses by biologists [73]. Following miRNA target prediction, biomolecular

validation is always necessary to confirm the miRNA-target gene interaction. Thus a protocol for validation of a miRNA target interaction is required. Different approaches are used for validation such as cloning of a dual-luciferase miRNA target expression vector, transfection of cells with this vector and a precursor miRNA (pre-miRNA), and subsequent luciferase assay [84]. Many other approaches such as degradome sequencing and methods in Table 1 are also used for validation. However the limitations of each validation approach need to be well understood. These limitations are well summarized in a recent review of experimental techniques for miRNA target identification [85]. Among the databases also giving supporting experimental evidence, miRTarBase [86] is the most updated resource for experimentally validated microRNA-target interactions.

While many miRNA targeting tools and several target databases are available, there still is a need to improve upon miRNA-targeting prediction in respect to prediction accuracy and toward quantifying the effect of the regulation. In the future, it would be beneficial for target prediction to include measures for target-site binding strength, target-site multiplicity on the target mRNA, and proximity to the stop of translation.

4 Detection of Gene: miRNA–Gene Interactions

MicroRNA detection and target prediction are important tasks. However, miRNAs can only convey function when co-expressed with their targets [87]. Therefore, it is important to refer to miRNA–mRNA interactions instead of analyzing miRNAs and their targets independently. Experimentally, such interactions can be analyzed using HITS-CLIP by cross-linking bound RNAs within protein complexes, isolating them, and sequencing the associated RNAs [53]. HITS-CLIP data was also instrumental in developing computational methods for the analysis of miRNA–mRNA regulatory interactions [88]. The HITS-CLIP methodology allows for the identification of functional interactions. Some limitations have been overcome with the CLASH protocol [54]. These approaches cannot replace computational methods, though, because it is not feasible to perform such experiments for all species, developmental stages, tissues, and external and internal stresses which affect regulation. Therefore, computational predictions of hairpins and their targets is an active area of research. A focus on miRNA–mRNA interactions will further shape the field in the future.

We here define a miRNA–mRNA interaction in terms of source gene interacting with its target gene, thereby abstracting all the detailed biological pathways and focusing only on the regulative role of the interaction. It needs to be noted that miRNAs can originate from anywhere in a genome. About 50% of the miRNAs

are located within transcription units (40% within introns) and the other 50% are located in intergenic regions where they are usually clustered [89]. For example, miRNAs can form clusters in intergenic regions of a genome and be transcribed in a coordinated manner [90]. They are also co-transcribed with genes within their exons or noncoding parts. For example, the DiGeorge syndrome critical region gene 8 (DGCR8) contains a hairpin in its first exon which can be recognized and cleaved by the microprocessor complex leading to a truncated product and a miRNA [91]. We predicted more than 300 putative targets of the hairpin in human and the following six target genes seem to be affected most with two target-binding sites each: RASGRP1, LYNX1, TBC1D16, KLHL28, IPO8, and DPF1. This also exemplifies that miRNAs can have multiple targets. On the other hand, mRNAs can be targeted by multiple miRNAs and can have multiple binding sites per miRNA. Together, this can lead to a huge interaction network (Fig. 2).

It is beneficial to filter this large network by actually expressed interactions [92]. A complication exists for the acquisition of sequencing data for miRNAs, which requires special sequencing strategies since they are very short. Such expression data is often not available in public data from, for example, the sequence read archive [93]. Another approach, which is not dependent on short read sequences, uses the expression of the enclosing transcription unit as a measure for the expression of the miRNA. MiRNAs are often part of genes (Fig. 3) and are co-transcribed with them. Therefore, the expression of the gene can indicate the expression of miRNAs co-expressed within UTRs or introns. MiRNAs from exons can be treated in the same manner, but then it also needs to be taken into account that the resulting mRNA is also structurally affected and its sequence altered.

4.1 Databases Containing MicroRNA Targets

MiRNA and their targets are available in databases such as TarBase and miRTarBase. These are based on predictions or text mining. These databases also provide information about the evidence for the targeting such as sequencing support. Degradome sequencing has become a means of assessing functional miRNA–mRNA interactions and plant data is available in DPMIND [94]. Most of the target databases represent miRNA targets for the miRNAs available in miRBase. However, in most cases the source gene or miRNA cluster is not specified. Within the page source miRBase displays the overlapping transcripts with the miRNAs it hosts but does not visualize it in the web site, yet. Perhaps, this is a future feature which still needs further scrutiny. Clearly, it is an important piece of information facilitating the integration of miRNA and gene regulatory networks. Previously, we have constructed a database compiling the miRNA and target information from TarBase, miRTarBase, and miRBase [95] which also includes the

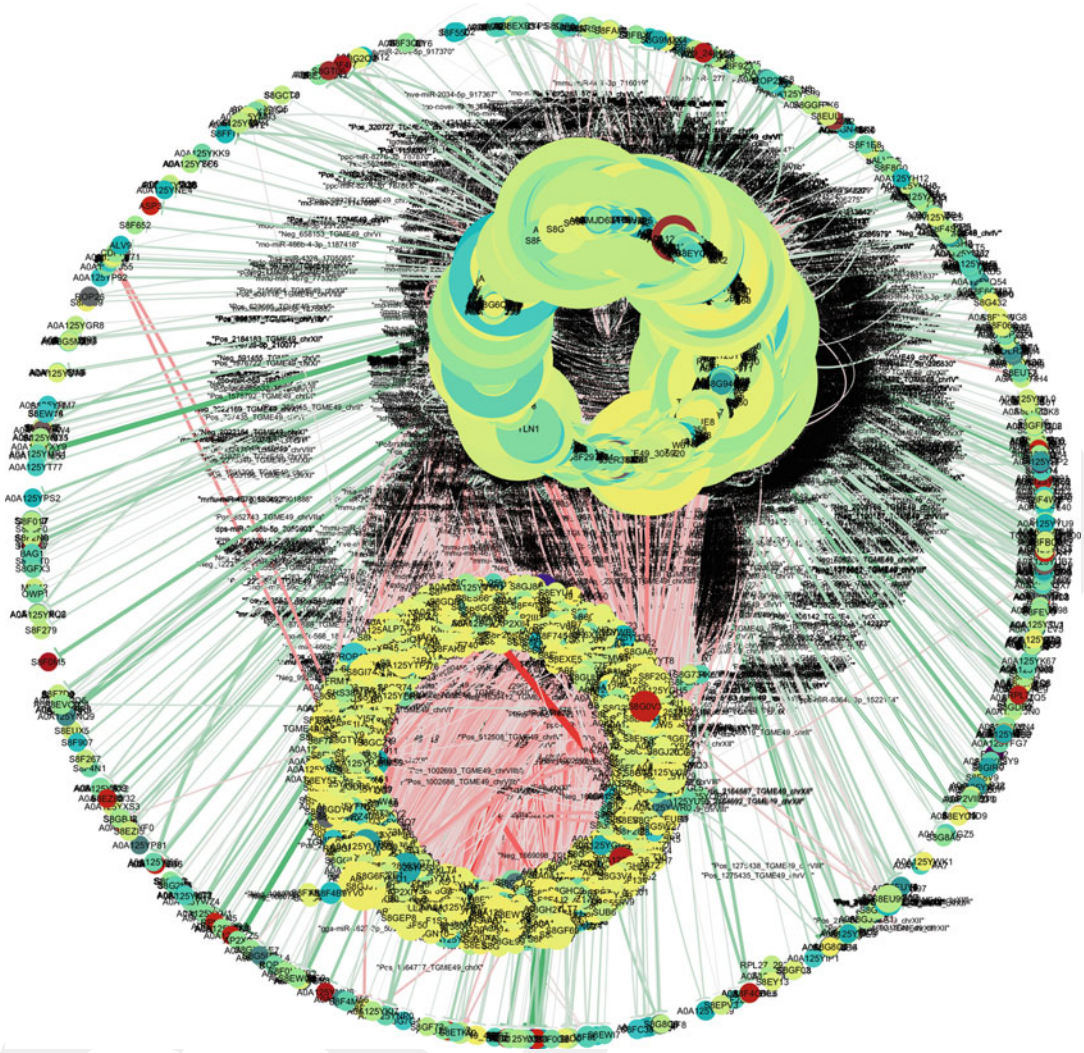

Fig. 2 Computationally predicted miRNA–mRNA interactions for *Toxoplasma gondii* forming a hairball that is difficult to interpret [92]

overlapping transcripts and can be queried using VANESA [96]. Figure 3 shows a regulatory network including genes, miRNAs, and proteins. Generally, pathway databases like Reactome [97] and KEGG [98] only include genes and their interactions. Implicitly, the interactions represent gene products which include transcription factors. VANESA facilitates the merging and enriching of KEGG pathways with miRNA interactions form the integrated database [99]. Thereby, pathways can be extended and can be investigated on multiple regulatory levels at the same time.

4.2 Other Approaches to Network Construction

The approach detailed above includes the origin of miRNAs in network construction in addition to targeting knowledge. To the best of our knowledge, no other approach takes this into account.

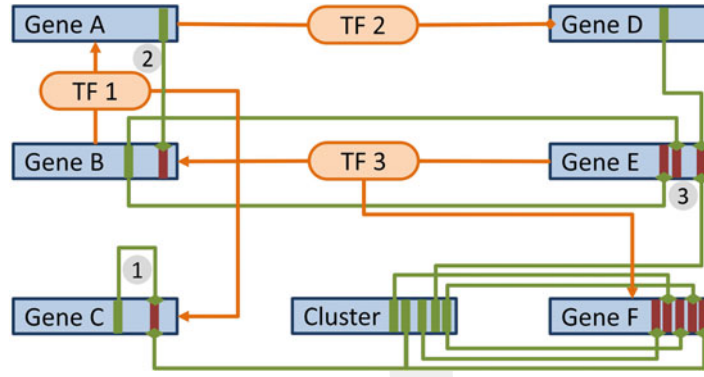

Fig. 3 A contrived example of a small regulatory network consisting of transcription factors (orange ovals), genes (blue rectangles), and miRNAs (green rectangles). Interactions are shown as lines with different arrowheads activating (arrows) or deactivating (diamonds). TFs are explicitly modeled here to show that three levels are cooperating to achieve regulation (genome, transcriptome, and proteome)

In general, networks are built from miRNAs (nodes) and their functional similarity (edges) to predict or extend disease networks. Le et al. refer to them as homogeneous networks and propose the use of bipartite graphs consisting of miRNAs and genes as nodes and their interactions as edges [100].

Regulatory networks that were established as pathways or discovered for various diseases or extended in ways as described above have also been employed for relating miRNAs to diseases. For this one of the two assumptions are generally used: (1) miRNAs that are associated with similar diseases must be similar, and (2) functionally similar miRNAs likely lead to a similar disease phenotype [101]. These assumptions lead to similarity measures or machine learning methods which suggest whether a miRNA is implicated in a particular pathway or disease [101].

In respect to (1) there are various methods to describe functional similarity. For instance, the shared targets among miRNAs can be used as a similarity measure [102]. Similarity for target gene regulation patterns has also been used [103]. Gene ontology enrichment of their targets was used to define similarity among miRNAs [104] and was further extended including protein interaction networks [105].

Considering (2), many machine learning approaches have been established. Different classifiers such as naïve Bayes [102] and support vector machines [103] were used. To remove the dependency on negative data without quality guarantee, a semi-supervised classifier was also used to prioritize candidate disease-related miRNAs [106].

Text mining of scientific publications to establish miRNA–disease associations has been employed to construct miRNA–disease

networks. Kandhro et al. extracted miRNA–lipid disease associations from literature and extended the resulting interaction network with further miRNA target predictions [107]. Similarly, Honardoost et al. used literature mining combined with database (miRWalk and miRTarBase) extension for the investigation of autoimmune disease-related miRNAs deregulated in Th17 cells [108].

Databases containing miRNA interactions such as miRWalk [67] have been used to construct miRNA regulatory networks using methodologies based in social network analysis to decipher miRNA involvement in the regulation of intestinal epithelial cellular pathways [109].

4.3 Regulatory Networks

Regulation is of crucial importance for the survival of the organism. With thousands of genes and miRNAs and even more proteins and metabolites that can interact on the molecular level, a large network of interactions results. It is currently not feasible to construct a comprehensive network let alone analyze it and, therefore, the research focus is on smaller subnetworks (Fig. 4). An example for regulation involving few partners is the microprocessor self-regulation mentioned above: DGCR8 contains a hairpin in its first exon which leads to a truncated protein product when the hairpin is excised. The microprocessor complex posttranscriptionally regulates its own expression by cleaving the hairpin [111]. This presents a feedback loop where Drosha deactivates the microprocessor in a dose-dependent manner. Very short feedback loops are possible for example Fig. 3(1) where a miRNA co-expressed with gene C inhibits the translation of the gene product. Other examples not involving miRNAs are the myocyte-enhancing factor 2 and twist genes in Drosophila which are single-gene feedback loops

Fig. 4 Recreated after [110]. A feed-forward loop (FFL) regulates PA and PtdIns (4,5)P2 production. After ARF6 has activated PIP5K and/or PLD, a feed-forward loop is activated in which PLD-dependent PA production leads to the activation of PIP5K, PtdIns(4,5)P2, and PLD. Lipid enzymes are shown in orange and lipid products in green. Black arrows denote activation of a downstream protein or process, and green arrows denote conversion to a lipid product

[112]. Such very small regulatory motifs come together to form the overall regulatory network of an organism.

From the viewpoint of biological relativity, all scales of life ranging from the genome to the phenome partake in regulation and together represent the genetic program [2]. This entails a comprehensive regulatory network which allows a holistic view of an organism. Naturally, not all parts of the network interact at every time while other parts may be very active. These interactions need to be under tight time and space control and it is possible to find regulatory motifs as substructures in the overall network such as feedback loops and feed-forward loops (Fig. 4). Such motifs tie together via gene products or miRNAs which interact with multiple targets.

4.4 Regulatory Motifs

For example, a gene may co-produce a miRNA within an intron, which in turn downregulates the protein abundance of the same gene (Fig. 3). This represents the shortest possible feedback loop (FBL) including miRNAs. An equally short path would result from a gene coding for a TF which downregulates the gene itself. An example for this is PHOX2B [113]. Combined feedback loops consisting of miRNAs, TFs, and genes (Fig. 3) can also be envisioned. Such structures can represent molecular switches between cell states. A double-negative feedback loop between the miR-200 family and ZEB1-SIP1, for example, represents such a switch controlling the epithelial to mesenchymal transition [114]. Feed-forward loops (FFL) involving TFs and miRNAs are formed when a TF and an miRNA co-regulate a common target gene. This can, for example, be useful for signal noise buffering [115]. Another function could be the suppression of "leaky" transcription of target genes by reinforcing transcriptional control with posttranscriptional control [116]. Such small motifs can further be combined to larger regulatory motifs. Zhang et al. comprehensively summarized the possible regulatory motifs concerning TFs and miRNAs [117].

Perhaps not very intuitive at first glimpse, a very short path regulatory structure exists (Fig. 5). The miRNA targets (A–C) are all targeted by the same miRNA (for sake of simplicity, all outside interactions are ignored). When the transcription of one of the mRNAs (e.g.: mRNA A) increases while the miRNA expression remains constant, more miRNA target sites become available, acting as decoys. Thereby, the miRNA regulation is reduced for all of its targets (Fig. 5). The effect is most noticeable if mRNA expression levels, the target site-binding strengths, their accessibility, and multiplicity are similar for all targets [109].

In summary, miRNAs and their targets can be detected computationally. MicroRNA source information is important for integration with pathways and pathway extension. Since not all miRNAs or targets are known, other approaches to network

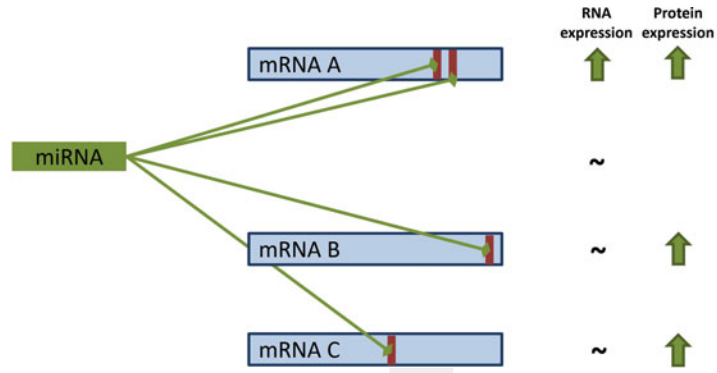

Fig. 5 Change in mRNA expression levels of one of the targets of one miRNA leads to increased protein expression for all its targets. For this example, assuming highly similar targets, mRNA A has two target sites for the miRNA, which would lead to inactivation of twice the increase in expression for mRNA A. Similarly, protein expression would increase by one-third of the expression change per mRNA

extension such as using literature mining are needed. Understanding regulatory pathways will be easier when regulatory motifs are annotated. Incorporation of (relative) expression levels for all players in such motifs will allow mathematical modeling of these substructures using for example Petri nets [118]. In the future, such motifs can be abstracted as circuits and combined into circuit diagrams.

5 Conclusion

Since the discovery of miRNAs in 1993 these small RNA molecules involved in posttranscriptional gene regulation have sparked a lot of research interest [19, 119]. They are generally thought to be involved in the downregulation of protein abundance, but have been shown to be involved in upregulation, as well. Recently, evidence is accumulating that they can be reimported into the nucleus where they are involved in transcriptional regulation. This versatility of miRNAs is not currently modeled in available pathway databases such as Reactome and KEGG. The methodology introduced at the beginning of the previous section which extends disease networks by miRNAs that are co-expressed with genes or target genes in known pathways can be extended with adding protein-protein interaction information. When that is done, all other methods described afterwards are implicitly contained in the solution if all miRNAs and their targets would be known or could be predicted. VANESA is a tool facilitating this approach [99, 120] including a database combining miRBase, miRTarBase, and TarBase as well as granting access to KEGG and the protein

interaction databases Mint [121], IntAct [122], and HPRD [123]. The system further supports modeling using Petri nets. However, extension of networks with other means such as by disease association as well as detection of regulatory motifs and mapping of expression levels is not currently possible. Incorporation of expression information ensures that miRNA and targets are co-expressed and that regulation is possible. It also indicates whether regulation would lead to measurable effects. For example, if target levels and target-site abundance among all expressed targets are high and miRNA expression is comparably low, no regulatory effect should be expected (Fig. 5).

Future studies involving miRNA regulation should thus consider the following points:

1. Limitations of miRNA and target prediction algorithms or databases hosting them
2. The sources for miRNAs (intergenic or genic)
3. Integration of miRNAs with known pathways
4. Extension of such pathways using various methods
5. Incorporation of expression data
6. Analysis of miRNA expression levels together with its target (s) expression
7. Detection of regulatory motifs and their mathematical modeling

Taking into account these information and making the resulting models FAIR similarly to data publishing guidelines [124] will help assign function to miRNAs and ensure that miRNAs will be employed as biomarkers and drugs for precision medicine in the future.

References

1. Crick F (1970) Central dogma of molecular biology. Nature 227:561–563
2. Noble D (2012) A theory of biological relativity: no privileged level of causation. Interface Focus 2:55–64. https://doi.org/10.1098/rsfs.2011.0067
3. Liu H, Lei C, He Q, Pan Z, Xiao D, Tao Y (2018) Nuclear functions of mammalian MicroRNAs in gene regulation, immunity and cancer. Mol Cancer 17:64. https://doi.org/10.1186/s12943-018-0765-5
4. Yousef M, Allmer J (2014) miRNomics: microRNA biology and computational analysis. Humana Press, Totowa, NJ
5. Iwakawa H, Tomari Y (2015) The functions of microRNAs: mRNA decay and translational repression. Trends Cell Biol 25:651–665. https://doi.org/10.1016/j.tcb.2015.07.011
6. Ørom UA, Nielsen FC, Lund AH (2008) MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. Mol Cell 30:460–471. https://doi.org/10.1016/j.molcel.2008.05.001
7. Grundhoff A, Sullivan CS (2011) Virus-encoded microRNAs. Virology 411:325–343. https://doi.org/10.1016/j.virol.2011.01.002
8. Skalsky RL, Cullen BR (2010) Viruses, microRNAs, and host interactions. Annu Rev Microbiol 64:123–141. https://doi.org/10.1146/annurev.micro.112408.134243

9. Saçar Demirci MD, Bağcı C, Allmer J (2016) Differential expression of Toxoplasma gondii microRNAs in murine and human hosts. In: Non-coding RNAs and inter-kingdom communication. Springer International Publishing, Cham, pp 143–159
10. Saçar MD, Bağcı C, Allmer J (2014) Computational prediction of MicroRNAs from toxoplasma gondii potentially regulating the hosts' gene expression. Genomics, Proteomics Bioinformatics 12:228–238. https://doi.org/10.1016/j.gpb.2014.09.002
11. Liu S, Weiner HL (2016) Control of the gut microbiome by fecal microRNA. Microb cell (Graz, Austria) 3:176–177. https://doi.org/10.15698/mic2016.04.492
12. Williams MR, Stedtfeld RD, Tiedje JM, Hashsham SA (2017) MicroRNAs-based interdomain communication between the host and members of the gut microbiome. Front Microbiol 8. https://doi.org/10.3389/fmicb.2017.01896
13. Baker M (2010) MicroRNA profiling: separating signal from noise. Nat Methods 7:687–692. https://doi.org/10.1038/nmeth0910-687
14. Chugh P, Dittmer DP (2012) Potential pitfalls in microRNA profiling. Wiley Interdiscip Rev RNA 3:601–616
15. Dong H, Lei J, Ding L, Wen Y, Ju H, Zhang X (2013) MicroRNA: function, detection, and bioanalysis. Chem Rev 113:6207–6233. https://doi.org/10.1021/cr300362f
16. Saçar MD, Allmer J (2014) Machine learning methods for microRNA gene prediction. In: Yousef M, Allmer J (eds) miRNomics: microRNA biology and computational analysis SE-10. Humana Press, pp 177–187
17. Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, Darnell JC, Darnell RB (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456:464–469. https://doi.org/10.1038/nature07488
18. Hamzeiy H, Allmer J, Yousef M (2014) Computational methods for microRNA target prediction. Methods Mol Biol 1107:207–221. https://doi.org/10.1007/978-1-62703-748-8_12
19. Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75:843–854
20. Berezikov E, Cuppen E, RH P (2006) Approaches to microRNA discovery. Nat Genet 38(Suppl):S2–S7. https://doi.org/10.1038/ng1794
21. Lim LP, Lau NC, Weinstein EG, Abdelhakim A, Yekta S, Rhoades MW, Burge CB, Bartel DP (2003) The microRNAs of Caenorhabditis elegans. Genes Dev 17:991–1008. https://doi.org/10.1101/gad.1074403
22. Wang X, Zhang J, Li F, Gu J, He T, Zhang X, Li Y (2005) MicroRNA identification based on sequence and structure alignment. Bioinformatics 21:3610–3614. https://doi.org/10.1093/bioinformatics/bti562
23. Nam J-W, Kim J, Kim S-K, Zhang B-T (2006) ProMiR II: a web server for the probabilistic prediction of clustered, nonclustered, conserved and nonconserved microRNAs. Nucleic Acids Res 34:W455–W458. https://doi.org/10.1093/nar/gkl321
24. Xue C, Li F, He T, Liu G-P, Li Y, Zhang X (2005) Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine. BMC Bioinformatics 6:310. https://doi.org/10.1186/1471-2105-6-310
25. Sewer A, Paul N, Landgraf P, Aravin A, Pfeffer S, Brownstein MJ, Tuschl T, van Nimwegen E, Zavolan M (2005) Identification of clustered microRNAs using an ab initio prediction method. BMC Bioinformatics 6:267. https://doi.org/10.1186/1471-2105-6-267
26. Hertel J, Stadler PF (2006) Hairpins in a Haystack: recognizing microRNA precursors in comparative genomics data. Bioinformatics 22:e197–e202. https://doi.org/10.1093/bioinformatics/btl257
27. Yousef M, Nebozhyn M, Shatkay H, Kanterakis S, Showe LC, Showe MK (2006) Combining multi-species genomic data for microRNA identification using a Naive Bayes classifier. Bioinformatics 22:1325–1334. https://doi.org/10.1093/bioinformatics/btl094
28. Huang T-H, Fan B, Rothschild MF, Hu Z-L, Li K, Zhao S-H (2007) MiRFinder: an improved approach and software implementation for genome-wide fast microRNA precursor scans. BMC Bioinformatics 8:341. https://doi.org/10.1186/1471-2105-8-341
29. Jiang P, Wu H, Wang W, Ma W, Sun X, Lu Z (2007) MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined

features. Nucleic Acids Res 35:W339–W344. https://doi.org/10.1093/nar/gkm368

30. Terai G, Komori T, Asai K (2081–2090) Kin T (2007) miRRim: a novel system to find conserved miRNAs with high sensitivity and specificity. https://doi.org/10.1261/rna.655107.been
31. Friedländer MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, Rajewsky N (2008) Discovering microRNAs from deep sequencing data using miRDeep. Nat Biotechnol 26:407–415. https://doi.org/10.1038/nbt1394
32. Hackenberg M, Sturm M, Langenberger D, Falcón-Pérez JM, Aransay AM (2009) miRanalyzer: a microRNA detection and analysis tool for next-generation sequencing experiments. Nucleic Acids Res 37:W68–W76. https://doi.org/10.1093/nar/gkp347
33. Oulas A, Boutla A, Gkirtzou K, Reczko M, Kalantidis K, Poirazi P (2009) Prediction of novel microRNA genes in cancer-associated genomic regions—a combined computational and experimental approach. Nucleic Acids Res 37:3276–3287. https://doi.org/10.1093/nar/gkp120
34. Kadri S, Hinman V, Benos PV (2009) HHMMiR: efficient de novo prediction of microRNAs using hierarchical hidden Markov models. BMC Bioinformatics 10(Suppl 1): S35. https://doi.org/10.1186/1471-2105-10-S1-S35
35. Mathelier A, Carbone A (2010) MIReNA: finding microRNAs with high accuracy and no learning at genome scale and from deep sequencing data. Bioinformatics 26:2226–2234. https://doi.org/10.1093/bioinformatics/btq329
36. Wu Y, Wei B, Liu H, Li T, Rayner S (2011) MiRPara: a SVM-based software tool for prediction of most probable microRNA coding regions in genome scale sequences. BMC Bioinformatics 12:107. https://doi.org/10.1186/1471-2105-12-107
37. Tempel S, Tahi F (2012) A fast ab-initio method for predicting miRNA precursors in genomes. Nucleic Acids Res 40:e80. https://doi.org/10.1093/nar/gks146
38. Gao D, Middleton R, Rasko JEJ, Ritchie W (2013) miREval 2.0: a web tool for simple microRNA prediction in genome sequences. Bioinformatics 29:3225–3226. https://doi.org/10.1093/bioinformatics/btt545
39. Lei J, Sun Y (2014) miR-PREFeR: an accurate, fast and easy-to-use plant miRNA prediction tool using small RNA-Seq data. Bioinformatics 30:2837–2839. https://doi.org/10.1093/bioinformatics/btu380
40. Tran VDT, Tempel S, Zerath B, Zehraoui F, Tahi F (2015) miRBoost: boosting support vector machines for microRNA precursor classification. RNA 21:775–785. https://doi.org/10.1261/rna.043612.113
41. Chen J, Wang X, Liu B (2016) iMiRNA-SSF: improving the identification of microRNA precursors by combining negative sets with different distributions. Sci Rep 6:19062. https://doi.org/10.1038/srep19062
42. Saçar Demirci MD, Baumbach J, Allmer J (2017) On the performance of pre-microRNA detection algorithms. Nat Commun 8:330. https://doi.org/10.1038/s41467-017-00403-z
43. Lu Yi, Aras AS, Halushka MK (2018) miRge 2.0: an updated tool to comprehensively analyze microRNA sequencing data, bioRxiv, https://doi.org/10.1101/250779
44. Gomes CPC, Cho J-H, Hood L, Franco OL, Pereira RW, Wang K (2013) A review of computational tools in microRNA discovery. Front Genet 4:81. https://doi.org/10.3389/fgene.2013.00081
45. van der Burgt A, Fiers MWJE, Nap J-P, van Ham RCHJ (2009) In silico miRNA prediction in metazoan genomes: balancing between sensitivity and specificity. BMC Genomics 10:204. https://doi.org/10.1186/1471-2164-10-204
46. Hofacker IL (2003) Vienna RNA secondary structure server. Nucleic Acids Res 31:3429–3431. https://doi.org/10.1093/nar/gkg599
47. Zeng C, Wang W, Zheng Y, Chen X, Bo W, Song S, Zhang W, Peng M (2010) Conservation and divergence of microRNAs and their functions in Euphorbiaceous plants. Nucleic Acids Res 38:981–995. https://doi.org/10.1093/nar/gkp1035
48. Liang H, Li W-H (2009) Lowly expressed human microRNA genes evolve rapidly. Mol Biol Evol 26:1195–1198. https://doi.org/10.1093/molbev/msp053
49. Saçar Demirci MD, Allmer J (2017) Delineating the impact of machine learning elements in pre-microRNA detection. PeerJ 5:e3131. https://doi.org/10.7717/peerj.3131
50. Marcinkowska M, Szymanski M, Krzyzosiak WJ, Kozlowski P (2011) Copy number variation of microRNA genes in the human genome. BMC Genomics 12:183. https://doi.org/10.1186/1471-2164-12-183

51. Erson-Bensan AE (2014) Introduction to microRNAs in biological systems. Methods Mol Biol 1107:1–14. https://doi.org/10.1007/978-1-62703-748-8_1
52. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jungkamp A-C, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141:129–141. https://doi.org/10.1016/j.cell.2010.03.009
53. Chi SW, Zang JB, Mele A, Darnell RB (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature 460:479–486. https://doi.org/10.1038/nature08170.Ago
54. Helwak A, Kudla G, Dudnakova T, Tollervey D (2013) Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell 153:654–665. https://doi.org/10.1016/j.cell.2013.03.043
55. Hsu S-D, Tseng Y-T, Shrestha S, Lin Y-L, Khaleel A, Chou C-H, Chu C-F, Huang H-Y, Lin C-M, Ho S-Y, Jian T-Y, Lin F-M, Chang T-H, Weng S-L, Liao K-W, Liao I-E, Liu C-C, Huang H-D (2014) miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. Nucleic Acids Res 42:D78–D85. https://doi.org/10.1093/nar/gkt1266
56. Vergoulis T, Vlachos IS, Alexiou P, Georgakilas G, Maragkakis M, Reczko M, Gerangelos S, Koziris N, Dalamagas T, Hatzigeorgiou AG (2012) TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. Nucleic Acids Res 40:D222–D229. https://doi.org/10.1093/nar/gkr1161
57. Krüger J, Rehmsmeier M (2006) RNAhybrid: microRNA target prediction easy, fast and flexible. Nucleic Acids Res 34:W451–W454. https://doi.org/10.1093/nar/gkl243
58. Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, Rajewsky N (2005) Combinatorial microRNA target predictions. Nat Genet 37:495–500. https://doi.org/10.1038/ng1536
59. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120:15–20. https://doi.org/10.1016/j.cell.2004.12.035
60. Miranda KC, Huynh T, Tay Y, Ang Y-S, Tam W-L, Thomson AM, Lim B, Rigoutsos I (2006) A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell 126:1203–1217. https://doi.org/10.1016/j.cell.2006.07.031
61. Sethupathy P, Corda B, Hatzigeorgiou AG (2006) TarBase: a comprehensive database of experimentally supported animal microRNA targets. RNA 12:192–197. https://doi.org/10.1261/rna.2239606
62. Yousef M, Jung S, Kossenkov AV, Showe LC, Showe MK (2007) Naïve Bayes for microRNA target predictions—machine learning for microRNA targets. Bioinformatics 23:2987–2992. https://doi.org/10.1093/bioinformatics/btm484
63. Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E (2007) The role of site accessibility in microRNA target recognition. Nat Genet 39:1278–1284. https://doi.org/10.1038/ng2135
64. Maragkakis M, Alexiou P, Papadopoulos GL, Reczko M, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, Simossis VA, Sethupathy P, Vergoulis T, Koziris N, Sellis T, Tsanakas P, Hatzigeorgiou AG (2009) Accurate microRNA target prediction correlates with protein repression levels. BMC Bioinformatics 10:295. https://doi.org/10.1186/1471-2105-10-295
65. Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T (2009) miRecords: an integrated resource for microRNA-target interactions. Nucleic Acids Res 37:D105–D110. https://doi.org/10.1093/nar/gkn851
66. Hsu S-D, Lin F-M, Wu W-Y, Liang C, Huang W-C, Chan W-L, Tsai W-T, Chen G-Z, Lee C-J, Chiu C-M, Chien C-H, Wu M-C, Huang C-Y, Tsou A-P, Huang H-D (2011) miRTarBase: a database curates experimentally validated microRNA-target interactions. Nucleic Acids Res 39:D163–D169. https://doi.org/10.1093/nar/gkq1107
67. Dweep H, Sticht C, Pandey P, Gretz N (2011) miRWalk—database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform 44:839–847. https://doi.org/10.1016/j.jbi.2011.05.002
68. Elefant N, Berger A, Shein H, Hofree M, Margalit H, Altuvia Y (2011) RepTar: a database of predicted cellular targets of host and viral miRNAs. Nucleic Acids Res 39:D188–D194. https://doi.org/10.1093/nar/gkq1233

69. Li J-H, Liu S, Zhou H, Qu L-H, Yang J-H (2014) starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res 42:D92–D97. https://doi.org/10.1093/nar/gkt1248
70. Chiu H-S, Llobet-Navas D, Yang X, Chung W-J, Ambesi-Impiombato A, Iyer A, Kim HR, Seviour EG, Luo Z, Sehgal V, Moss T, Lu Y, Ram P, Silva J, Mills GB, Califano A, Sumazin P (2015) Cupid: simultaneous reconstruction of microRNA-target and ceRNA networks. Genome Res 25:257–267. https://doi.org/10.1101/gr.178194.114
71. Bandyopadhyay S, Ghosh D, Mitra R, Zhao Z (2015) MBSTAR: multiple instance learning for predicting specific functional binding sites in microRNA targets. Sci Rep 5:8004. https://doi.org/10.1038/srep08004
72. Liu S, Li J-H, Wu J, Zhou K-R, Zhou H, Yang J-H, Qu L-H (2015) StarScan: a web server for scanning small RNA targets from degradome sequencing data. Nucleic Acids Res 43: W480–W486. https://doi.org/10.1093/nar/gkv524
73. Riffo-Campos Á, Riquelme I, Brebi-Mieville P (2016) Tools for sequence-based miRNA target prediction: what to choose? Int J Mol Sci 17:1987. https://doi.org/10.3390/ijms17121987
74. Lewis BP, Shih I, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. Cell 115:787–798
75. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS (2003) MicroRNA targets in Drosophila. Genome Biol 5:R1. https://doi.org/10.1186/gb-2003-5-1-r1
76. Kiriakidou M, Nelson PT, Kouranov A, Fitziev P, Bouyioukos C, Mourelatos Z, Hatzigeorgiou A (2004) A combined computational-experimental approach predicts human microRNA targets. Genes Dev 18:1165–1178. https://doi.org/10.1101/gad.1184704
77. Peterson SM, JA T, Ufkin ML, Sathyanarayana P, Liaw L, Congdon CB (2014) Common features of microRNA target prediction tools. Front Genet 5:23. https://doi.org/10.3389/fgene.2014.00023
78. Yousef M, Allmer J, Khalifa W (2016) Feature selection for microRNA target prediction comparison of one-class feature selection methodologies. In: BIOINFORMATICS 2016—7th international conference on bioinformatics models, methods and algorithms, Proceedings; Part of 9th international joint conference on biomedical engineering systems and technologies, BIOSTEC 2016
79. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS (2004) Human microRNA targets. PLoS Biol 2:e363. https://doi.org/10.1371/journal.pbio.0020363
80. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2
81. Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R (2004) Fast and effective prediction of microRNA/target duplexes. RNA 10:1507–1517. https://doi.org/10.1261/rna.5248604
82. Lai EC (2004) Predicting and validating microRNA targets. Genome Biol 5:115. https://doi.org/10.1186/gb-2004-5-9-115
83. Yousef M, Nigatu D, Levy D, Allmer J, Henkel W (2017) Categorization of species based on their microRNAs employing sequence motifs, information-theoretic sequence feature extraction, and k-mers. EURASIP J Adv Signal Process 2017:70. https://doi.org/10.1186/s13634-017-0506-8
84. Heyn J, Hinske LC, Ledderose C, Limbeck E, Kreth S (2013) Experimental miRNA target validation. Methods Mol Biol 936:83–90. https://doi.org/10.1007/978-1-62703-083-0_7
85. Thomson DW, Bracken CP, Goodall GJ (2011) Experimental strategies for microRNA target identification. Nucleic Acids Res 39:6845–6853. https://doi.org/10.1093/nar/gkr330
86. Chou C-H, Shrestha S, Yang C-D, Chang N-W, Lin Y-L, Liao K-W, Huang W-C, Sun T-H, Tu S-J, Lee W-H, Chiew M-Y, Tai C-S, Wei T-Y, Tsai T-R, Huang H-T, Wang C-Y, Wu H-Y, Ho S-Y, Chen P-R, Chuang C-H, Hsieh P-J, Wu Y-S, Chen W-L, Li M-J, Wu Y-C, Huang X-Y, Ng FL, Buddhakosai W, Huang P-C, Lan K-C, Huang C-Y, Weng S-L, Cheng Y-N, Liang C, Hsu W-L, Huang H-D (2018) miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. Nucleic Acids Res 46:D296–D302. https://doi.org/10.1093/nar/gkx1067
87. Saçar MD, Allmer J (2013) Current limitations for computational analysis of miRNAs in cancer. Pakistan J Clin Biomed Res 1:3–5
88. Koo J, Zhang J, Chaterji S (2018) Tiresias: context-sensitive approach to decipher the presence and strength of microRNA

regulatory interactions. Theranostics 8:277–291. https://doi.org/10.7150/thno.22065

89. Kim VN, Han J, Siomi MC (2009) Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol 10:126–139. https://doi.org/10.1038/nrm2632
90. Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, Brownstein MJ, Tuschl T, Margalit H (2005) Clustering and conservation patterns of human microRNAs. Nucleic Acids Res 33:2697–2706. https://doi.org/10.1093/nar/gki567
91. Mechtler P, Johnson S, Slabodkin H, Cohanim AB, Brodsky L, Kandel ES (2017) The evidence for a microRNA product of human DROSHA gene. RNA Biol 14:1508–1513. https://doi.org/10.1080/15476286.2017.1342934
92. Acar İE, Saçar Demirci MD, Groß U, Allmer J (2018) The expressed MicroRNA—mRNA interactions of Toxoplasma gondii. Front Microbiol 8. https://doi.org/10.3389/fmicb.2017.02630
93. Leinonen R, Sugawara H, Shumway M (2011) The sequence read archive. Nucleic Acids Res 39:D19–D21. https://doi.org/10.1093/nar/gkq1019
94. Fei Y, Wang R, Li H, Liu S, Zhang H, Huang J (2017) DPMIND: degradome-based Plant MiRNA-target interaction and network database. Bioinformatics. https://doi.org/10.1093/bioinformatics/btx824
95. Kozomara A, Griffiths-Jones S (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 42:D68–D73. https://doi.org/10.1093/nar/gkt1181
96. Brinkrolf C, Janowski SJ, Kormeier B, Lewinski M, Hippe K, Borck D, Hofestädt R (2014) VANESA—a software application for the visualization and analysis of networks in system biology applications. J Integr Bioinform 11:239. https://doi.org/10.2390/biecoll-jib-2014-239
97. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, Caudy M, Garapati P, Gillespie M, Kamdar MR, Jassal B, Jupe S, Matthews L, May B, Palatnik S, Rothfels K, Shamovsky V, Song H, Williams M, Birney E, Hermjakob H, Stein L, D'Eustachio P (2014) The Reactome pathway knowledgebase. Nucleic Acids Res 42:D472–D477. https://doi.org/10.1093/nar/gkt1102
98. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30
99. Hamzeiy H, Suluyayla R, Brinkrolf C, Janowski SJ, Hofestaedt R, Allmer J (2017) Visualization and analysis of microRNAs within KEGG pathways using VANESA. J Integr Bioinform 14. https://doi.org/10.1515/jib-2016-0004
100. Le DH, Verbeke L, Son LH, Chu DT, Pham VH (2017) Random walks on mutual microRNA-target gene interaction network improve the prediction of disease-associated microRNAs. BMC Bioinformatics 18:1–13. https://doi.org/10.1186/s12859-017-1924-1
101. Zeng X, Zhang X, Zou Q (2016) Integrative approaches for predicting microRNA function and prioritizing disease-related microRNA using biological interaction networks. Brief Bioinform 17:193–203. https://doi.org/10.1093/bib/bbv033
102. Jiang Q, Hao Y, Wang G, Juan L, Zhang T, Teng M, Liu Y, Wang Y (2010) Prioritization of disease microRNAs through a human phenome-microRNAome network. BMC Syst Biol 4(Suppl 1):S2. https://doi.org/10.1186/1752-0509-4-S1-S2
103. Jiang Q, Hao Y, Wang G, Zhang T, Wang Y (2010) Weighted network-based inference of human microRNA-disease associations. In: 2010 Fifth international conference on frontier of computer science and technology. IEEE, pp 431–435
104. Wang D, Wang J, Lu M, Song F, Cui Q (2010) Inferring the human microRNA functional similarity and functional network based on microRNA-associated diseases. Bioinformatics 26:1644–1650. https://doi.org/10.1093/bioinformatics/btq241
105. Xu J, Li C-X, Li Y-S, Lv J-Y, Ma Y, Shao T-T, Xu L-D, Wang Y-Y, Du L, Zhang Y-P, Jiang W, Li C-Q, Xiao Y, Li X (2011) MiRNA-miRNA synergistic network: construction via co-regulating functional modules and disease miRNA topological features. Nucleic Acids Res 39:825–836. https://doi.org/10.1093/nar/gkq832
106. Chen X, Yan G-Y (2015) Semi-supervised learning for potential human microRNA-disease associations inference. Sci Rep 4:5501. https://doi.org/10.1038/srep05501
107. Kandhro AH, Shoombuatong W, Nantasenamat C, Prachayasittikul V, Nuchnoi P (2017) The microRNA interaction network of lipid diseases. Front Genet 8:1–14. https://doi.org/10.3389/fgene.2017.00116
108. Honardoost MA, Naghavian R, Ahmadinejad F, Hosseini A, Ghaedi K

(2015) Integrative computational mRNA-miRNA interaction analyses of the autoimmune-deregulated miRNAs and well-known Th17 differentiation regulators: an attempt to discover new potential miRNAs involved in Th17 differentiation. Gene 572:153–162. https://doi.org/10.1016/j.gene.2015.08.043

109. Robinson JM, Henderson WA (2018) Modelling the structure of a ceRNA-theoretical, bipartite microRNA-mRNA interaction network regulating intestinal epithelial cellular pathways using R programming. BMC Res Notes 11:1–7. https://doi.org/10.1186/s13104-018-3126-y
110. van den Bout I, Divecha N (2009) PIP5K-driven PtdIns(4,5)P2 synthesis: regulation and cellular functions. J Cell Sci 122:3837–3850. https://doi.org/10.1242/jcs.056127
111. Han J, Pedersen JS, Kwon SC, Belair CD, Kim Y, Yeom K, Yang W, Haussler D, Blelloch R, Kim VN (2009) Posttranscriptional crossregulation between Drosha and DGCR8. Cell 136:75–84. https://doi.org/10.1016/j.cell.2008.10.053
112. Crews ST, Pearson JC (2009) Transcriptional autoregulation in development. Curr Biol 19: R241–R246. https://doi.org/10.1016/j.cub.2009.01.015
113. Cargnin F, Flora A, Di Lascio S, Battaglioli E, Longhi R, Clementi F, Fornasari D (2005) PHOX2B regulates its own expression by a transcriptional auto-regulatory mechanism. J Biol Chem 280:37439–37448. https://doi.org/10.1074/jbc.M508368200
114. Bracken CP, Gregory PA, Kolesnikoff N, Bert AG, Wang J, Shannon MF, Goodall GJ (2008) A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. Cancer Res 68:7846–7854. https://doi.org/10.1158/0008-5472.CAN-08-1942
115. Osella M, Bosia C, Corá D, Caselle M (2011) The role of incoherent microRNA-mediated feedforward loops in noise buffering. PLoS Comput Biol 7. https://doi.org/10.1371/journal.pcbi.1001101
116. Tsang J, Zhu J, van Oudenaarden A (2007) MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. Mol Cell 26:753–767. https://doi.org/10.1016/j.molcel.2007.05.018
117. Zhang HM, Kuang S, Xiong X, Gao T, Liu C, Guo AY (2013) Transcription factor and microRNA co-regulatory loops: Important regulatory motifs in biological processes and diseases. Brief Bioinform 16:45–58. https://doi.org/10.1093/bib/bbt085
118. Yousef M, Trinh HV, Allmer J (2014) Intersection of microRNA and gene regulatory networks and their implication in cancer. Curr Pharm Biotechnol 15:445–454. https://doi.org/10.2174/1389201015666140519120855
119. Wightman B, Ha I, Ruvkun G (1993) Post-transcriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell 75:855–862
120. Hamzeiy H, Suluyayla R, Brinkrolf C, Janowski SJ, Hofestädt R, Allmer J (2018) Visualization and analysis of miRNAs implicated in amyotrophic lateral sclerosis within gene regulatory pathways. Stud Heal Technol Inform 253:183–187
121. Licata L, Briganti L, Peluso D, Perfetto L, Iannuccelli M, Galeota E, Sacco F, Palma A, Nardozza AP, Santonico E, Castagnoli L, Cesareni G (2012) MINT, the molecular interaction database: 2012 Update. Nucleic Acids Res 40
122. Kerrien S, Aranda B, Breuza L, Bridge A, Broackes-Carter F, Chen C, Duesbury M, Dumousseau M, Feuermann M, Hinz U, Jandrasits C, Jimenez RC, Khadake J, Mahadevan U, Masson P, Pedruzzi I, Pfeiffenberger E, Porras P, Raghunath A, Roechert B, Orchard S, Hermjakob H (2012) The IntAct molecular interaction database in 2012. Nucleic Acids Res 40
123. Liu B, Hu B (2010) HPRD: a high performance RDF database. Int J Parallel Emergent Distrib Syst 25:123–133
124. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PA, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3:160018. https://doi.org/10.1038/sdata.2016.18