nature genetics

microRNA target predictions in animals

Nikolaus Rajewsky

In recent years, microRNAs (miRNAs) have emerged as a major class of regulatory genes, present in most metazoans and important for a diverse range of biological functions. Because experimental identification of miRNA targets is difficult, there has been an explosion of computational target predictions. Although the initial round of predictions resulted in very diverse results, subsequent computational and experimental analyses suggested that at least a certain class of conserved miRNA targets can be confidently predicted and that this class of targets is large, covering, for example, at least 30% of all human genes when considering about 60 conserved vertebrate miRNA gene families. Most recent approaches have also shown that there are correlations between domains of miRNA expression and mRNA levels of their targets. Our understanding of miRNA function is still extremely limited, but it may be that by integrating mRNA and miRNA sequence and expression data with other comparative genomic data, we will be able to gain global and yet specific insights into the function and evolution of a broad layer of post-transcriptional control.

The first round of miRNA target predictions in animals

In 2003, it was shown that the fly miRNA *bantam* targets and negatively regulates the pro-apoptotic gene *hid*[1,2]. Using genetic approaches, other miRNA targets had been found in *Caenorhabditis elegans* before, but *hid* was the first target identified by performing a genome-wide, sequence-based bioinformatic screen for targets of a miRNA. Several other groups published target predictions in *Drosophila melanogaster*[3,4] and vertebrates[5–7]. These methods were based on (i) a very limited set of experimentally identified likely target sites for very few miRNAs (roughly 20 sites in *C. elegans* for two miRNAs, all located in the 3′ untranslated regions (UTRs) of mRNAs); (ii) the initial observation of Eric Lai that known posttranscriptional regulatory motifs in 3′ UTRs, such as the K-box, are perfectly complementary to the 5′ end of some fly miRNAs[8]; and (iii) various *in vitro* experiments, including the result that the multiplicity of sites in a 3′ UTR seemed to exponentially boost the efficacy of target repression (reviewed in ref. 9). All methods scored not only complementarity of the miRNA to the target site but also evaluated the predicted free energy of the miRNA/mRNA duplex. This seemed natural, because (i) the perhaps best-validated miRNA target sites (located in the nematode gene *lin-41* and targeted by *let-7*) have extensive complementarity to the miRNA; and (ii) several entire mature miRNA sequences have been almost perfectly conserved over hundreds of million years. However, it was recognized that the most significant contribution to target recognition seemed to come from short stretches (6–8 bp) of consecutive, perfect Watson-Crick miRNA-mRNA complementarity[2,4,7]. Using elegant bioinformatic methods, it was found that these perfect matches are typically located in the 5′ end of the miRNA (called 'seed' sites[7]; see also ref. 2). The term 'nucleus' was used for 6- to 8-bp consecutive perfect (Watson-Crick) base pairings between miRNA and target site that were found to be the key component of target recognition and were often found in the 5′ end of the miRNAs but did not *a priori* restrict the position of a nucleus to the 5′ end of the miRNA[4]. It was also put forward that these nuclei were initiating a rapid zip up of the miRNA/mRNA duplex to overcome thermal diffusion, followed by a stabilizing thermodynamic step of further annealing of the miRNA to the target site[4]. This result indicated that the free energy of the entire miRNA/mRNA duplex is generally a bad predictor for miRNA target sites. Most of the proposed methods used some way of assessing the statistical significance of predicted targets, for example by randomizing 3′ UTR sequences[4], but the perhaps most convincing test used random sequences ('mock' miRNAs) as controls[7]. Mock miRNAs were chosen such that the abundance of perfect matches to their 5′ end sequences in 3′ UTRs of a single genome was comparable to true miRNAs. The central argument was that these mock sequences were unlikely to be biologically relevant and that thus the observation that there were many more conserved 'hits' for real miRNAs than for mock miRNAs would indicate that many of the hits for real miRNAs were indeed biologically relevant. The ratio of 'real' versus 'mock' hits then provided an estimate of the signal-to-noise ratio of the target predictions. This argument has been widely accepted; however, it is worth pointing out that if a miRNA has only very few targets because having a target site might be functionally disadvantageous for a mRNA, then the number of its predicted conserved sites could be fewer or similar to the number of hits produced by 'mock' miRNAs.

Although the general nature of all methods were related (for a more detailed comparison see ref. 10), animal miRNA target sites are small and have only limited complementarity to their target sites, and thus even small differences in the algorithms produced a great diversity in target predictions. Apart from differences in the algorithms, it is important to note that it is not trivial to define 3′ UTR sequences. Although many non-redundant full length cDNAs have been

Center for Comparative Functional Genomics Department of Biology, 100 Washington Square East, New York, New York 10003, USA
e-mail: nikolaus.rajewsky@nyu.edu

Published online 30 May 2006; doi:10.1038/ng1798

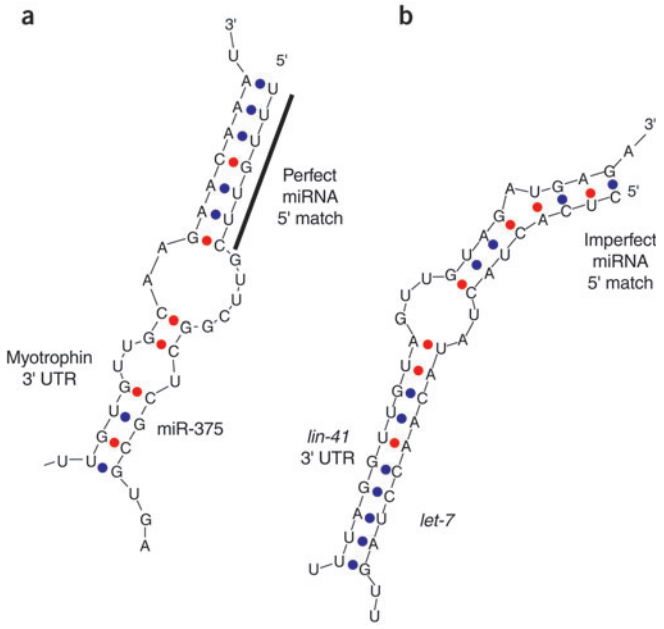

Figure 1 Two classes of miRNA target sites in animals. (**a**) Class 1 targets have perfect, consecutive Watson-Crick base pairings between the 5′ end of the miRNA and the 3′ UTR target sites but insignificant complementarity in the remainder of the miRNA sequence. Shown is a portion of the murine myotrophin (Mtpn) 3′ UTR that has been shown to be regulated by miR-375 (ref. 48). (**b**) Class 2 targets have an imperfect miRNA 5′ match but significant complementarity of the remainder of the miRNA sequence. Shown is a portion of the *C. elegans lin-41* 3′ UTR that has been shown to be regulated by *let-7* (ref. 49, and references therein).

annotated (10,000 in flies and ~35,000 in humans) different data sets differ considerably and also ignore important and often tissue-specific isoforms. Thus, running the same algorithm on the human RefSeq data set of 3′ UTRs and separately on the dataset of 'known gene' 3′ UTRs results in 10–20% variability in predicted regulatory relationships between genes and miRNAs (N.R., unpublished observations). Furthermore, alternative adenylation sites are typically ignored in current annotations. Many genes have multiple polyA sites in a terminal exon, and 3′UTRs are thus composed of a constitutive and an alternative part of possibly varying length. In two independent ongoing analyses (J. Thierry-Mieg and D. Thierry-Mieg, Aceview, personal communication; and W. Majoros and U. Ohler, personal communication), roughly 70% and 40%, respectively, of genes fall into this category, and alternative polyA sites may possibly generate more transcript diversity than alternative splicing altogether. The biological significance of this diversity remains to be evaluated, but because (i) 3′ UTR lengths of highly expressed genes are tissue dependent[11] (ii) miRNAs can direct deadenylation of mRNAs[12,13] and (iii) in genes with several 3′ UTR isoforms, the majority of human miRNA target predictions fall into alternative parts of 3′ UTRs (W. Majoros and U. Ohler, personal communication), it seems that miRNA targets are affected by 3′ UTR isoform regulation or even, very speculatively, that miRNAs are involved in 3′ UTR isoform regulation. Thus, it will be important to construct more detailed 3′ UTR annotations in the future (see, for example, Aceview (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly)).

Almost all algorithms used evolutionary conservation of target sites as a filter for biologically important targets, and the rules for scoring site conservation, as well as the definition and analysis of orthologous 3′ UTR sequences, differed considerably between methods. Another basic problem was how to incorporate the length of 3′ UTRs into the score assigned to a predicted site. If a 3′ UTR is very long, it is not surprising to find sites just by chance; on the other hand, it was (and is) unknown if a site in a small 3′ UTR has more or less efficacy in repressing the target than a site located in a long 3′ UTR. Finally, the nucleotide content of 3′ UTRs is highly variable (see ref. 14, for example), and methods that do not take this feature into account may produce biased results. In summary, all genome-wide methods predicted crudely a few hundred targets for all miRNAs in each animal clade, only a tiny fraction of these had any experimental support and the overlap between target sites predicted by different algorithms was marginal.

Current miRNA target predictions

A new generation of miRNA target predictions emerged in 2005, mainly based on systematic target-site mutation experiments[15] and more extensive bioinformatic analyses[16–20] that took advantage of the power of cross-species comparisons based on the recent availability of high-quality, genome-wide alignments of several newly sequenced metazoans. The mutation experiments indicated that there exist at least two classes of miRNA targets (**Fig. 1a,b**). One class of targets sites shows perfect Watson-Crick complementarity to the 5′ end of the miRNA and does not require significant further base pairings (**Fig. 1a**). The other class has imperfect 5′ matches but compensates via additional base pairings in the 3′ end of the miRNA (**Fig. 1b**). It was found that the number of statistically detectable target sites in the first class seemed an order of magnitude greater than the number of sites in the second class, and conserved sites in the first class covered a major fraction of genes in vertebrates[17,18,20], flies[15,16,19] and nematodes[21]. For example, two independent approaches suggested that roughly 30% of all human 3′ UTRs (after subtracting the estimated number of false positives) are predicted to be regulated by conserved vertebrate miRNAs via Watson-Crick complementarity of seven conserved consecutive nucleotides between the target site and the 5′ end of the miRNA[17,18]. The total human target predictions of TargetScanS[18] and the ones by PicTar[17] have a high overlap (roughly 80–90% identical sites when run on the same dataset of 3′ UTRs), although the ranking of the roughly 200 predicted targets per miRNA can differ substantially, and TargetScanS does not detect 'class 2' target sites. The number of predicted targets per miRNA is highly variable and ranges from a handful to more than 800 unique transcripts. Interestingly, it was also found that a high fraction of all significantly conserved short 'blocks' of vertebrate 3′ UTR motifs are complementary to the 5′ end of known miRNAs, indicating that 3′ UTR sequence analysis can be used to predict miRNA genes[20] and showing that a significant fraction of 'sharply' conserved 3′ UTR sequence blocks are likely under functional constraints mediated by miRNAs. However, it should be noted that the sequence space covered by sequence motifs complementary to the 5′ end of known miRNAs is relatively small (about 10% of the entire human 3′ UTR sequence space), and many other regulatory factors that bind 3′ UTRs are known to bind more 'fuzzy' *cis*-regulatory motifs, perhaps in many cases constrained by the secondary structure of the mRNA. Thus, a large fraction (perhaps even a majority) of 3′ UTR sequences could be under functional constraints unrelated to miRNAs. Elucidating functional constraints on 3′ UTR sequences by phylogenetic analyses thus remains an open and challenging topic.

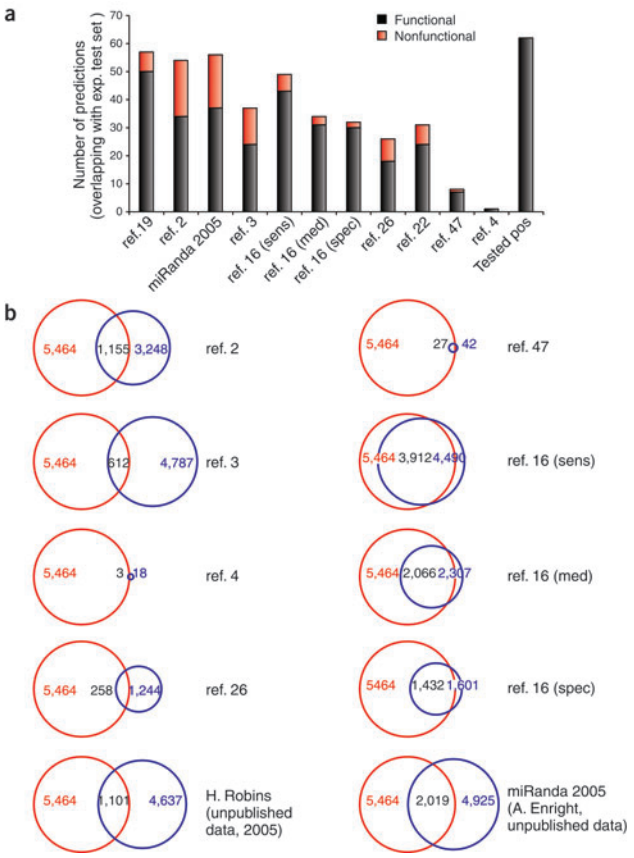

Figure 2 Experimental test of predicted miRNA/mRNA regulatory relationships in *Drosophila melanogaster*. (**a**) Performance of the different published methods on 133 experimentally tested miRNA target pairs. As most methods did not consider 3′ UTRs that were not confirmed by ESTs, these pairs were excluded. Shown are the absolute numbers of tested miRNA target pairs that are predicted. Black bars indicate the prediction of functional pairs (true positives; 'Tested pos' shows the total number of functional pairs) and red bars flag the prediction of nonfunctional pairs (false positives). With reference to ref. 16, 'sens', 'med', and 'spec' refer to the high sensitivity, medium sensitivity/specificity, and high specificity settings of the PicTar algorithm. The miR-278 predictions of miRanda 2005 were not available. As most nonfunctional tests were for this miRNA, and excluding it would have artificially penalized miRanda 2005, the miRanda 2003 predictions were used for miR-278. (**b**) Total overlap of predicted target genes for all 5′ nonredundant cloned *D. melanogaster* miRNAs[41]. Predictions made by the Cohen laboratory[19] are in red circles. Considered were the miRNAs bantam, miR-1, miR-11, miR-12, miR-124, miR-14, miR-184, miR-210, miR-263b, miR-275, miR-276a, miR-277, miR-279, miR-281, miR-285, miR-2a, miR-3, miR-304, miR-305, miR-306, miR-307, miR-308, miR-311, miR-314, miR-315, miR-316, miR-317, miR-31a, miR-34, miR-4, miR-5, miR-7, miR-8, miR-92a, miR-9a, miR-iab-4-3p, miR-iab-4-5p (miR-278 and miR-303 were not available from miRanda 2005 or PicTar, respectively). The absolute number of predictions might differ from those in the respective papers, because depicted predictions by the Cohen laboratory were restricted to a single (arbitrary) representative for each miRNA family. Rajewsky and Socci[4], for example, predicted 39 pairs, of which 21 are not considered in this analysis, because many were not chosen as a family representative (for example, miR-6/tll, which, like several others, is in fact recovered by PicTar). This figure was reprinted from Supplementary Figure 1 of ref. 19 with permission from the authors and publisher (Elsevier). This figure legend is a slightly modified version of that published in ref. 19.

It has been proposed that incorporating the secondary structure of 3′ UTRs may help to predict biologically important miRNA targets[22,23]. Although this idea is intuitive, as sequences inside a 3′ UTR that are stably base paired should be less likely to be bound by miRNAs, it is not yet clear if the proposed algorithms help to improve specificity or sensitivity of algorithms that do not take into account 3′ UTR secondary structure (refs. 19,24, and N.R, unpublished observations). A major problem seems that current programs to predict RNA secondary structure such as MFOLD[25] are typically unreliable when the sequence length exceeds a few hundred nucleotides. On the other hand, folding smaller sequence windows of segments of mRNAs[23] may miss the true *in vivo* structure. Moreover, other mRNA binding factors can change the mRNA secondary structure. For example, it is well known that mRNA binding proteins participate in establishing circularized mRNA upon translation.

How good are the algorithms?

At present, it is difficult to judge which of the algorithms produces the most reliable and/or sensitive target predictions. However, certain trends have emerged. The algorithm developed by the Cohen laboratory[15,19], PicTar[16,17,21] and TargetScanS[18] produce similar overall sets of predicted target sites; most other algorithms produce results that are significantly removed from these[19]. The convergence of algorithms, of course, neither proves the reliability of their predictions nor indicates that they detect all classes of target sites. However, a systematic comparison was recently carried out of miRNA target predictions produced by several algorithms (excluding TargetScanS, for which no predictions have been published in fly) versus ~130 experimentally assayed miRNA-mRNA regulatory relationships in *Drosophila melanogaster* (**Fig. 2** (from ref. 19)). The majority of these experiments were conducted by the Cohen laboratory at the European Molecular Biology Laboratory (EMBL) using an *in vivo* reporter system and miRNA misexpression. Under the nontrivial assumption that this assay can faithfully distinguish between true (endogenous) and nonexistent regulatory relationships, the accuracy of the EMBL and PicTar algorithms (defined as the probability that a prediction was consistent with the experiment) was high (~90%). The sensitivity of these algorithms (defined as the ability to detect true regulatory relationships) was also high (~70–80%). As expected, the overlap between both algorithms was extensive. None of the other current algorithms (including the latest miRanda predictions, RNAhybrid[26] and the algorithm in ref. 22) could simultaneously match this accuracy and sensitivity, but predicted many additional targets. However, there is a possibility that a bias in the selection of predictions that were assayed in the experiments artificially inflated sensitivity and accuracy estimates in favor of the EMBL and PicTar algorithm. Therefore, it seems of importance (as pointed out by ref. 19) that systematic experiments be carried out that test targets exclusively predicted by the other algorithms. In general, because more and more target prediction methods are being published (refs. 27,28 and the 'RNA22 method' of the Rigoutsos group (K.C. Miranda *et al.*, unpublished data) it seems important for the future of the field to keep track of overlaps between different prediction methods and to compare predictions to assayed regulatory relationships. Regrettably, stringent experimental tests of individual miRNA target sites *in vivo* under endogenous conditions have been exceedingly rare. Most tests used luciferase reporter systems and over- or misexpressed miRNAs, and it is unclear how many of the targets validated by these tests are indeed true endogenous targets. It is therefore important to systematically record all miRNA target validation experiments with a detailed description of the assay used (for example, Tarbase[29],

Table 1 Available online resources for miRNA target predictions

Method	Organism	Website
Precomputed predictions on searchable websites		
miRNA target predictions at EMBL	Flies	http://www.russell.embl-heidelberg.de/miRNAs/
miRanda	Flies, vertebrates	http://www.microrna.org//miranda.html
mirBase	Vertebrates, insects, nematodes	http://microrna.sanger.ac.uk/targets/v2/
PicTar	Vertebrates, flies, nematodes	http://pictar.bio.nyu.edu
TargetScan, TargetScanS	Vertebrates	http://genes.mit.edu/targetscan
Ref. 27	Flies, nematodes	http://tavazoielab.princeton.edu/mirnas/
RNA hybrid	Flies	http://www.techfak.uni-bielefeld.de/persons/marc/mirna/targets/drosophila
Tools for locating miRNA targets		
RNAhybrid		http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/welcome.html
DIANA-MicroT		http://diana.pcbi.upenn.edu/DIANA-microT
RNA22		http://cbcsrv.watson.ibm.com/rna22.html
Databases of targets with experimental support		
Tarbase		http://www.diana.pcbi.upenn.edu/tarbase.html
Argonaute		http://www.ma.uni-heidelberg.de/apps/zmf/argonaute/interface
miRNAMAP		http://mirnamap.mbc.nctu.edu.tw/

For other published miRNA target predictions, see ref. 28 (nematodes), ref. 47 (*D. melanogaster*) and ref. 20 (vertebrates).

Argonaute[30] or miRNAMAP[31]). Tarbase allows submission of experimental data, similarly to ORegAnno, an open-access database and curation system for gene regulatory elements in general[32]. **Table 1** lists searchable websites with precomputed target predictions, websites that allow *de novo* prediction of miRNA targets and additional downloadable tables with precomputed predictions.

Combinatorial regulation by miRNAs?

It has been widely proposed that miRNAs may act together to regulate a target mRNA. This concept has been borrowed from the well-known paradigm that many metazoan mRNAs are under control of combinations of transcription factors that together modulate the rate of transcription. It has been further corroborated by *in vitro* experiments[33] as well as the observation that many 3′ UTRs had predicted targets for different miRNAs[3,5,17]. It has also been shown in several cases that multiple sites for the same miRNA boost target repression, and this most simplistic case of combinatorial control has been incorporated into many of the target prediction algorithms. However, it is still unclear if different miRNAs act *in vivo* in any kind of synergistic fashion. Although there has been experimental evidence that different coexpressed vertebrate miRNAs can regulate the murine *Mtpn* gene in a concentration-dependent manner[17], these experiments were performed in cell lines using overexpressed miRNAs. Computational analysis has shown that multiple sites for different fly miRNAs in the same 3′ UTR are significantly differently distributed than expected by chance[19]. However, it could also be that sites for different miRNAs in the same 3′ UTR merely indicate that the mRNA is regulated independently by these miRNAs in different tissues or during development. In summary, the concept of different coexpressed miRNAs coordinately regulating the same mRNA is appealing because it would imply the potential for enhanced specificity of gene regulation mediated by miRNAs and has some experimental support, but it remains to be proven.

Toward lineage-specific miRNA target predictions

Many of the existing target predictions were filtered for conservation in orthologous 3′ UTR sequences. The term 'conserved target site' can refer to different scenarios, and it is important to distinguish between them. Consider two aligned orthologous 3′ UTR sequences from two sufficiently diverged organisms. A target site is commonly flagged as 'conserved' if it appears at the same position in the alignment (that is, the site is aligned). Alternatively, target sites can also be called conserved if they simply appear somewhere in both sequences but are not necessarily aligned. Since algorithms that are used to align 3′ UTR sequences depend on parameters and typically also on the assumption that the order of both sequences has been conserved, strategies that identify conserved sites based on overlapping alignment positions are generally more reliable for species in appropriate phylogenetic distances but always have the tendency to 'lose' sites, especially for more distantly related species. It is also possible that sites 'drift' in sequences; that is, mutations can eliminate a site at a certain position and create a compensatory site at another position. Searching for sites simply based on their presence typically results in decreased signal-to-noise ratios for relatively closely related species; for example, between *D. melanogaster* and *D. pseudoobscura* (N.R., unpublished observations), but it may be the superior method when analyzing more distantly related species. It is also important to carefully consider the phylogeny when defining site conservation. For example, a site that is conserved between human and dog but not between human and mouse may be reasonably flagged as conserved, although mouse and human have a more recent common ancestor than human and dog. The reason is that mouse genomic sequence is in general less closely related to the human sequence than dog to human, most likely owing to accelerated evolution in the rodent lineage[34]. Several of the sites flagged as 'nonconserved' in ref. 24 may therefore very well be conserved. Ultimately, a species-specific phylogenetic

model for neutral 3′ UTR evolution will be required to define site conservation.

Several studies have attempted to study the conservation of predicted miRNA-mRNA regulatory relationships across very large evolutionary distances[16,27] (K. Chen and N.R., unpublished observations); for example, between flies and vertebrates[16]. These studies suggest that a surprisingly small number of regulatory relationships are common between different animal clades (although many of the miRNA genes are well conserved between them) and may suggest extensive rewiring of miRNA-mediated post-transcriptional gene regulation. In the meantime, many more species within each clade have been sequenced (12 *Drosophila* species and 17 vertebrates), so it now seems possible to refine miRNA target predictions toward lineage-specific predictions. At least in flies, it seems relatively straightforward to test some lineage-specific predictions, opening the exciting possibility of understanding more about the possible role of miRNAs in molding organismal diversity.

Correlating miRNA expression and mRNA levels: clues to miRNA function

Recent studies have demonstrated that miRNAs can not only repress translation of mRNAs but can also induce their degradation, even if the target sites have incomplete complementarity to the miRNA[35,36]. This degradation may not necessarily be a result of cleavage but may occur through deadenylation. Therefore, several groups have attempted to detect correlations between the expression domains of miRNAs and the mRNA expression of their targets. Affymetrix microarray data before and after *in vivo* knockdown of the highly liver specific murine miRNA miR-122 (ref. 37) revealed that mostly weak, but significantly upregulated mRNAs were highly specifically and significantly enriched in the 5′ 'recognition motif' of miR-122. Conversely, downregulated transcripts had a tendency to be depleted in this motif. This depletion can be explained with a simple evolutionary argument: mRNAs that are likely to be (indirectly) activated by miR-122 tend to avoid harboring target sites for miR-122. The existence of these 'anti-targets' had been postulated previously[38]. Interestingly, it was shown that the downregulated mRNAs were likely to mediate the only miR-122 knockdown phenotype observed (cholesterol biosynthesis reduction), suggesting that the identification of miRNA 'anti-targets' is perhaps as important as the identification of direct targets. The computational identification of anti-targets based on pure sequence analysis seems difficult, as it is a formidable task to computationally prove selection against harboring a specific small motif in a given 3′ UTR. However, perhaps not very surprisingly, it has now been shown that the 3′ UTR sequences of certain classes of ubiquitously expressed genes (ribosomal genes) are specifically depleted in miRNA target sites[19], and it has been demonstrated that the endogenous expression of several highly specific miRNAs is typically negatively correlated with the endogenous mRNA levels of their targets[11,19,24]. In all three cases, targets were chiefly defined based on perfect complementarity to the 5′ end of the miRNA (typically seven consecutive matches) but with effectively no other constraints on the miRNA-mRNA target[15,17,18]. However, certain differences between these studies emerged. It was suggested that fly miRNAs are expressed when their targets are not ('mutual exclusive expression'[19]), but it seemed that at least in human and mouse tissues, many predicted targets of miRNAs were manifestly expressed in the same tissue where miRNAs were expressed. However, they were expressed at significantly lower levels compared with most other tissues[11,24].

Because evolutionarily conserved targets can be under selection only if they 'see' the miRNA, miRNA and cognate targets must be coexpressed at some point, and 'mutual exclusive expression' could then be explained as an *a posteriori* effect. Alternatively, and very speculatively, it may be that miRNAs could negatively interfere with transcription of their targets, perhaps via binding to target sites during nascent mRNA synthesis or by binding to genomic DNA. Whatever the global picture, all three studies suggest that the mere presence (or absence) of a miRNA 5′ recognition motif in 3′ UTRs is negatively (or positively, respectively) correlated with the expression of these UTRs and the miRNA. Many open questions and caveats remain, as these correlations or "signatures of miRNAs on the mRNA levels of their targets"[11] obviously indicate only trends and, furthermore, cannot be reliably detected for all tissue specific miRNAs. It should also be noted that these studies cannot draw any conclusions about translational regulation mediated by miRNAs. However, it now seems possible to identify at least a large class of targets and anti-targets by knockdown (or overexpression) of miRNAs (or sets of miRNAs in Dicer knockout experiments) and microarray analysis of all mRNAs[13,37]. A computational method (REDUCE[39]) that finds *de novo* sequence motifs that best correlate with the observed changes in mRNA levels has been shown to yield meaningful identification of functional post-transcriptional motifs in 3′ UTRs (including miRNA target sites) without any cross-species comparisons[11,40]. Future developments of this method could aim at integrating transcription factor binding site motifs and post-transcriptional (3′ and 5′ UTR) motifs toward a better quantitative modeling of changes in mRNA levels.

Toward integrating miRNAs into regulatory networks

By now, it seems that a large class of miRNA targets can be confidently detected, so it is natural to integrate these predicted regulatory relationships into existing functional genomics data. An initial attempt in this direction has been published recently for nematodes[21]. Exploration of the 'wiring' of miRNA regulatory relationships together with known protein-protein interaction data, phenotypic data, transcriptional regulatory interactions and other functional genomics data may help to further elucidate the function of miRNAs at a system-wide level. In particular, target predictions should ultimately incorporate expression levels of miRNAs as well as expression levels of their targets. Large-scale attempts to profile miRNA expression in many tissues and during development have been published (see, for example, refs. 41–46, and references therein) or are under way. Expression levels of mRNAs are already readily available in many cases or can be obtained in miRNA knockdown, overexpression or gain-of-function experiments. In summary, it may be that by integrating genome-wide computational and experimental data we have the unprecedented opportunity to study function and evolution of a broad layer of gene regulatory control mediated by miRNAs.

ACKNOWLEDGMENTS

Writing a perspective in a new and fast-moving field is necessarily selective on which papers are being cited. However, I would like to apologize to all authors whose papers were not included. I also want to thank A. Stark, U. Ohler and S. Lall for a critical reading of the manuscript and helpful suggestions, and my collaborators and members of the miRNA/RNAi community for many interesting and stimulating discussions. J. Thierry-Mieg and D. Thierry-Mieg as well as W. Majoros and U. Ohler shared unpublished data with me. Finally, I thank A. Stark and S. Cohen for permission to use **Figure 2**.

COMPETING INTERESTS STATEMENT

The author declares that he has no competing financial interests.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B. & Cohen, S.M. Bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in *Drosophila*. *Cell* **113**, 25–36 (2003).
2. Stark, A., Brennecke, J., Russell, R.B. & Cohen, S.M. Identification of *Drosophila* microRNA targets. *PLoS Biol.* **1**, E60 (2003).
3. Enright, A.J. *et al.* MicroRNA targets in *Drosophila*. *Genome Biol.* **5**, R1 (2003).
4. Rajewsky, N. & Socci, N.D. Computational identification of microRNA targets. *Dev. Biol.* **267**, 529–535 (2004).
5. John, B. *et al.* Human MicroRNA targets. *PLoS Biol.* **2**, e363 (2004).
6. Kiriakidou, M. *et al.* A combined computational-experimental approach predicts human microRNA targets. *Genes Dev.* **18**, 1165–1178 (2004).
7. Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P. & Burge, C.B. Prediction of mammalian microRNA targets. *Cell* **115**, 787–798 (2003).
8. Lai, E.C. Micro RNAs are complementary to 3′ UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat. Genet.* **30**, 363–364 (2002).
9. Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
10. Bentwich, I. Prediction and validation of microRNAs and their targets. *FEBS Lett.* **579**, 5904–5910 (2005).
11. Sood, P., Krek, A., Zavolan, M., Macino, G. & Rajewsky, N. Cell-type-specific signatures of microRNAs on target mRNA expression. *Proc. Natl. Acad. Sci. USA*, published online 13 February 2006 (doi:10.1073/pnas.0511045103).
12. Wu, L., Fan, J. & Belasco, J.G. MicroRNAs direct rapid deadenylation of mRNA. *Proc. Natl. Acad. Sci. USA* **103**, 4034–4039 (2006).
13. Giraldez, A.J. *et al.* Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312**, 75–79 (2006).
14. Robins, H. & Press, W.H. Human microRNAs target a functionally distinct population of genes with AT-rich 3′ UTRs. *Proc. Natl. Acad. Sci. USA* **102**, 15557–15562 (2005).
15. Brennecke, J., Stark, A., Russell, R.B. & Cohen, S.M. Principles of microRNA-target recognition. *PLoS Biol.* **3**, e85 (2005).
16. Grün, D., Wang, Y.L., Langenberger, D., Gunsalus, K.C. & Rajewsky, N. microRNA target predictions across seven *Drosophila* species and comparison to mammalian targets. *PLoS Comput. Biol.* **1**, e13 (2005).
17. Krek, A. *et al.* Combinatorial microRNA target predictions. *Nat. Genet.* **37**, 495–500 (2005).
18. Lewis, B.P., Burge, C.B. & Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20 (2005).
19. Stark, A., Brennecke, J., Bushati, N., Russell, R.B. & Cohen, S.M. Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3′UTR evolution. *Cell* **123**, 1133–1146 (2005).
20. Xie, X. *et al.* Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. *Nature* **434**, 338–345 (2005).
21. Lall, S. *et al.* A genome-wide map of conserved microRNA targets in *C. elegans*. *Curr. Biol.* **16**, 460–471 (2006).
22. Robins, H., Li, Y. & Padgett, R.W. Incorporating structure to predict microRNA targets. *Proc. Natl. Acad. Sci. USA* **102**, 4006–4009 (2005).
23. Zhao, Y., Samal, E. & Srivastava, D. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* **436**, 214–220 (2005).
24. Farh, K.K. *et al.* The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science* **310**, 1817–1821 (2005).
25. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
26. Rehmsmeier, M., Steffen, P., Hochsmann, M. & Giegerich, R. Fast and effective prediction of microRNA/target duplexes. *RNA* **10**, 1507–1517 (2004).
27. Chan, C.S., Elemento, O. & Tavazoie, S. Revealing posttranscriptional regulatory elements through network-level conservation. *PLoS Comput. Biol.* **1**, e69 (2005).
28. Watanabe, Y. *et al.* Computational analysis of microRNA targets in *Caenorhabditis elegans*. *Gene* **365**, 2–10 (2006).
29. Sethupathy, P., Corda, B. & Hatzigeorgiou, A.G. TarBase: A comprehensive database of experimentally supported animal microRNA targets. *RNA* **12**, 192–197 (2006).
30. Shahi, P. *et al.* Argonaute–a database for gene regulation by mammalian microRNAs. *Nucleic Acids Res.* **34**, D115–D118 (2006).
31. Hsu, P.W. *et al.* miRNAMAP: genomic maps of microRNA genes and their target genes in mammalian genomes. *Nucleic Acids Res.* **34**, D135–D139 (2006).
32. Montgomery, S.B. *et al.* ORegAnno: an open access database and curation system for literature-derived promoters, transcription factor binding sites and regulatory variation. *Bioinformatics* **22**, 637–640 (2006).
33. Doench, J.G. & Sharp, P.A. Specificity of microRNA target selection in translational repression. *Genes Dev.* **18**, 504–511 (2004).
34. Lindblad-Toh, K. *et al.* Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803–819 (2005).
35. Lim, L.P. *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769–773 (2005).
36. Bagga, S. *et al.* Regulation by let-7 and lin-4 microRNAs results in target mRNA degradation. *Cell* **122**, 553–563 (2005).
37. Krutzfeldt, J. *et al.* Silencing of microRNAs *in vivo* with 'antagomirs'. *Nature* **438**, 685–689 (2005).
38. Bartel, D.P. & Chen, C.Z. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat. Rev. Genet.* **5**, 396–400 (2004).
39. Bussemaker, H.J., Li, H. & Siggia, E.D. Regulatory element detection using correlation with expression. *Nat. Genet.* **27**, 167–171 (2001).
40. Foat, B.C., Houshmandi, S.S., Olivas, W.M. & Bussemaker, H.J. Profiling condition-specific, genome-wide regulation of mRNA stability in yeast. *Proc. Natl. Acad. Sci. USA* **102**, 17675–17680 (2005).
41. Aravin, A.A. *et al.* The small RNA profile during *Drosophila melanogaster* development. *Dev. Cell* **5**, 337–350 (2003).
42. Chen, P.Y. *et al.* The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes Dev.* **19**, 1288–1293 (2005).
43. Baskerville, S. & Bartel, D.P. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* **11**, 241–247 (2005).
44. Barad, O. *et al.* MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. *Genome Res.* **14**, 2486–2494 (2004).
45. Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
46. Hammond, S.M. microRNA detection comes of age. *Nat. Methods* **3**, 12–3 (2006).
47. Burgler, C. & Macdonald, P.M. Prediction and verification of microRNA targets by MovingTargets, a highly adaptable prediction method. *BMC Genomics* **6**, 88 (2005).
48. Poy, M.N. *et al.* A pancreatic islet-specific microRNA regulates insulin secretion. *Nature* **432**, 226–230 (2004).
49. Vella, M.C. *et al.* The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3′ UTR. *Genes Dev.* **18**, 132–137 (2004).