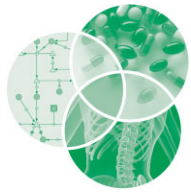
Advanced Review



Genome-wide approaches in the study of microRNA biology

Melissa L. Wilbert and Gene W. Yeo*

MicroRNAs (miRNAs), a class of ~21–23 nucleotide long non-coding RNAs (ncRNAs), have critical roles in diverse biological processes that encompass development, proliferation, apoptosis, stress response, and fat metabolism. miRNAs recognize their target mRNA transcripts by partial sequence complementarity and collectively have been estimated to regulate the majority of human genes. Consequently, misregulation of miRNAs or disruption of their target sites in genes has been implicated in a variety of human diseases ranging from cancer metastasis to neurological disorders. With the development and availability of genomic technologies and computational approaches, the field of miRNA biology has advanced tremendously over the last decade. Here we review the genome-wide approaches that have allowed for the discovery of new miRNAs, the characterization of their targets, and a systems-level view of their impact. *WIREs Syst Biol Med* 2011 DOI: 10.1002/wsbm.128

INTRODUCTION

The field of miRNA biology emerged with the discovery that the gene *lin-4*, which controls developmental timing in the nematode *Caenorhabditis elegans*, surprisingly did not code for protein, but instead acted as a ~22 nt RNA transcript.[1,2] Experiments showed that this small RNA molecule regulated its target, *lin-14*, by base-pairing to the 3′ untranslated region (3′ UTR) of *lin-14* mRNA via partial sequence complementarity.[2,3] This work established the dogmatic understanding of miRNAs as posttranscriptional regulators that target the 3′ UTRs of protein-coding genes to repress gene expression. Since then, miRNAs have been shown to play a variety of regulatory roles and target other genic regions in addition to 3′ UTRs.[4–12] The discovery that the second characterized miRNA, *let-7*,[13] was evolutionarily conserved across *bilateria*,[14] and that its expression was regulated through the course of development, sparked concerted efforts to identify other miRNAs and elucidate their functions. A framework to study the consequences of miRNA regulation emerged from the findings that 3′ UTRs widely contain evolutionarily conserved elements that are complementary to the 5′ end of certain miRNAs and mediate repression by these miRNAs.[2,3,5,13,15–20] From these foundations, the field of miRNA biology has quickly progressed through the application of genome-wide approaches for identification of miRNAs and their targets. Here we review these tools and their contributions to our understanding of miRNAs as global cellular regulators.

COMPUTATIONAL APPROACHES TO miRNA DISCOVERY

Soon after the first experimental efforts identified a significant cohort of miRNAs,[14,21–23] distinguishing characteristics of these transcripts were utilized for the development of *in silico* miRNA prediction tools. Mature miRNAs are derived from a longer (~70 nt) precursor hairpin structure (pre-miRNA).[21–28] The most likely secondary structure of putative pre-miRNA sequences can be predicted using RNA folding algorithms, such as mfold and RNAfold,[29] to identify these characteristic stem-loop structures. These predictions, along with other common features among known miRNAs, such as base-pairing within the hairpin, evolutionary conservation, nucleotide usage, and structural features, were the basis for the first miRNA discovery tool, MiRscan.[30] Using MiRscan, Lim and colleagues were able to identify novel miRNAs by ranking predicted miRNA hairpin precursors conserved between *C. elegans* and

*Correspondence to: geneyeo@ucsd.edu

Department of Cellular and Molecular Medicine, Stem Cell Program, Institute for Genomic Medicine, University of California, San Diego, La Jolla, CA, USA

DOI: 10.1002/wsbm.128

C. briggsae based on these criteria.[30] Of these features, they found that base-pairing potential and sequence conservation across genomes were the most effective predictors of validated miRNAs. Later successful application of MiRscan to miRNA discovery in vertebrate genomes demonstrated that generic features of miRNAs are more broadly conserved than the sequences themselves.[31] Additionally, contextual sequence and conservation features flanking miRNA genes were observed to improve the sensitivity of MiRscan.[32] This work also improved the prediction of miRNAs encoded within introns through the analysis of introns in orthologous genes. Other groups have improved their algorithms by defining very specific patterns of conservation along the arms and loops of miRNA precursors, as well as the surrounding sequence.[33,34] Less strict conservation considerations, filtering for negative repetitive sequences, and different emphasis on homology have also been applied to these approaches.[35,36]

Due to selective pressures to maintain miRNA complementary sites,[37,38] conservation of these target sequences has served as a foundation for another group of miRNA discovery tools.[39–41] Generally, these algorithms first identify short conserved motifs in the 3′ UTRs of protein-coding genes. Next, genomic searches for conserved sequences complementary to these motifs are used to find potential miRNA genes. Alternatively, in plants a search of intergenic sequences for motifs complementary to mRNAs can be applied.[40] Where matches are found, the flanking region is tested for its ability to form stable hairpin structures. A disadvantage of these homology and conservation-based algorithms is their general inability to identify species-specific or otherwise divergent miRNAs. This is especially a concern given the growing body of evidence supporting the rapid evolution of miRNAs.[42–44]

To improve the sensitivity of miRNA prediction, *ab initio* algorithms have been developed. These approaches do not require conservation criteria, although some applications have chosen to maintain these considerations.[45–47] Instead, these methods rely on intrinsic characteristics of known miRNAs to define sets of features that can accurately describe the structure and sequence of these transcripts. Supervised learning algorithms are then used to weigh these features based on positive and negative training sets. Positive datasets include known miRNAs, whereas negative datasets are typically pseudo-hairpins—genomic sequences that form stem-loop like structures but do not contain a mature miRNA. The trained algorithm is used to measure the similarity of candidate miRNAs to known miRNAs. Careful definition of datasets, for example, through consideration of other non-coding RNAs (ncRNAs), and rigorous training have been shown to improve these approaches.[48] The choice of the machine-learning algorithm, either kernel-based (e.g., support machine vector, hidden Markov model, relaxed kernel density estimator) or logic based (e.g., decision trees, decision rules), is another significant variation of these methods. Support machine vectors (SMVs) are known for their success in a wide range of bioinformatics applications, including miRNA discovery.[49–52] However, it has been demonstrated that the use of a relaxed kernel density estimator, over an SMV, can be superior for species-specific miRNA discovery by exploiting more local information in the training dataset.[53] Jiang and colleagues have also described a random forest-based algorithm that outperforms previous SMV-based tools.[54] Some groups have combined kernel- and logic-based approaches, as well as other traditional sequence comparisons and structure prediction tools.[45,55] *Ab initio* approaches are particularly useful for application to rapidly evolving organisms and/or those with little homology to other genomes, as is the case with viruses such as the Epstein-Barr virus.[56]

HIGH-THROUGHPUT SEQUENCING FOR THE IDENTIFICATION OF miRNAS

Despite advances in *in silico* miRNA gene finding, sensitivity and false-positive rates remain a concern. Early approaches for genome-wide *in vivo* miRNA capture were limited by their abilities to clone and sequence these transcripts.[21–24,57–59] With the availability of next-generation sequencing technology, millions of small RNAs can be sequenced relatively inexpensively, offering an improved avenue for miRNA discovery.[60] An advantage of these approaches is their ability to detect evidence of expressed miRNAs, in a cell-type and context-specific manner. See Figure 1 for a review of *in silico* versus small RNA-seq (smRNA-seq) miRNA discovery approaches. One of the first passes at miRNA discovery using next-generation sequencing successfully identified the majority of miRNAs annotated in *C. elegans* at that time, some of which had only previously been detected through computational predictions or PCR assays, as well as 18 novel miRNAs.[61] In addition, smRNA-seq-based algorithms are less reliant on feature extraction from known miRNAs than purely *in silico* approaches, in particular with regard to conservation. For example, analyses of small RNA transcripts from human and chimpanzee brains more than doubled the number of known miRNA genes, and revealed a surprisingly large percent of novel miRNAs conserved only in mammals (30%) or

Approach		Criteria	Advantages	Disadvantages	
In silico prediction algorithms	Conservation- and homology-based algorithms	Conservation and homology; Hairpin predictions; Structure and sequence features of known miRNAs and surrounding sequences	Genome-wide miRNA predictions; Useful for prediction of evolutionarily-imporant miRNAs	Poor for detection of miRNAs in rapidly evolving genomes	Algorithm overfitting using features of known miRNAs can compromise sensitivity for novel miRNA identification
	Ab initio algorithms that do not use conservation criteria	Structure and sequence features of known miRNAs; Optional: hairpin predictions	Greater sensitivity for non-conserved miRNAs	Cannot predict miRNA expression patterns	
Predictions using high-throughput small RNA sequencing data		Hairpin predictions; Lengths and abundance of reads represent miRNA processing byproducts	High information content of reads; Identification of expressed miRNAs; Cell-type and context specific readout; Identification of isomiRs and non-conserved miRNAs	Volume of data and processing may be a technical challenge for some groups	

FIGURE 1 | Genome-wide approaches for miRNA discovery.

in a single species (8%).[62] This is compared to known miRNAs, also cloned by the study, whose original discovery was largely driven by conservation-based approaches and thus were significantly more enriched for transcripts conserved among vertebrates and mammals (75%), with only about 1% specific to humans. This work revealed the power of sequencing to identify miRNAs at an unprecedented level of sensitivity.

In 2008, the Rajewsky group developed miRDeep, the first publicly available software package for the discovery of miRNAs using smRNA-seq data.[63] The general strategy of smRNA-seq based miRNA discovery is to search genomic sequences for evidence of hairpin structures, and then determine if sequencing reads aligned to these structures mimic miRNA-processing byproducts. These byproducts are short sequences of largely invariant lengths corresponding to the mature miRNA, the complementary sequence (miRNA*) within the hairpin, and the loop of the hairpin. The most abundant of these transcripts is expected to be the mature miRNA sequence, as the miRNA* and hairpin loop are generally degraded during processing. The distinguishing abundance and lengths of these transcripts provide a relatively high level of information content that can be captured using high-throughput sequencing. This is an advantage for algorithms using this sequencing data to identify miRNAs, as a major barrier for purely *in silico* miRNA prediction algorithms has been the relatively low information content of short genomic miRNA sequences and even their longer precursors. However, interpretation of smRNA-seq data remains less than straightforward as biases arise from technical aspects related to sequencing accuracy, genomic mapping efficacy, and small RNA library preparation. Heterogeneous pools of small RNAs including degradation products and other ncRNAs, as well as sequence variations arising from single nucleotide polymorphisms and posttranscriptional RNA editing, are confounding factors that should be considered for accurate miRNA annotation.[64] Like purely *in silico* approaches, overfitting of algorithms using features of known miRNAs is also a concern, as this can compromise their ability to reliably identify novel miRNAs.

Despite these confounding factors, the Rajewsky group successfully identified new miRNAs using miRDeep in *C. elegans*, human, and dog.[63] Concurrently, the Graveley and Yeo groups reported a set of novel planarian (*S. mediterranea*) miRNAs mined from smRNA-seq data using a regularized least-squares classification algorithm, MIResque,[65] as did the Rajewsky group using miRDeep.[66] Later that year, the first web-based tool for miRNA prediction using sequencing data became available.[67] As in traditional miRNA finding, different groups have chosen to combine various tools for precursor predictions and free energy calculations to develop miRNA discovery pipelines.[68,69] Purification of miRNAs along with bound protein complexes[24] coupled with high-throughput sequencing is a powerful way to identify active transcripts.[70–76] To date, about 750 human miRNAs have been reported, as documented in miRBase, an online database repository for known miRNAs.[77,78] A caveat of these annotations is that additional experimental validation may be necessary to confirm that their transcripts are processed by the miRNA machinery and represent *bona fide* miRNAs.[64] As our knowledge of miRNA biology

improves and sequencing power increases, it is likely that the number and quality of miRNA annotations will continue to increase, particularly in large and/or poorly annotated genomes.[64,79]

In addition to identification of novel miRNAs, deep sequencing has also been proved to be a powerful tool for quantifying miRNA expression and defining variation within miRNA sequences. The comprehensiveness and sensitivity in identification of novel and lowly expressed transcripts are significant advantages of smRNA-seq data over traditional microarray platforms. The ability to determine absolute expression values that can be compared within and across experiments represents another distinct advantage of sequencing data over the relative quantifications given by microarray measurements.[80] Furthermore, sequencing approaches do not suffer from cross-hybridization artifacts observed in microarray experiments. Small RNA-seq has now been applied to measure miRNA expression and variation in a wide range of studies.[64,81–89] These studies capture the cell-type and context-specific expression of miRNAs, as suggested by previous approaches. However, Landgraf and colleagues point out that the number of miRNAs contributing to these unique profiles is relatively small compared to those that are ubiquitously expressed.[81] In particular, the miRNA expression profiles in embryonic and induced pluripotent stem cells, and differentiated states from these cells, have been an area of intense interest.[68,74,90–96] Importantly, it has been suggested that miRNAs may serve as precise indicators and effective determinants of pluripotency and cell fate.[97–107]

The nucleotide-level resolution that RNA sequencing provides has also influenced the way researchers approach the study of posttranscriptional modifications of miRNAs. MiRNAs have been shown to be the subjects of editing effects such as adenosine to inosine (A-to-I) editing, mediated by the ADAR proteins,[108] and posttranscriptional end modifications such as uridylation.[109] Through the use of smRNA-seq, the prevalence of these and other sequence modifications, and their implication for miRNA regulation, has come to light.[64,81,110,111] In response, the term 'isomiRs' has been coined to describe these miRNA-variants.[68] Editing can alter miRNA processing and targets and thereby their specific regulatory effects.[112] A growing body of evidence supports the idea of cell-type-specific miRNA variant profiles, emphasizing the unique physiological and pathological roles of these transcripts.[68,113] Accordingly, web-based databases and tools have been developed to catalog, process, and analyze this information, particularly in the context of biological pathways.[113] A detailed understanding of cell-type-specific miRNA expression and variation is vital for extrapolation of their roles in orchestrating regulatory networks of these cells.

miRNA TARGET PREDICTION ALGORITHMS

The critical link for understanding the diverse roles of miRNAs in the cell is the mRNA transcripts they target. It is the elucidation of these targets and rules that govern mRNA target recognition that has been, and will continue to be, instrumental in our understanding of miRNA biology and how their misregulation can lead to disease. The original basis for miRNA target identification arose from early experiments that demonstrated that partial sequence complementarity was sufficient for miRNA–mRNA interactions.[2,3,13,16] Searching genomic sequences for short stretches of nucleotides with near perfect complementarity to known miRNAs was successful for miRNA target finding in plants, but not in animals where imperfect pairing is prevalent.[114] In 2003, Lewis and colleagues observed that miRNA target sites tend to be more conserved across evolution than what would be expected by chance.[18] By refining the complementary region to bases 2–8 from the 5′ end of the miRNA (the 'seed' region) and requiring pairing to conserved 3′ UTR mRNA sequences, their algorithm, TargetScan, successfully predicted many miRNA targets above background false positives.[18] The functional importance of this 'seed rule' was subsequently verified through experimental methods.[7,37,20,115] Other algorithms that incorporated multiple sequence alignments of 3′ UTRs to identify conserved miRNA seed-complementary regions showed improved success in miRNA target finding.[4,5,116–118] Indeed, biochemical data support the conclusion that conservation criteria enhance prediction of functional and effective target sites.[119] However, functional miRNA target sites are not confined to those that are conserved. In fact, some evidence suggests that nonconserved target sites are more prevalent than conserved ones.[120] In response to this, a number of prediction tools have removed or relaxed conservation requirements in their predictions.[10,121–123]

Although conserved seed-pairing requirements helped to improve prediction algorithms, additional contextual features play a role in target recognition. The identification of these features has been accomplished through interplay between experimental and computational approaches. Mining of their own and previous experimental work led to the observation that a 6 nt seed match (miRNA bases 2–7)[5] could be enriched by an extended seed with a complementary site at position 8 (7mer-m8), and/or

an A opposite the miRNA position 1 (7mer-A1), or with an A opposite miRNA position 9.[5,7,18,118,124] The functionality of the 7mer-A1 and 7mer-m8 sites was later confirmed at the translational level by proteomics data.[119,125] Identification of these 7–8 nt seed matches and/or multiple seed-complementary sites in a single UTR improves the specificity of target predictions.[5,7,118,126] These multiple, closely spaced seed-complementary sites on a single mRNA transcript have been shown to act cooperatively, increasing target site efficiency.[7,10,20,72,115,127,128] Using expression data to identify cooperative targeting of coexpressed miRNAs may further improve identification of true miRNA targets.[10,118,125] Some target prediction tools have also incorporated tolerance of a G:U wobble,[77,122,123] though this adds little to their performance over algorithms requiring stringent Watson-Crick base-pairing.[5,119,125] 'Supplemental' and 'compensatory' base-pairing around sites 13–16 of the miRNA has been suggested to complement perfect seed matches or ameliorate targeting when the seed match is not perfect, respectively.[7,10,13,20,129] Though some analyses have associated this pairing with increased target site efficacy,[10] the infrequency of these events and weaker conservation have prevented them from having a significant impact on target prediction algorithms overall.[126] More recently, results from biochemical target-finding approaches have suggested unique base-pairing mechanisms for individual miRNAs and specific to genic regions.[75,130] A class of rare but functional 'centered sites' with base-pairing between miRNA bases 4–14 or 5–15 has been similarly observed.[131] These observations may help to explain some of the confusion surrounding supplemental and compensatory binding. Incorporation of miRNA-specific targeting rules is expected to improve future target prediction tools.

Other considerations such as free energy binding of the miRNA–mRNA duplex, secondary structure accessibility, nucleotide content in and around the putative target site, and position of seed-complementary sites within the mRNA transcript have also been associated with target site efficacy and incorporated into various target prediction algorithms.[4,10,116–118,121,123,124,132–135] Many of these features that have, and will, contribute to miRNA target prediction algorithms were mined from data produced by experimental approaches to genome-wide target finding, as discussed below. Overall, target prediction algorithms have represented a major advancement in the field of miRNA biology. Importantly, by using rules learned from miRNA discoveries in a number of organisms, they enabled identification of novel targets in organisms where no known targets existed. For a complete review of miRNA target prediction algorithms, see a review by Bartel.[126]

EXPERIMENTAL IDENTIFICATION OF miRNA TARGETS

Despite advances in the sensitivity of target prediction algorithms, the overlap of predicted targets between different platforms remains relatively small. Additionally, these *in silico* predictions are not cell-type or context specific. In response to these limitations, *in vivo* and *in vitro* approaches to genome-wide miRNA target discovery have been developed.

Transcriptome and Proteome Profiling for Target Finding

The advent of microarray technology presented an opportunity for researchers to examine the large-scale impact of miRNA regulation and infer direct miRNA targets (Table 1). In 2005 Lim and colleagues showed that HeLa cells transfected with muscle or brain-specific miRNAs, miR-1 or miR-124 respectively, took on mRNA expression patterns characteristic of those tissue types.[37] This change in expression was presumably mediated by these miRNAs, since down-regulated transcripts were enriched for miR-1 or miR-124 seed-matched sites. These results suggested that miRNAs help to define tissue-specific gene expression and corroborated the 'seed rule' as the mechanism by which miRNAs function. Prior to this work, translational repression was thought to be the main regulatory function of miRNAs in animals. However, because significant changes in mRNA levels were observed in response to miRNA overexpression, it seems that many miRNAs have broad effects at the level of RNA stability. Finally, these results were the first biochemical support for the concept initially suggested by *in silico* target predictions that miRNAs may have a large number of targets. A subsequent study in zebrafish provided *in vivo* evidence that each miRNA may directly target hundreds of mRNA transcripts.[43] Furthermore, Linsley and colleagues used complementary gain- and loss-of-function approaches to show that target transcripts can be coordinately regulated by miRNAs with homologous seed regions.[136] The inability of computational algorithms to predict the majority of experimentally identified targets[136] and the high percentage of non-conserved targets found[43] reveal significant shortcomings of *in silico* approaches, in particular those that rely on conservation.

Approaches for perturbation of endogenously expressed miRNAs are an important complement to overexpression studies. Many studies, such as the ones described above, have utilized dicer mutants and knockdowns for loss-of-function miRNA studies.[43,136] Without dicer processing, mature miRNAs cannot be excised from precursor hairpins, blocking miRNAs dependant on this pathway.[25–28] Specific miRNA mutants or knockdown have also proved extremely useful for the study of individual miRNAs.[6,8,137,138] miRNA silencing can be accomplished using engineered antisense oligonucleotides such as 'antagomirs'[137] or locked nucleic acids (LNA),[139] siRNAs,[6] miRNA sponges,[140] and most recently small molecule inhibitors (SMIRS).[141,142] Combined exogenous gain-of-function and endogenous loss-of-function studies have been proved to be both effective methods for identification of the same, physiologically relevant targets, and targeting rules.[10,119,125,128,136] MiRNA inhibitors have also formed the foundation of miRNA-centric therapies, which have progressed quickly from their inception to effective disease treatment in primates in less than 5 years.[137,143]

Although studies documenting transcriptome changes were informative with respect to miRNA regulation and targets, they were incapable of demonstrating the effect of miRNAs on translational regulation. Nakamoto and colleagues used a polyribosome profiling approach to show that transcripts released from translational repression after miR-30a-3p knockdown represented *in vivo* targets of this miRNA.[6] Subsequent works have employed stable isotope labeling by amino acids in cell culture (SILAC) to detect more comprehensive changes in the proteome in response to miRNA regulation.[119,125,144] This approach uses mass spectrometry to measure mass ratios of protein peptides from control cells versus SILAC-labeled cells.[145,146] Using SILAC and miR-1 overexpression, Vinther and colleagues found that targets repressed at the protein level recapitulated seed enrichment observations and largely overlapped with the set of targets identified through transcriptional changes.[37,144] However, some repressed transcripts appeared to be controlled only at the protein level demonstrating the utility of this approach for identification of these translationally repressed targets. By expanding their SILAC-based approaches to thousands, rather than hundreds, of proteins, studies by both the Bartel and Rajewsky groups provided valuable insight into genome-wide translational regulation by miRNAs.[119,125] Their data corroborated much of known miRNA targeting rules that had been suggested by computational and transcriptional studies, but at the protein level. As shown by these previous studies, they also identified functional targeting outside of 3′ UTRs and provided proteome-wide evidence that these coding region sites were less effective in repressing gene expression than 3′ UTR target sites. Many sites identified by these studies were non-conserved,[119] as indicated by previous works, and furthermore these sites had weaker effects on targets than conserved sites overall.[125] Their works showed that while a group of genes is primarily regulated by miRNA at the level of translation, mRNA destabilization seemed to be the prevailing method of regulation in these mammalian models, as supported by subsequent work.[147] Because protein level changes were modest overall, data from these works have contributed to the model that miRNAs fine tune gene expression in the cell.[119,125] This control extended to hundreds of direct targets and thousands of downstream effectors, providing essential evidence that miRNAs have an expansive impact on protein production (Table 1).

IP-Based Approaches for miRNA Target Finding

Instead of looking at gene expression as indirect readouts of miRNA regulation to infer targeting, researchers began to identify ways to capture actual target transcripts. In the cell, the mature miRNA associates closely with the Argonaute protein while it guides the RNA-induced silencing complex (RISC) to target mRNAs. Tenenbaum and colleagues demonstrated that immunoprecipitation (IP) of RNA-binding proteins (RBPs), followed by microarray analysis, could be used for the study of associated RNA transcripts, providing insight into posttranscriptional regulation in a cell-type-specific manner.[148] Many similar methods originally developed for the study of a wide range of RBPs have now been applied to miRNA–ribonucleoprotein complexes (miRNPs).[149–153] Table 2 summarizes the application of these approaches to the study of miRNA targeting. Easow and colleagues used IP of tagged Ago1 protein to demonstrate that this type of approach could be used in an unbiased search for miRNA targets in *Drosophila*.[8] Their work follows the general model of these IP-based studies, that is, targets are identified as transcripts enriched by protein IP over controls in the presence of endogenous or overexpressed miRNAs or depleted in miRNA-deficient samples. The ability to identify miRNA seed-complementary sequences, typically in 3′ UTRs of these transcripts, provides further evidence that these represent *bona fide* miRNA targets.

In these approaches for miRNA target finding, the choice of antibody or protein target, control conditions used for comparison, and readout methods have

TABLE 1 | Gene Expression Array and Proteomics Approaches to miRNA Target Identification

Method	miRNA Expression	Organism	Cell Type	Significant Conclusions	Reference
MiRNA overexpression, measure changes in mRNA levels via microarray	miR-1, miR-124, mutant miR-124 transfection	Human	HeLa	MiRNA targets can be detected in cell lines using microarrays; miRNAs cause wide-spread RNA-level changes; tissue-specific gene expression can be controlled by miRNAs; seed sequences have a functional role in targeting; miRNAs have many targets with diverse functions	Lim et al.[37]
Intravenous antagomir injection, measure tissue-specific mRNA expression via microarray	miR-122 silencing through antagomir	Mouse	Liver tissue	Antagomirs efficiently antagonize function of specific miRNAs; enrichment for seed-matched sites in 3′ UTRs of genes up-regulated in response to miRNA silencing	Krutsfeldt et al.[137]
MiRNA knockdown, use microarrays to measure total mRNA and infer increased translation by detection of transcripts shifted to the heavy polyribosome fraction	miR-30a-3p knockdown with siRNA	Human	HepG2	Detected a shift in target transcripts towards the heavy polyribosome fraction upon miRNA knockdown; some targets had modest mRNA increase upon miRNA knockdown; seed-matched sites in 3′ UTRs and coding regions of targets; poor algorithm prediction of experimentally identified targets	Nakamoto et al.[6]
Compare mRNA expression in embryos lacking miR-430 with WT or miR-430 rescue embryos via microarray analysis	miRNA deficiency; miR-430 rescue; endogenous	Zebrafish	Whole embryo; Dicer mutant	MiRNAs can widely cause deadenylation of mRNAs; a single miRNA can have hundreds of targets; target sites are under rapid evolution; 3′UTRs of regulated transcripts enriched for seed-target sites	Giraldez et al.[43]
MiRNA overexpression, measure protein level changes by SILAC	miR-1 transfection	Human	HeLa	Overlap with transcriptome changes from Lim 2005; some targets only detected by protein level changes; enrichment for seed-matched sites in 3′ UTRs of affected proteins	Vinther et al.[144]
Individual or combination miRNA overexpression or depletion, measure changes in mRNA levels via microarray, monitor the affect on cell cycle progression	Transfection of 24 individual miRNA; miR-16 expression by shRNA and genomic fragment of endogenous locus; miR-16 and miR-106b depletion by anti-miRs	Human	HCT116 Dicer^{ex5}, DLD-1 Dicer^{ex5}, and WT; A549; MCF7; TOV21G; HeLa; SW1417	MiRNAs with the same seed coordinately regulate the same transcripts; gain- and loss-of-function approaches largely identify the same targets; miR-16 miRNAs negatively regulate cell cycle progression; siRNA pools against miR-16 targets can phenocopy the miR-16-induced phenotype; ~60% of targets were missed by computational predictions; miRNA regulation is dependent on seed sequence	Linsley et al.[136]

TABLE 1 | Continued

Method	miRNA Expression	Organism	Cell Type	Significant Conclusions	Reference
MiRNA depletion in a mouse model, measure mRNA expression level differences in CD4+T cells	miR-155 deficiency	Mouse	Th1, Th2 - bic/miRNA-155 mutant	bic/microRNA-155 has key role in immune system homeostasis and function; miR-155 seed-targets enriched in 3′UTRs of targets over other miRNAs; verified c-Maf targeting by miR-155	Rodriguez et al.[138]
MiRNA overexpression, measure mRNA level changes via microarray, computational analyses	Transfection of 11 individual miRNAs	Human	HeLa	Analyses of target features showed site efficacy improved by: positioning away from the stop codon and center of 3′UTRs; local AU enrichment; proximity of targets for coexpressed miRNAs	Grimson et al.[10]
MiRNA overexpression in cell or in vivo miR deficiency, measure protein levels via SILAC and mRNA via microarray	miR-125, miR-1, miR-181 transfection; miR-233 deficient mutant; endogenous	Human Mouse	HeLa; miR-233 knockout mouse neutrophils	Greater regulatory effect of 3′UTR targets versus coding regions; protein repression associated with 7-8mer seed-matched sites; same targeting principles revealed by ectopic and endogenous miRNA expression; widespread evidence of non-conserved target sites; overall prevalence of mRNA destabilization as mechanism of miRNA regulation	Baek et al.[119]
MiRNA overexpression or depletion, measure protein levels by pulse-labeling amino acids at different time points (pSILAC), mRNA changes measured via microarray	miR-1, miR-155, miR-16, miR-30a, let-7b transfection; let-7b LNA depletion	Human	HeLa	Pulse-labeling captures early effects of miRNA regulation on translation; miRNAs regulate expression of hundreds of proteins at a modest level; synergy of multiple seed sequences; most targets repressed at both the mRNA and protein level; miRNA overexpression experiments are physiologically relevant; target sites in 3′UTRs exhibit the greatest regulatory effect	Selbach et al.[125]

been customized by researchers performing the work. Instead of microarray analysis, the Meister group used traditional sequencing methods to identify targets of both Ago1 and Ago2.[154] While they found that these proteins bound largely non-overlapping sets of target mRNAs, subsequent work has suggested the Ago proteins do bind redundant target sets.[71] As seen for array-based approaches, this work emphasized the utility of biochemical approaches in identifying targets missed by prediction algorithms.[154] In addition to Ago1–4, Landthaler and colleagues used IP of associated TNRC6 GW182 family member proteins to identify miRNA targets in human cells.[71] Similarly, Zhang and colleagues targeted GW182 homologs in *C. elegans* to identify known and novel miRNA targets in this organism.[70] Karginov and colleagues also used known *C. elegans* miRNA targets to validate the efficacy of their approach and its reliance on seed-site complementarity to identify target transcripts.[128] In addition to IP enrichment, they measured mRNA expression level changes in response to miR-124a overexpression. Analysis of this data showed that miR-124a seed sequences were enriched not only in down-regulated targets, but also in targets showing no mRNA expression changes. This suggested that the latter group of genes was regulated by translational repression, and this was confirmed for 21 out of 30 of these targets.

A consensus of parallel Ago-IP and transcriptome analyses is that IP datasets are more effective for miRNA target identification than comparing changes in mRNA expression alone.[72,128] Taking these studies

one-step further, Hendrickson and colleagues combined their IP-based method with a polysome profiling approach to estimate translational rate from ribosome density and occupancy.[155] Results of this work agree with previous findings that mRNA decay is the primary cause of protein reduction upon miRNA targeting,[119,125,144] and also suggest that translational repression is quickly followed by mRNA degradation, explaining the correlation between mRNA abundance and translation. Additionally, the inhibition of translational initiation or stimulation of ribosome drop-off near the translation start site is proposed as the primary means of translational regulation by miRNAs. As suggested by computational predictions and previous experimental work,[4–7,9,10,119,125,156] a number of these studies also noted evidence for miRNA targets in coding regions and 5′ UTRs, in addition to 3′ UTRs, implying diverse regulatory roles of miRNAs.[8,72,155]

As was seen with transcriptome and proteome profiling approaches, these IP-based methods were scrutinized for measures predictive of miRNA targeting that could be mined from their data.[157] By combining information from both types of studies, along with their own IP experiments, Hausser and colleagues were able to define features indicative of targeting and those indicative of mRNA degradation.[157] They found the best indicators of targeting were the structural accessibility of the seed-match and target site, and the free energy of miRNA–mRNA hybridization. Characteristics of sites causing mRNA degradation were nucleotide composition in and around the target site, with U content being most predictive of degradation. Secondly, in some, but not all datasets, these sites had more accessible secondary structure. Target sites causing mRNA degradation were more conserved than those with targeting alone, supporting the conclusion that degradation, rather than translational repression, is the main function of miRNAs. Overall, it was concluded that a combination of sequence and structural aspects makes the best model for miRNA target prediction.

CLIP and HT-Sequencing Based miRNA Target Finding

The next-generation of genome-wide biochemical assays for miRNA target finding was ushered in with reports from the Darnell and Pasquinelli/Yeo laboratories.[73,75] Their approaches, and others like them, benefit from several adaptations to the original IP-based methods, allowing for improved target finding. A major disadvantage of traditional methods that relied on co-IP of RBPs with target RNAs is the co-precipitation of unrelated RBPs along with their RNA targets and spurious *in vitro* RBP–RNA interactions.[158] To avoid these problems, the cross-linking immunoprecipitation (CLIP) methodology developed by the Darnell lab exploits the property that UV-irradiation induces covalent bonds between proteins and nucleic acids when contact distances are within angstroms.[150] See Figure 2 for a general description of these approaches.

Through CLIP coupled with high-throughput sequencing (HITS-CLIP or CLIP-seq), the identification of RNA-binding protein target sites is now possible at nucleotide-level resolution.[73,75] This level of resolution of target sites was not possible using hybridization of long transcripts on cDNA arrays. Using sequencing data from mouse brain tissue, Chi and colleagues calculated Ago-CLIP reads that clustered together, representing Argonaute binding sites, and defined an average 'Ago-mRNA footprint'.[73] In support of earlier work, but with greater resolution than previously obtainable, Ago-binding was identified primarily in 3′ UTRs, and also in coding regions, introns, and ncRNAs. Overall, they were able to demonstrate that their approach could identify previously supported miR-124 target sites, especially those most likely to have a significant influence on downregulation of targets at the mRNA and protein levels. Looking at the function of proposed targets, regulatory maps of miR-124, miR-125, and miR-9 corresponded well to what is known about their involvement in neural growth and differentiation.

The significance of improved target site resolution made possible through the use of high-throughput sequencing was emphasized by the identification of new targeting features by the Pasquinelli/Yeo group.[75] Using a similar approach but in whole animals, Zisoulis and colleagues profiled binding sites of the worm ortholog of Argonaute, ALG-1, in *C. elegans*. An advantage of this system was that it allowed Zisoulis and colleagues to capture *in vivo* miRNA regulation in a context-specific manner through cross-linking of developmentally synchronized whole animals. In addition, an *alg-1* mutant strain of worm was available as a negative control, and known miRNA target sites served as positive controls. Known features of miRNA target sites were distinguished in the ALG-1 binding data such as greater conservation, secondary structure accessibility, and enrichment for seed-complementary sequences. Beyond identification of ALG-1 binding outside of 3′ UTRs, they showed coding-exon target sites had significant pairing capacity to the central region of miRNAs and lacked CU rich motifs found in 3′ UTR binding sites. These sequence variations may be linked to the functional difference of

TABLE 2 | Immunoprecipitation-Based Approaches to miRNA Target Identification

Method	miRNA Expression	miRISC Protein Targeted	Organism	Cell Type	Significant Conclusions	Reference
MiRNA overexpression, measure changes in mRNA levels via microarray	Endogenous	Ago1; Ago2	Human	HEK293T	~40% of targets were missed by prediction algorithms; Ago1 and Ago2 bind largely non-overlapping sets of mRNAs	Beitzinger et al.[154]
IP and microarray, enrichment of target mRNAs in FLAG/HA Ago1 transfected cells, miR-1 transfected cells and WT versus *miR-1* mutant embryos	Endogenous; miR-1 transfection; miR-1 deficiency	FLAG/HA tagged Ago1	Drosophila	S2 cells; whole organism; *miR-1* mutant	Significant enrichment for subsets of mRNAs in experimental IPs; evidence of miRNA seed-matched sites enriched in mRNA 3′UTRs; some functional targets in coding regions	Easow et al.[8]
IP and 454 sequencing and microarray, targets identified as mRNAs enriched by IP, miRNA expression determined by sequencing	Endogenous	AIN-1; AIN2 (GW182)	*C. elegans*	Whole organism	Known miRNAs and their targets enriched in AIN-IP datasets; significant overlap between miRNA targets associated with AIN-1 and AIN-2; AIN-IP mRNAs enriched for predicted and validated targets	Zhang et al.[70]
MiRNA expression or depletion, IP and RT-PCR or microarray, total mRNA and IP enriched mRNAs compared via RT-PCR for a subset of known targets and via microarray for global targets	miR-124a transfection; miR-124a antisense oligos; endogenous	c-myc tagged Ago2	Human; mouse	293S; MEF; mouse cortical neurons	Seed sites can act in an additive manner; IP is a more specific and comprehensive predictor of miRNA-mRNA interactions than transcript changes alone	Karginov et al.[128]
MiRNA overexpression, IP or total RNA extraction, microarray detection of enriched target mRNAs; size fractionation and 454 sequencing for miRNA expression	Endogenous; miR-122 transfection	FLAG/HA tagged AGO1-4; FLAG/HA tagged TNRC6A-C (GW182)	Human	HEK239T	TNRC6 proteins are part of the RISC complex; AGO1-4 bind similar sets of mRNAs, and these largely overlap with TNRC6A-C bound transcripts; Seed-match enrichment in target 3′UTRs corresponds to miRNA expression levels and target conservation; evidence of functional redundancy between AGO and TRNRC6 proteins	Landthaler et al.[71]

TABLE 2 | Continued

Method	miRNA Expression	miRISC Protein Targeted	Organism	Cell Type	Significant Conclusions	Reference
MiRNA overexpression, IP, size fractionation for miRNA enrichment, microarray detection of miRNAs and target mRNAs and transcriptome changes	miR-1, miR-124 transfection; endogenous	FLAG tagged Ago2	Human	HEK293T	Ago2 IPs contain representative cellular miRNAs; miR-1 or miR-124 transfection and Ago2 IP return specific and unique target sets; enrichment for longer and multiple seed matches in 3′ UTRs, coding regions and 5′ UTRs; more miR-specific targets found than through transcriptome changes alone; TargetScan is the most effective target prediction algorithm but not as effective as biochemical approaches	Hendrickson et al.[72]
MiRNA overexpression, IP or total mRNA extraction, microarray detection of enriched target mRNAs or transcriptome changes, polysome profiling with microarrays to measure translation rate	miR-124 transfection; endogenous	Ago1-4	Human	HEK293T	Translational profiling is indicative of protein synthesis; change in mRNA abundance is the primary cause of reduced protein expression; miRNAs regulate translational initiation or ribosome drop-off; seed-matches were correlated with, but not required for, reduction in mRNA and translation; miRNA-mediated regulation of mRNA abundance and translation appear related; found common features of translational programs between humans and *S. cerevisiae*	Hendrickson et al.[155]
MiRNA overexpression, IP or total mRNA extraction, microarray detection of enriched target mRNAs or transcriptome changes	Endogenous; miR-124 or miR-7 transfection	FLAG/HA tagged EIF2C2 (Ago2)	Human	HEK239T	mRNA degradation may be the primary action of miRNA targeting; predictive features of targeting are seed-match and target site secondary structure accessibility and free energy of miRNA-mRNA hybridization; predictive features of degradation are nucleotide composition and structural accessibility surrounding the target site; transcriptome datasets are more informative than proteomics data for predictive features of target sites	Hausser et al.[157]

TABLE 2 | Continued

Method	miRNA Expression	miRISC Protein Targeted	Organism	Cell Type	Significant Conclusions	Reference
UV cross-linking, IP and Illumina sequencing of mRNA and miRNA sequences (HITS-CLIP)	Endogenous	Ago1-4	Mouse	P13 whole brain	Nucleotide-level resolution of binding sites; reproducible set of Ago-bound miRNAs; binding coding exons, introns, in and downstream of 3′UTRs and in non-coding RNAs	Chi et al.[73]
UV cross-linking, IP and Illumina sequencing of mRNA and miRNA sequences (CLIP-seq)	Endogenous	ALG-1	C. elegans	Whole organism; *alg-1* mutant	Nucleotide-level resolution of binding sites; most binding sites in coding exons and 3′UTRs, binding sites also in introns and 5′UTRs; reproducible set of Ago-bound miRNAs; CU motif in and around 3′UTR binding sites; distinguishing features of 3′UTR binding sites versus coding exons; specific base-pairing preferences for individual miRNAs; miRNA regulatory genes enriched in targets	Zisoulis et al.[75]
Incubation with photoactivatable ribonucleosides, UV cross-linking, IP and Illumina sequencing of mRNA and miRNA sequences (PAR-CLIP)	Endogenous	FLAG/HA tagged AGO1-4, FLAG/HA tagged TNRC6A-C (GW182)	Human	HEK293T	Nucleotide-level resolution of binding sites, cross-linked nucleotides can be distinguished; PAR-CLIP clusters center on miRNA-mRNA complementary sites; binding in all genic regions, most in coding and 3′UTR exons; some evidence of flexible base-pairing; functional distinction between 3′UTR sites and coding exons sites	Hafner et al.[76]

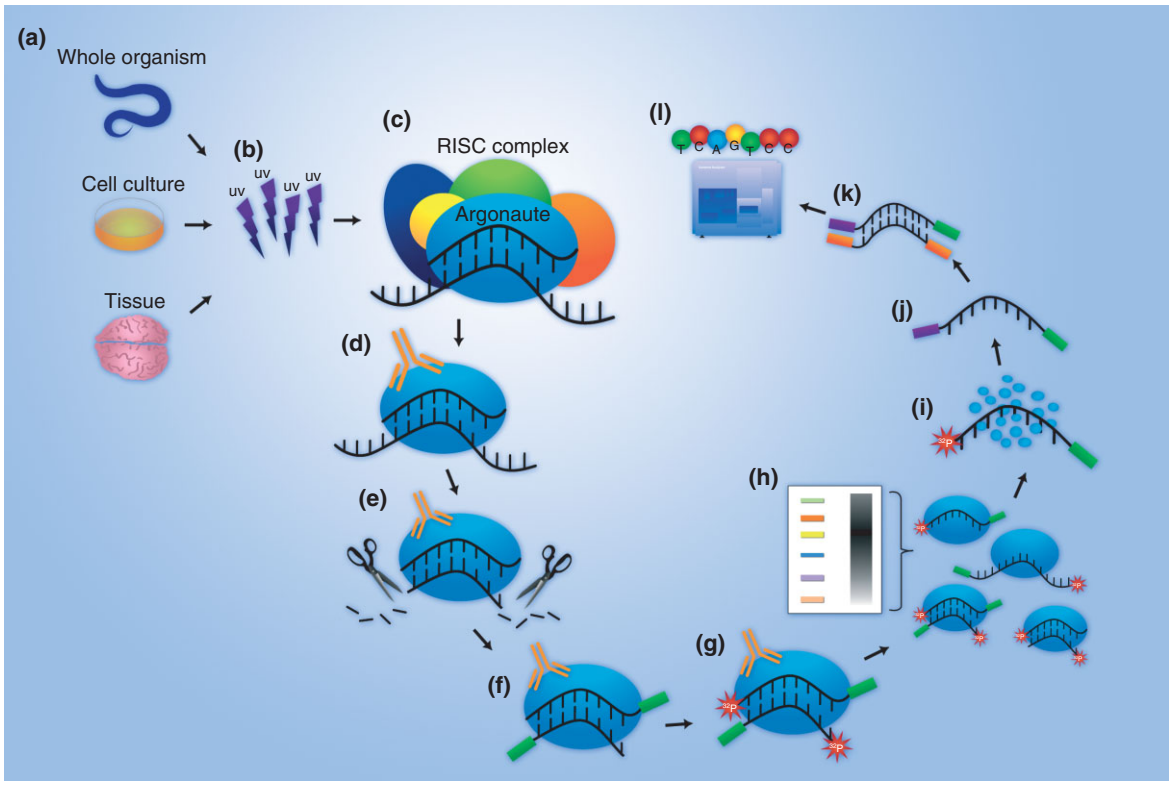


FIGURE 2 | CLIP-seq overview. The CLIP-seq method for identifying miRNA target transcripts can begin with different types of starting material, (a) either whole organisms, cell culture optionally grown in the presence of photoreactable ribonucleosides, or tissue samples. (b) UV-irradiation is used to covalently cross-link proteins with interacting nucleotides. (c) Proteins of the RISC complex associate with the Argonaute protein bound to miRNA and target mRNA. (d) Immunoprecipitation of the Argonaute or other RISC member protein that associates with miRNAs can be used to co-precipitate bound RNA transcripts. (e) Unbound RNA transcripts are degraded, for example with MNase treatment. (f) A 3′ linker for adaptation to the Illumina sequencing platform is ligated to precipitated RNAs. (g) Radiolabeling and (h) SDS-PAGE are used to purify and select RBP–RNA complexes of interest. (i) Proteins bound to RNA transcripts are degraded before (j) ligation of the 5′ sequencing linker and (k) preparation of a cDNA library via RT-PCR and followed by further PCR amplification. (l) Sequencing is typically performed on the Illumina/Solexa system; reads returned correspond to the miRNA and mRNA originally bound in the RISC complex.

3′ UTR target sites noted in this and previous works. Genic location of target sites may also be related to functional category of the target gene, as some distinction in functional categories was seen between groups of genes with 3′ UTR targets versus those with coding-exon targets. Overall, enrichment was seen for targets in miRNA regulatory genes, indicating an autoregulatory function of miRNAs. Evidence was also found for specific seed-pairing rules for individual miRNAs. For example, *lin-4* showed strongest pairing capacity for sites 4–9 and 14–19, rather than canonical seed sites 2–7. This flexibility of miRNA base-pairing may explain the lack of seed matches in a large number of Ago-CLIP clusters identified by the Darnell group.[73] MiRNA-specific targeting rules may be a common occurrence, and as such, will help to improve *in silico* target predictions and our understanding of the *in vivo* role of these miRNAs.[130]

Recently, the Tuschl group added a modification to the cross-linking protocol by incubating cells with the photoactivatable ribonucleoside 4-thiouridine (4SU) prior to UV-irradiation.[76] The effect of this is that cross-linked sites can be distinguished by thymidine to cytidine transitions in the cDNA prepared from extracted RNA. They utilized this approach, termed PAR-CLIP (photoactivatable ribonucleoside-enhanced crosslinking and IP), to improve cross-linking efficiency, remove noise, and refine target sites.

In addition to the use of high-throughput sequencing and cross-linking, next-generation target-finding approaches include a number of other modifications aimed at decreasing noise and further refining target sites. Partial RNA digestion, for example by MNase or RNase T1, has been used to degrade RNA species not protected by the covalently bound protein. This both refines the bound sequence to improve target resolution and generates small fragments of RNA that can be readily sequenced. Separation of RNA–protein complexes by SDS-PAGE, followed by autoradiography visualization allows for extraction

of RNA–protein species of interest and removes noise from non-specific contaminating RNAs.[159] In preparation of the cDNA library for sequencing, choice of primers and PCR conditions must also be considered. Although the Illumina/Solexa platform has been used for these approaches thus far, one can envision how new methods of single-cell, padlock probe, or other undeveloped sequencing approaches may complement these studies in the future. Each of these steps leaves room for optimization depending on starting material, protein target, antibody, sequencing quality, and more.

After sequencing, informatics plays a critical role in these types of approaches. Appropriate algorithms and parameters must be considered when mapping sequencing reads and even the genomic material used in mapping may vary. For example, RNA-seq data can be used to extend annotated UTRs in order to capture an accurate search space for potential binding sites and for choosing appropriate genomic control sequences.[75] Findings of the Darnell group that 8% of their clusters fell within 10 kb downstream of genes suggests that they could have benefited from additional transcriptome information.[73] Once reads have been mapped, custom algorithms are used to distinguish significant clusters of reads from noise as evidence of miRNA–protein binding. Each of the studies described above has applied unique approaches to this problem, using different biological and genomic controls for comparison in cluster finding and/or to extract distinguishing features of binding sites. The choice of appropriate controls in these high-throughput and informatic approaches is still an area of development, and remains an essential consideration for reliable interpretations of data. A shortcoming of all of these approaches is their inability to maintain the association of miRNA–mRNA transcripts as a direct readout. While the search space is greatly reduced by biochemical assays, researchers still use computational approaches to infer these interactions through sequence evidence indicative of miRNA binding, such as seed-complementary sites. The use of unique experimental and computational optimizations and controls by these groups in otherwise similar approaches emphasizes the need for researchers to have a clear understanding of the biological data being analyzed, as well as the strengths and limitations of the molecular and informatic approaches being utilized. One consensus that can be drawn from these studies is that they have provided valuable insight into miRNA-targeting rules and biological roles. These technologies are continuing to evolve and are being applied to unique questions of miRNA biology. More than ever, this combination of biochemical techniques with high-throughput sequencing and bioinformatic analyses emphasizes the synergy of genome-wide approaches that have moved the study of miRNAs forward.

miRNAS IN NORMAL AND DISEASE NETWORKS

As miRNAs have proved to be integral components of regulatory circuits in cells, systems biologists are incorporating miRNAs into known networks of protein–protein interactions, or co-regulated genes. For example, miRNA interactions with transcription factors have already been shown to have important implications for the way regulatory programmes are propagated in the cell.[4,100,103,114,117,138] Several groups have considered this interplay between miRNAs and transcription factors to build more accurate and informative regulatory networks.[160,161] These works have brought to light the role of miRNAs as necessary for maintaining the robustness and stability of cellular systems when they are challenged with environmental stimuli. This critical connection to maintenance of homeostasis may help to explain the implication of miRNAs in disease systems ranging from immunological deficiencies,[138] cancers and tumor metastasis,[162] host–pathogen interactions,[163] and even neurological diseases such as autism spectrum disorders[164] and Alzheimer's disease.[165] Effective diagnosis and therapies for these, and other diseases, may need to combine knowledge of both the protein and miRNA components of these pathways. In disease systems where the cause of misregulation has eluded researchers, the inclusion of miRNAs in network models may provide additional information about protein interconnectivity and even reveal that these transcripts are the missing link in interconnected pathways.[166]

Genetic variation within miRNAs and their target sites can also affect miRNA regulation, directly resulting in disease phenotypes and instances of drug resistance.[167–171] In response, the term miRSNP has been used to describe this class of functional polymorphisms.[172] Computational combinations of polymorphism data and miRNA target predictions have speculated the potential widespread influence of miRSNPs on human disease susceptibility, drug response, and overall phenotypic individuality.[173–176] Through adaptation of the latest biochemical approaches to miRNA target finding to this area of study, future works may be able to identify miRSNPs with greater accuracy and thereby explain the association of certain miRNA-affecting polymorphisms with disease phenotypes. As has been

proposed with protein–protein interaction networks, the use of miRNA–protein interaction networks to add power to genome-wide association studies may similarly allow researchers to identify causal variants that have otherwise remained elusive. These associations would have promising implications for the use of miRNAs as biomarkers and oligonucleotide therapies for disease treatment.[177]

CONCLUSION

The field of miRNA biology has been progressive in its application of genome-wide approaches and analytical methods used to extrapolate the biological roles and regulatory circuits controlled by miRNAs. Improved knowledge of miRNAs and their target networks has, and will continue to, help complete our understanding of the cellular regulatory circuitry and shed light on the functional role of miRNAs in diverse disease models. The implication of miRNAs as global regulators of cell fate specification, disease pathways, and their utility as biomarkers makes these small transcripts an essential component of future biomedical research. In order to deduce the specific networks controlled by individual miRNAs, biochemical target-finding approaches will need to be adapted to allow readout of explicit miRNA–mRNA interactions. Individual miRNA targeting rules will undoubtedly be elucidated through a combination of this experimental work and computational analyses. This information will be essential for the effective and safe use of miRNAs as drug targets, offering an attractive opportunity to control the cooperative actions of multiple genes involved in a single disease pathway. The next generation of miRNA-based therapies will be ushered in by development of improved deliverable and stable synthetic miRNAs and miRNA inhibitors, tools that will also be useful for the continued study of miRNA biology. For these aspirations to be realized, researchers can expect that the synergy between molecular techniques and computational approaches, which has been such a critical element of the miRNA field, will be more essential than ever.

ACKNOWLEDGEMENTS

We thank A. E. Pasquinelli, S. Aigner, J. Nathanson, K. B. Massirer, M. T. Lovci, and T. J. Stark for their comments on the manuscript, and Marcia Wilbert for the figures. We apologize to colleagues whose work has not been cited directly due to space limitations. M.L.W. is supported by the Genetics Training Program at the University of California, San Diego, T32GM008666, and a graduate fellowship from Genentech. This work was also supported by grant GM084317 from the US National Institutes of Health to G.W.Y. and the Stem Cell Program at the University of California, San Diego (G.W.Y.).

REFERENCES

1. Chalfie M, Horvitz HR, Sulston JE. Mutations that lead to reiterations in the cell lineages of *C. elegans*. *Cell* 1981, 24:59–69.
2. 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: 843–854.
3. 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, 75:855–862.
4. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human MicroRNA targets. *PLoS Biol* 2004, 2:e363. doi:10.1371/journal.pbio.0020363.
5. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 2005, 120:15–20. doi:10.1016/j.cell.2004.12.035.
6. Nakamoto M, Jin P, O'Donnell WT, Warren ST. Physiological identification of human transcripts translationally regulated by a specific microRNA. *Hum Mol Genet* 2005, 14:3813–3821. doi:10.1093/hmg/ddi397.
7. Brennecke J, Stark A, Russell RB, Cohen SM. Principles of microRNA-target recognition. *PLoS Biol* 2005, 3:e85. doi:10.1371/journal.pbio.0030085.
8. Easow G, Teleman AA, Cohen SM. Isolation of microRNA targets by miRNP immunopurification. *RNA* 2007, 13:1198–1204. doi:10.1261/rna.563707.
9. Lytle JR, Yario TA, Steitz JA. Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5′ UTR as in the 3′ UTR. *Proc Natl Acad Sci U S A* 2007, 104:9667–9672. doi:10.1073/pnas.0703820104.
10. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* 2007, 27:91–105. doi:10.1016/j.molcel.2007.06.017.

11. Vasudevan S, Tong Y, Steitz JA. Switching from repression to activation: microRNAs can up-regulate translation. *Science* 2007, 318:1931–1934. doi:10.1126/science.1149460.
12. Place RF, Li LC, Pookot D, Noonan EJ, Dahiya R. MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proc Natl Acad Sci U S A* 2008, 105:1608–1613. doi:10.1073/pnas.0707594105.
13. Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* 2000, 403:901–906. doi:10.1038/35002607.
14. Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, Hayward DC, Ball EE, Degnan B, Muller P, et al. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* 2000, 408:86–89. doi:10.1038/35040556.
15. Wightman B, Burglin TR, Gatto J, Arasu P, Ruvkun G. Negative regulatory sequences in the lin-14 3′-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development. *Genes Dev* 1991, 5:1813–1824.
16. Moss EG, Lee RC, Ambros V. The cold shock domain protein LIN-28 controls developmental timing in *C. elegans* and is regulated by the lin-4 RNA. *Cell* 1997, 88:637–646.
17. Slack FJ, Basson M, Liu Z, Ambros V, Horvitz HR, Ruvkun G. The lin-41 RBCC gene acts in the *C. elegans* heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Mol Cell* 2000, 5:659–669.
18. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. *Cell* 2003, 115:787–798.
19. Brennecke J, Hipfner DR, Stark A, Russell RB, Cohen SM. Bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. *Cell* 2003, 113:25–36.
20. Doench JG, Sharp PA. Specificity of microRNA target selection in translational repression. *Genes Dev* 2004, 18:504–511. doi:10.1101/gad.1184404.
21. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. *Science* 2001, 294:853–858. doi:10.1126/science.1064921.
22. Lau NC, Lim LP, Weinstein EG, Bartel DP. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science* 2001, 294:858–862. doi:10.1126/science.1065062.
23. Lee RC, Ambros V. An extensive class of small RNAs in Caenorhabditis elegans. *Science* 2001, 294:862–864. doi:10.1126/science.1065329.
24. Mourelatos Z, Dostie J, Paushkin S, Sharma A, Charroux B, Abel L, Rappsilber J, Mann M, Dreyfuss G. miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev* 2002, 16:720–728. doi:10.1101/gad.974702.
25. Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha I, Baillie DL, Fire A, Ruvkun G, Mello CC. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 2001, 106:23–34.
26. Hutvagner G, McLachlan J, Pasquinelli AE, Balint E, Tuschl T, Zamore PD. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 2001, 293:834–838. doi:10.1126/science.1062961.
27. Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* 2001, 15:2654–2659. doi:10.1101/gad.927801.
28. Knight SW, Bass BL. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. *Science* 2001, 293:2269–2271. doi:10.1126/science.1062039.
29. Hofacker IL. Vienna RNA secondary structure server. *Nucleic Acids Res* 2003, 31:3429–3431.
30. Lim LP, Lau NC, Weinstein EG, Abdelhakim A, Yekta S, Rhoades MW, Burge CB, Bartel DP. The microRNAs of Caenorhabditis elegans. *Genes Dev* 2003, 17:991–1008. doi:10.1101/gad.1074403.
31. Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP. Vertebrate microRNA genes. *Science* 2003, 299:1540. doi:10.1126/science.1080372.
32. Ohler U, Yekta S, Lim LP, Bartel DP, Burge CB. Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification. *RNA* 2004, 10:1309–1322. doi:10.1261/rna.5206304.
33. Lai EC, Tomancak P, Williams RW, Rubin GM. Computational identification of Drosophila microRNA genes. *Genome Biol* 2003, 4:R42. doi:10.1186/gb-2003-4-7-r42.
34. Berezikov E, Guryev V, van de Belt J, Wienholds E, Plasterk RH, Cuppen E. Phylogenetic shadowing and computational identification of human microRNA genes. *Cell* 2005, 120:21–24. doi:10.1016/j.cell.2004.12.031.
35. Grad Y, Aach J, Hayes GD, Reinhart BJ, Church GM, Ruvkun G, Kim J. Computational and experimental identification of *C. elegans* microRNAs. *Mol Cell* 2003, 11:1253–1263.
36. Wang X, Zhang J, Li F, Gu J, He T, Zhang X, Li Y. MicroRNA identification based on sequence and structure alignment. *Bioinformatics* 2005, 21:3610–3614. doi:10.1093/bioinformatics/bti562.

37. Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 2005, 433:769–773. doi:10.1038/nature03315.
38. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009, 19:92–105. doi:10.1101/gr.082701.108.
39. Chan CS, Elemento O, Tavazoie S. Revealing posttranscriptional regulatory elements through network-level conservation. *PLoS Comput Biol* 2005, 1:e69. doi:10.1371/journal.pcbi.0010069.
40. Adai A, Johnson C, Mlotshwa S, Archer-Evans S, Manocha V, Vance V, Sundaresan V. Computational prediction of miRNAs in Arabidopsis thaliana. *Genome Res* 2005, 15:78–91. doi:10.1101/gr.2908205.
41. Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES, Kellis M. Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. *Nature* 2005, 434:338–345. doi:10.1038/nature03441.
42. Berezikov E, Cuppen E, Plasterk RH. Approaches to microRNA discovery. *Nat Genet* 2006, 38:S2–S7. doi:10.1038/ng1794.
43. Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ, Schier AF. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* 2006, 312:75–79. doi:10.1126/science.1122689.
44. Liang H, Li WH. Lowly expressed human microRNA genes evolve rapidly. *Mol Biol Evol* 2009, 26:1195–1198. doi:10.1093/molbev/msp053.
45. Nam JW, Shin KR, Han J, Lee Y, Kim VN, Zhang BT. Human microRNA prediction through a probabilistic co-learning model of sequence and structure. *Nucleic Acids Res* 2005, 33:3570–3581. doi:10.1093/nar/gki668.
46. Hertel J, Stadler PF. Hairpins in a Haystack: recognizing microRNA precursors in comparative genomics data. *Bioinformatics* 2006, 22:e197–e202. doi:10.1093/bioinformatics/btl257.
47. Terai G, Komori T, Asai K, Kin T. miRRim: a novel system to find conserved miRNAs with high sensitivity and specificity. *RNA* 2007, 13:2081–2090. doi:10.1261/rna.655107.
48. Batuwita R, Palade V. microPred: effective classification of pre-miRNAs for human miRNA gene prediction. *Bioinformatics* 2009, 25:989–995. doi:10.1093/bioinformatics/btp107.
49. Sewer A, Paul N, Landgraf P, Aravin A, Pfeffer S, Brownstein MJ, Tuschl T, van Nimwegen E, Zavolan M. Identification of clustered microRNAs using an ab initio prediction method. *BMC Bioinformatics* 2005, 6:267. doi:10.1186/1471-2105-6-267.
50. Xue C, Li F, He T, Liu GP, Li Y, Zhang X. Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine. *BMC Bioinformatics* 2005, 6:310. doi: 10.1186/1471-2105-6-310.
51. Ng KL, Mishra SK. De novo SVM classification of precursor microRNAs from genomic pseudo hairpins using global and intrinsic folding measures. *Bioinformatics* 2007, 23:1321–1330. doi:10.1093/bioinformatics/btm026.
52. Yang Y, Wang YP, Li KB. MiRTif: a support vector machine-based microRNA target interaction filter. *BMC Bioinformatics* 2008, 9(suppl 12):S4. doi:10.1186/1471-2105-9-S12-S4.
53. Chang DT, Wang CC, Chen JW. Using a kernel density estimation based classifier to predict species-specific microRNA precursors. *BMC Bioinformatics* 2008, 9(suppl 12):S2. doi:10.1186/1471-2105-9-S12-S2.
54. Jiang P, Wu H, Wang W, Ma W, Sun X, Lu Z. MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. *Nucleic Acids Res* 2007, 35:W339–344. doi:10.1093/nar/gkm368.
55. Hsieh CH, Chang DT, Hsueh CH, Wu CY, Oyang YJ. Predicting microRNA precursors with a generalized Gaussian components based density estimation algorithm. *BMC Bioinformatics* 2010, 11(suppl 1):S52. doi:10.1186/1471-2105-11-S1-S52.
56. Pfeffer S, Sewer A, Lagos-Quintana M, Sheridan R, Sander C, Grasser FA, van Dyk LF, Ho CK, Shuman S, Chien M, et al. Identification of microRNAs of the herpesvirus family. *Nat Methods* 2005, 2: 269–276. doi:10.1038/nmeth746.
57. Ambros V, Lee RC, Lavanway A, Williams PT, Jewell D. MicroRNAs and other tiny endogenous RNAs in *C. elegans*. *Curr Biol* 2003, 13:807–818.
58. Aravin AA, Lagos-Quintana M, Yalcin A, Zavolan M, Marks D, Snyder B, Gaasterland T, Meyer J, Tuschl T. The small RNA profile during Drosophila melanogaster development. *Dev Cell* 2003, 5: 337–350.
59. Kim J, Krichevsky A, Grad Y, Hayes GD, Kosik KS, Church GM, Ruvkun G. Identification of many microRNAs that copurify with polyribosomes in mammalian neurons. *Proc Natl Acad Sci U S A* 2004, 101:360–365. doi:10.1073/pnas.2333854100.
60. Lu C, Tej SS, Luo S, Haudenschild CD, Meyers BC, Green PJ. Elucidation of the small RNA component of the transcriptome. *Science* 2005, 309:1567–1569. doi:10.1126/science.1114112.
61. Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, Ge H, Bartel DP. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* 2006, 127: 1193–1207. doi:10.1016/j.cell.2006.10.040.

62. Berezikov E, Thuemmler F, van Laake LW, Kondova I, Bontrop R, Cuppen E, Plasterk RH. Diversity of microRNAs in human and chimpanzee brain. *Nat Genet* 2006, 38:1375–1377. doi:10.1038/ng1914.

63. Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, Rajewsky N. Discovering microRNAs from deep sequencing data using miRDeep. *Nat Biotechnol* 2008, 26:407–415. doi:10.1038/nbt1394.

64. Chiang HR, Schoenfeld LW, Ruby JG, Auyeung VC, Spies N, Baek D, Johnston WK, Russ C, Luo S, Babiarz JE, et al. Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev* 2010, 24:992–1009. doi:10.1101/gad.1884710.

65. Lu YC, Smielewska M, Palakodeti D, Lovci MT, Aigner S, Yeo GW, Graveley BR. Deep sequencing identifies new and regulated microRNAs in Schmidtea mediterranea. *RNA* 2009, 15:1483–1491. doi:10.1261/rna.1702009.

66. Friedlander MR, Adamidi C, Han T, Lebedeva S, Isenbarger TA, Hirst M, Marra M, Nusbaum C, Lee WL, Jenkin JC, et al. High-resolution profiling and discovery of planarian small RNAs. *Proc Natl Acad Sci U S A* 2009, 106:11546–11551. doi:10.1073/pnas.0905222106.

67. Hackenberg M, Sturm M, Langenberger D, Falcon-Perez JM, Aransay AM. miRanalyzer: a microRNA detection and analysis tool for next-generation sequencing experiments. *Nucleic Acids Res* 2009, 37:W68–W76. doi:10.1093/nar/gkp347.

68. Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M, et al. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res* 2008, 18:610–621. doi:10.1101/gr.7179508.

69. Stark MS, Tyagi S, Nancarrow DJ, Boyle GM, Cook AL, Whiteman DC, Parsons PG, Schmidt C, Sturm RA, Hayward NK. Characterization of the Melanoma miRNAome by Deep Sequencing. *PLoS One* 2010, 5:e9685. doi:10.1371/journal.pone.0009685.

70. Zhang L, Ding L, Cheung TH, Dong MQ, Chen J, Sewell AK, Liu X, Yates JR III Han M. Systematic identification of *C. elegans* miRISC proteins, miRNAs, and mRNA targets by their interactions with GW182 proteins AIN-1 and AIN-2. *Mol Cell* 2007, 28:598–613. doi:10.1016/j.molcel.2007.09.014.

71. Landthaler M, Gaidatzis D, Rothballer A, Chen PY, Soll SJ, Dinic L, Ojo T, Hafner M, Zavolan M, Tuschl T. Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. *RNA* 2008, 14:2580–2596. doi:10.1261/rna.1351608.

72. Hendrickson DG, Hogan DJ, Herschlag D, Ferrell JE, Brown PO. Systematic identification of mRNAs recruited to argonaute 2 by specific microRNAs and corresponding changes in transcript abundance. *PLoS One* 2008, 3:e2126. doi:10.1371/journal.pone.0002126.

73. Chi SW, Zang JB, Mele A, Darnell RB. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 2009, 460:479–486. doi:10.1038/nature08170.

74. Goff LA, Davila J, Swerdel MR, Moore JC, Cohen RI, Wu H, Sun YE, Hart RP. Ago2 immunoprecipitation identifies predicted microRNAs in human embryonic stem cells and neural precursors. *PLoS One* 2009, 4:e7192. doi:10.1371/journal.pone.0007192.

75. Zisoulis DG, Lovci MT, Wilbert ML, Hutt KR, Liang TY, Pasquinelli AE, Yeo GW. Comprehensive discovery of endogenous Argonaute binding sites in Caenorhabditis elegans. *Nat Struct Mol Biol* 2010, 17:173–179. doi:10.1038/nsmb.1745.

76. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A. Ascano M Jr, Jungkamp AC, Munschauer M. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 2010, 141:129–141. doi:10.1016/j.cell.2010.03.009.

77. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. miRBase: tools for microRNA genomics. *Nucleic Acids Res* 2008, 36:D154–D158. doi:10.1093/nar/gkm952.

78. Griffiths-Jones S. miRBase: microRNA sequences and annotation. *Curr Protoc Bioinform* 2010, 9:1–10. doi:10.1002/0471250953.bi1209s29.

79. Skalsky RL, Vanlandingham DL, Scholle F, Higgs S, Cullen BR. Identification of microRNAs expressed in two mosquito vectors, Aedes albopictus and Culex quinquefasciatus. *BMC Genomics* 2010, 11:119. doi:10.1186/1471-2164-11-119.

80. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 2008, 5:621–628. doi:10.1038/nmeth.1226.

81. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A, Kamphorst AO, Landthaler M, et al. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* 2007, 129:1401–1414. doi:10.1016/j.cell.2007.04.040.

82. Lui WO, Pourmand N, Patterson BK, Fire A. Patterns of known and novel small RNAs in human cervical cancer. *Cancer Res* 2007, 67:6031–6043. doi:10.1158/0008-5472.CAN-06-0561.

83. Tarasov V, Jung P, Verdoodt B, Lodygin D, Epanchintsev A, Menssen A, Meister G, Hermeking H. Differential regulation of microRNAs by p53 revealed by massively parallel sequencing: miR-34a is a p53 target that induces apoptosis and G1-arrest. *Cell Cycle* 2007, 6:1586–1593.

84. Kuchenbauer F, Morin RD, Argiropoulos B, Petriv OI, Griffith M, Heuser M, Yung E, Piper J, Delaney A,

Prabhu AL, et al. In-depth characterization of the microRNA transcriptome in a leukemia progression model. *Genome Res* 2008, 18:1787–1797. doi:10.1101/gr.077578.108.

85. Nygaard S, Jacobsen A, Lindow M, Eriksen J, Balslev E, Flyger H, Tolstrup N, Moller S, Krogh A, Litman T. Identification and analysis of miRNAs in human breast cancer and teratoma samples using deep sequencing. *BMC Med Genomics* 2009, 2:35. doi:10.1186/1755-8794-2-35.
86. Isik M, Korswagen HC, Berezikov E. Expression patterns of intronic microRNAs in Caenorhabditis elegans. *Silence* 2010, 1:5. doi:10.1186/1758-907X-1-5.
87. Remenyi J, Hunter CJ, Cole C, Ando H, Impey S, Monk CE, Martin KJ, Barton GJ, Hutvagner G, Arthur JS. Regulation of the miR-212/132 locus by MSK and CREB in response to neurotrophins. *Biochem J* 2010, 428:281–291. doi:10.1042/BJ20100024.
88. Shao NY, Hu HY, Yan Z, Xu Y, Hu H, Menzel C, Li N, Chen W, Khaitovich P. Comprehensive survey of human brain microRNA by deep sequencing. *BMC Genomics* 2010, 11:409. doi:10.1186/1471-2164-11-409.
89. Umbach JL, Cullen BR. In-depth analysis of Kaposi's sarcoma-associated herpesvirus microRNA expression provides insights into the mammalian microRNA-processing machinery. *J Virol* 2010, 84:695–703. doi:10.1128/JVI.02013-09.
90. Houbaviy HB, Murray MF, Sharp PA. Embryonic stem cell-specific MicroRNAs. *Dev Cell* 2003, 5: 351–358.
91. Suh MR, Lee Y, Kim JY, Kim SK, Moon SH, Lee JY, Cha KY, Chung HM, Yoon HS, Moon SY, et al. Human embryonic stem cells express a unique set of microRNAs. *Dev Biol* 2004, 270:488–498. doi:10.1016/j.ydbio.2004.02.019.
92. Wilson KD, Venkatasubrahmanyam S, Jia F, Sun N, Butte AJ, Wu JC. MicroRNA profiling of human-induced pluripotent stem cells. *Stem Cells Dev* 2009, 18:749–758. doi:10.1089/scd.2008.0247.
93. Calabrese JM, Seila AC, Yeo GW, Sharp PA. RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. *Proc Natl Acad Sci U S A* 2007, 104:18097–18102. doi:10.1073/pnas.0709193104.
94. Bar M, Wyman SK, Fritz BR, Qi J, Garg KS, Parkin RK, Kroh EM, Bendoraite A, Mitchell PS, Nelson AM, et al. MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. *Stem Cells* 2008, 26:2496–2505. doi:10.1634/stemcells.2008-0356.
95. Laurent LC, Chen J, Ulitsky I, Mueller FJ, Lu C, Shamir R, Fan JB, Loring JF. Comprehensive microRNA profiling reveals a unique human embryonic stem cell signature dominated by a single seed sequence. *Stem Cells* 2008, 26:1506–1516. doi:10.1634/stemcells.2007-1081.
96. Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* 2008, 22: 2773–2785. doi:10.1101/gad.1705308.
97. Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* 2008, 455:1124–1128. doi:10.1038/nature07299.
98. Tay YM, Tam WL, Ang YS, Gaughwin PM, Yang H, Wang W, Liu R, George J, Ng HH, Perera RJ, et al. MicroRNA-134 modulates the differentiation of mouse embryonic stem cells, where it causes post-transcriptional attenuation of Nanog and LRH1. *Stem Cells* 2008, 26:17–29. doi:10.1634/stemcells.2007-0295.
99. Lin SL, Chang DC, Chang-Lin S, Lin CH, Wu DT, Chen DT, Ying SY. Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state. *RNA* 2008, 14:2115–2124. doi:10.1261/rna.1162708.
100. Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T, Johnstone S, Guenther MG, Johnston WK, Wernig M, Newman J, et al. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 2008, 134: 521–533. doi:10.1016/j.cell.2008.07.020.
101. Xu N, Papagiannakopoulos T, Pan G, Thomson JA, Kosik KS. MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell* 2009, 137:647–658. doi:10.1016/j.cell.2009.02.038.
102. Lin CH, Jackson AL, Guo J, Linsley PS, Eisenman RN. Myc-regulated microRNAs attenuate embryonic stem cell differentiation. *EMBO J* 2009, 28: 3157–3170. doi:10.1038/emboj.2009.254.
103. Cordes KR, Srivastava D. MicroRNA regulation of cardiovascular development. *Circ Res* 2009, 104: 724–732. doi:10.1161/CIRCRESAHA.108.192872.
104. Judson RL, Babiarz JE, Venere M, Blelloch R. Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nat Biotechnol* 2009, 27:459–461. doi:10.1038/nbt.1535.
105. Stadler BM, Ivanovska I, Mehta K, Song S, Nelson A, Tan Y, Mathieu J, Darby GC, Blau CA, Ware C, et al. Characterization of microRNAs Involved in Embryonic Stem Cell States. *Stem Cells Dev* 2010, 19:935–950. doi:10.1089/scd.2009.0426.
106. Melton C, Judson RL, Blelloch R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature* 2010, 463:621–626. doi:10.1038/nature08725.
107. Mallanna SK, Rizzino A. Emerging roles of microRNAs in the control of embryonic stem cells and the generation of induced pluripotent stem cells. *Dev Biol* 2010, 344:16–25. doi:10.1016/j.ydbio.2010.05.014.

108. Luciano DJ, Mirsky H, Vendetti NJ, Maas S. RNA editing of a miRNA precursor. *RNA* 2004, 10: 1174–1177. doi:10.1261/rna.7350304.

109. Aravin A, Tuschl T. Identification and characterization of small RNAs involved in RNA silencing. *FEBS Lett* 2005, 579:5830–5840. doi:10.1016/j.febslet.2005.08.009.

110. Reid JG, Nagaraja AK, Lynn FC, Drabek RB, Muzny DM, Shaw CA, Weiss MK, Naghavi AO, Khan M, Zhu H, et al. Mouse let-7 miRNA populations exhibit RNA editing that is constrained in the 5′-seed/ cleavage/anchor regions and stabilize predicted mmu-let-7a:mRNA duplexes. *Genome Res* 2008, 18:1571–1581. doi:10.1101/gr.078246.108.

111. Ebhardt HA, Tsang HH, Dai DC, Liu Y, Bostan B, Fahlman RP. Meta-analysis of small RNA-sequencing errors reveals ubiquitous post-transcriptional RNA modifications. *Nucleic Acids Res* 2009, 37: 2461–2470. doi:10.1093/nar/gkp093.

112. Kawahara Y, Zinshteyn B, Sethupathy P, Iizasa H, Hatzigeorgiou AG, Nishikura K. Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* 2007, 315:1137–1140. doi:10.1126/science.1138050.

113. Pantano L, Estivill X, Marti E. SeqBuster, a bioinformatic tool for the processing and analysis of small RNAs datasets, reveals ubiquitous miRNA modifications in human embryonic cells. *Nucleic Acids Res* 2010, 38:e34. doi:10.1093/nar/gkp1127.

114. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. Prediction of plant microRNA targets. *Cell* 2002, 110:513–520.

115. Lai EC, Tam B, Rubin GM. Pervasive regulation of Drosophila Notch target genes by GY-box-, Brd-box-, and K-box-class microRNAs. *Genes Dev* 2005, 19:1067–1080. doi:10.1101/gad.1291905.

116. Stark A, Brennecke J, Russell RB, Cohen SM. Identification of Drosophila microRNA targets. *PLoS Biol* 2003, 1:E60. doi:10.1371/journal.pbio.0000060.

117. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. *Genome Biol* 2003, 5:doi:10.1186/gb-2003-5-1-r1.:R1.

118. Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, et al. Combinatorial microRNA target predictions. *Nat Genet* 2005, 37:495–500. doi:10.1038/ng1536.

119. Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP. The impact of microRNAs on protein output. *Nature* 2008, 455:64–71. doi:10.1038/nature07242.

120. Farh KK, Grimson A, Jan C, Lewis BP, Johnston WK, Lim LP, Burge CB, Bartel DP. The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science* 2005, 310:1817–1821. doi:10.1126/science.1121158.

121. Robins H, Li Y, Padgett RW. Incorporating structure to predict microRNA targets. *Proc Natl Acad Sci U S A* 2005, 102:4006–4009. doi:10.1073/pnas.0500775102.

122. Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, Lim B, Rigoutsos I. A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* 2006, 126:1203–1217. doi:10.1016/j.cell.2006.07.031.

123. Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. The role of site accessibility in microRNA target recognition. *Nat Genet* 2007, 39:1278–1284. doi:10.1038/ng2135.

124. Nielsen CB, Shomron N, Sandberg R, Hornstein E, Kitzman J, Burge CB. Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA* 2007, 13:1894–1910. doi:10.1261/rna.768207.

125. Selbach M, Schwanhausser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N. Widespread changes in protein synthesis induced by microRNAs. *Nature* 2008, 455:58–63. doi:10.1038/nature07228.

126. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell* 2009, 136:215–233. doi:10.1016/j.cell.2009.01.002.

127. Vella MC, Choi EY, Lin SY, Reinert K, Slack FJ. The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3′UTR. *Genes Dev* 2004, 18:132–137. doi:10.1101/gad.1165404.

128. Karginov FV, Conaco C, Xuan Z, Schmidt BH, Parker JS, Mandel G, Hannon GJ. A biochemical approach to identifying microRNA targets. *Proc Natl Acad Sci U S A* 2007, 104:19291–19296. doi:10.1073/pnas.0709971104.

129. Yekta S, Shih IH, Bartel DP. MicroRNA-directed cleavage of HOXB8 mRNA. *Science* 2004, 304: 594–596. doi:10.1126/science.1097434.

130. Wang WX, Wilfred BR, Xie K, Jennings MH, Hu Y, Stromberg AJ, Nelson PT. Individual microRNAs (miRNAs) display distinct mRNA targeting "rules". *RNA Biol* 2010, 7:373–380.

131. Shin C, Nam JW, Farh KK, Chiang HR, Shkumatava A, Bartel DP. Expanding the microRNA targeting code: functional sites with centered pairing. *Mol Cell* 2010, 38:789–802. doi:10.1016/j.molcel.2010.06.005.

132. Lall S, Grun D, Krek A, Chen K, Wang YL, Dewey CN, Sood P, Colombo T, Bray N, Macmenamin P, et al. A genome-wide map of conserved microRNA targets in *C. elegans*. *Curr Biol* 2006, 16:460–471. doi:10.1016/j.cub.2006.01.050.

133. Zhao Y, Samal E, Srivastava D. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* 2005, 436:214–220. doi:10.1038/nature03817.

134. Long D, Lee R, Williams P, Chan CY, Ambros V, Ding Y. Potent effect of target structure on microRNA function. *Nat Struct Mol Biol* 2007, 14:287–294. doi:10.1038/nsmb1226.

135. Hammell M, Long D, Zhang L, Lee A, Carmack CS, Han M, Ding Y, Ambros V. mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein-enriched transcripts. *Nat Methods* 2008, 5:813–819. doi:10.1038/nmeth.1247.

136. Linsley PS, Schelter J, Burchard J, Kibukawa M, Martin MM, Bartz SR, Johnson JM, Cummins JM, Raymond CK, Dai H, et al. Transcripts targeted by the microRNA-16 family cooperatively regulate cell cycle progression. *Mol Cell Biol* 2007, 27:2240–2252. doi:10.1128/MCB.02005-06.

137. Krutzfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, Manoharan M, Stoffel M. Silencing of microRNAs in vivo with 'antagomirs'. *Nature* 2005, 438:685–689. doi:10.1038/nature04303.

138. Rodriguez A, Vigorito E, Clare S, Warren MV, Couttet P, Soond DR, van Dongen S, Grocock RJ, Das PP, Miska EA, et al. Requirement of bic/microRNA-155 for normal immune function. *Science* 2007, 316:608–611. doi:10.1126/science.1139253.

139. Orom UA, Kauppinen S, Lund AH. LNA-modified oligonucleotides mediate specific inhibition of microRNA function. *Gene* 2006, 372:137–141. doi:10.1016/j.gene.2005.12.031.

140. Ebert MS, Neilson JR, Sharp PA. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nat Methods* 2007, 4:721–726. doi:10.1038/nmeth1079.

141. Thomas JR, Hergenrother PJ. Targeting RNA with small molecules. *Chem Rev* 2008, 108:1171–1224. doi:10.1021/cr0681546.

142. Zhang S, Chen L, Jung EJ, Calin GA. Targeting microRNAs with small molecules: from dream to reality. *Clin Pharmacol Ther* 2010, 87:754–758. doi:10.1038/clpt.2010.46.

143. Lanford RE, Hildebrandt-Eriksen ES, Petri A, Persson R, Lindow M, Munk ME, Kauppinen S, Orum H. Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. *Science* 2010, 327:198–201. doi:10.1126/science.1178178.

144. Vinther J, Hedegaard MM, Gardner PP, Andersen JS, Arctander P. Identification of miRNA targets with stable isotope labeling by amino acids in cell culture. *Nucleic Acids Res* 2006, 34:e107. doi:10.1093/nar/gkl590.:e107.

145. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002, 1:376–386.

146. Zhu H, Pan S, Gu S, Bradbury EM, Chen X. Amino acid residue specific stable isotope labeling for quantitative proteomics. *Rapid Commun Mass Spectrom* 2002, 16:2115–2123. doi:10.1002/rcm.831.

147. Guo H, Ingolia NT, Weissman JS, Bartel DP. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 2010, 466:835–840. doi:10.1038/nature09267.

148. Tenenbaum SA, Carson CC, Lager PJ, Keene JD. Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. *Proc Natl Acad Sci U S A* 2000, 97:14085–14090. doi:10.1073/pnas.97.26.14085.

149. Brown V, Jin P, Ceman S, Darnell JC, O'Donnell WT, Tenenbaum SA, Jin X, Feng Y, Wilkinson KD, Keene JD, et al. Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* 2001, 107:477–487.

150. Ule J, Jensen KB, Ruggiu M, Mele A, Ule A, Darnell RB. CLIP identifies Nova-regulated RNA networks in the brain. *Science* 2003, 302:1212–1215. doi:10.1126/science.1090095.

151. Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 2008, 456:\break 464–469. doi:10.1038/nature07488.

152. Yeo GW, Coufal NG, Liang TY, Peng GE, Fu XD, Gage FH. An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. *Nat Struct Mol Biol* 2009, 16:130–137. doi:10.1038/nsmb.1545.

153. Sanford JR, Wang X, Mort M, Vanduyn N, Cooper DN, Mooney SD, Edenberg HJ, Liu Y. Splicing factor SFRS1 recognizes a functionally diverse landscape of RNA transcripts. *Genome Res* 2009, 19:381–394. doi:10.1101/gr.082503.108.

154. Beitzinger M, Peters L, Zhu JY, Kremmer E, Meister G. Identification of human microRNA targets from isolated argonaute protein complexes. *RNA Biol* 2007, 4:76–84.

155. Hendrickson DG, Hogan DJ, McCullough HL, Myers JW, Herschlag D, Ferrell JE, Brown PO. Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. *PLoS Biol* 2009, 7:e1000238. doi:10.1371/journal.pbio.1000238.

156. Stark A, Lin MF, Kheradpour P, Pedersen JS, Parts L, Carlson JW, Crosby MA, Rasmussen MD, Roy S, Deoras AN, et al. Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. *Nature* 2007, 450:219–232. doi:10.1038/nature06340.

157. Hausser J, Landthaler M, Jaskiewicz L, Gaidatzis D, Zavolan M. Relative contribution of sequence and

structure features to the mRNA binding of Argonaute/EIF2C-miRNA complexes and the degradation of miRNA targets. *Genome Res* 2009, 19:2009–2020. doi:10.1101/gr.091181.109.

158. Mili S, Steitz JA. Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* 2004, 10:1692–1694. doi:10.1261/rna.7151404.
159. Ule J, Jensen K, Mele A, Darnell RB. CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* 2005, 37:376–386. doi:10.1016/j.ymeth.2005.07.018.
160. Yu X, Lin J, Zack DJ, Mendell JT, Qian J. Analysis of regulatory network topology reveals functionally distinct classes of microRNAs. *Nucleic Acids Res* 2008, 36:6494–6503. doi:10.1093/nar/gkn712.
161. Tu K, Yu H, Hua YJ, Li YY, Liu L, Xie L, Li YX. Combinatorial network of primary and secondary microRNA-driven regulatory mechanisms. *Nucleic Acids Res* 2009, 37:5969–5980. doi:10.1093/nar/gkp638.
162. Zhang H, Li Y, Lai M. The microRNA network and tumor metastasis. *Oncogene* 2010, 29:937–948. doi:10.1038/onc.2009.406.
163. Cullen BR. Viral and cellular messenger RNA targets of viral microRNAs. *Nature* 2009, 457:421–425. doi:10.1038/nature07757.
164. Sarachana T, Zhou R, Chen G, Manji HK, Hu VW. Investigation of post-transcriptional gene regulatory networks associated with autism spectrum disorders by microRNA expression profiling of lymphoblastoid cell lines. *Genome Med* 2010, 2:23. doi:10.1186/gm144.
165. Ho L, Fivecoat H, Wang J, Pasinetti GM. Alzheimer's disease biomarker discovery in symptomatic and asymptomatic patients: experimental approaches and future clinical applications. *Exp Gerontol* 2010, 45:15–22. doi:10.1016/j.exger.2009.09.007.
166. Liang H, Li WH. MicroRNA regulation of human protein protein interaction network. *RNA* 2007, 13:1402–1408. doi:10.1261/rna.634607.
167. Iwai N, Naraba H. Polymorphisms in human pre-miRNAs. *Biochem Biophys Res Commun* 2005, 331:1439–1444. doi:10.1016/j.bbrc.2005.04.051.
168. Abelson JF, Kwan KY, O'Roak BJ, Baek DY, Stillman AA, Morgan TM, Mathews CA, Pauls DL, Rasin MR, Gunel M, et al. Sequence variants in SLITRK1 are associated with Tourette's syndrome. *Science* 2005, 310:317–320. doi:10.1126/science.1116502.
169. Mayr C, Hemann MT, Bartel DP. Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science* 2007, 315:1576–1579. doi:10.1126/science.1137999.
170. Mishra PJ, Humeniuk R, Mishra PJ, Longo-Sorbello GS, Banerjee D, Bertino JR. A miR-24 microRNA binding-site polymorphism in dihydrofolate reductase gene leads to methotrexate resistance. *Proc Natl Acad Sci U S A* 2007, 104:13513–13518. doi:10.1073/pnas.0706217104.
171. Nicoloso MS, Sun H, Spizzo R, Kim H, Wickramasinghe P, Shimizu M, Wojcik SE, Ferdin J, Kunej T, Xiao L, et al. Single-nucleotide polymorphisms inside microRNA target sites influence tumor susceptibility. *Cancer Res* 2010, 70:2789–2798. doi:10.1158/0008-5472.CAN-09-3541.
172. Mishra PJ, Mishra PJ, Banerjee D, Bertino JR. MiRSNPs or MiR-polymorphisms, new players in microRNA mediated regulation of the cell: Introducing microRNA pharmacogenomics. *Cell Cycle* 2008, 7:853–858.
173. Saunders MA, Liang H, Li WH. Human polymorphism at microRNAs and microRNA target sites. *Proc Natl Acad Sci U S A* 2007, 104:3300–3305. doi:10.1073/pnas.0611347104.
174. Yu Z, Li Z, Jolicoeur N, Zhang L, Fortin Y, Wang E, Wu M, Shen SH. Aberrant allele frequencies of the SNPs located in microRNA target sites are potentially associated with human cancers. *Nucleic Acids Res* 2007, 35:4535–4541. doi:10.1093/nar/gkm480.
175. Xu Z, Taylor JA. SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. *Nucleic Acids Res* 2009, 37:W600–W605. doi:10.1093/nar/gkp290.
176. Sun G, Yan J, Noltner K, Feng J, Li H, Sarkis DA, Sommer SS, Rossi JJ. SNPs in human miRNA genes affect biogenesis and function. *RNA* 2009, 15:1640–1651. doi:10.1261/rna.1560209.
177. Saunders MA, Lim LP. (micro)Genomic medicine: microRNAs as therapeutics and biomarkers. *RNA Biol* 2009, 6:324–328.