The functions of animal microRNAs

Victor Ambros

Dartmouth Medical School, Department of Genetics, Hanover, New Hampshire 03755, USA (e-mail: vra@dartmouth.edu)

MicroRNAs (miRNAs) are small RNAs that regulate the expression of complementary messenger RNAs. Hundreds of miRNA genes have been found in diverse animals, and many of these are phylogenetically conserved. With miRNA roles identified in developmental timing, cell death, cell proliferation, haematopoiesis and patterning of the nervous system, evidence is mounting that animal miRNAs are more numerous, and their regulatory impact more pervasive, than was previously suspected.

Noncoding' or 'non-messenger' RNAs are diverse molecules with structural, enzymatic and regulatory functions. Among those with regulatory activity are miRNAs, which are about 22 nucleotides in length and found in all metazoa studied so far[1]. Of the 100–200 genes for distinct miRNAs[1] minimally contained in animal genomes, only a handful have known functions (Table 1). However, these functions suggest that miRNAs are important for the control of animal development and physiology.

As regulators of gene expression, miRNAs can work by essentially two modes[2–7]. In plants, miRNAs base pair with messenger RNA targets by precise or nearly precise complementarity, and direct cleavage and destruction of the target mRNA through a mechanism involving the RNA interference (RNAi) machinery[6,7] (see review in this issue by Meister and Tuschl, page 343). In contrast, most animal miRNAs are imprecisely complementary to their mRNA targets (Fig. 1), and they inhibit protein synthesis through an unknown mechanism that preserves the stability of the mRNA target: some studies even suggest that the translationally repressed target mRNAs remain associated with ribosomes[8,9].

Most miRNA genes seem to be solitary, and are expressed under the control of their own promoters and regulatory sequences. Other miRNA genes are arranged in clusters, and may be co-regulated with other members of the cluster. The *Drosophila* and *Caenorhabditis elegans* genomes each contain at least 100 different miRNA genes[10–15], and vertebrate genomes contain about 250 miRNA genes, as shown by complementary DNA cloning[16,17] and computational predictions[18]. Some miRNAs are abundant, with an estimated 10,000 molecules per cell[15], whereas others are only just detectable by hybridization to total RNA samples.

Here, we focus on recent findings from genomic experiments, whereby large numbers of miRNA genes are identified by cDNA sequencing and computational approaches; forward genetic experiments, whereby mutant genes are isolated from an organism showing abnormal phenotypic characteristics; and reverse genetic experiments, whereby a specific gene is knocked out (and/or overexpressed) to identify its function. These three approaches have identified roles for animal miRNA genes in the regulation of animal development and physiology (Fig. 2). Recently published computational predictions of potential targets of vertebrate and insect miRNAs will also be discussed.

Forward genetic analysis of miRNA function

Despite the success of cDNA cloning and bioinformatics approaches in identifying hundreds of miRNA genes, forward genetics remains one of the most fruitful approaches for identifying miRNA genes with critical roles in the regulation of development and physiology (Fig. 2). The canonical miRNA genes, *lin-4* and *let-7* of *C. elegans*, were first identified by loss-of-function mutations that cause defects in developmental timing in the worm larvae[19,20]. *lin-4* and *let-7* were then cloned on the basis of their mutant phenotypes, and the genes were found to encode small (21–22 nucleotide) noncoding RNAs that are processed from hairpin precursors. These hairpin precursors are a characteristic feature of the miRNA class of regulatory RNAs[21]. Regulatory targets of the *lin-4* and *let-7* miRNAs were identified from the genetic analysis of other genes with developmental timing phenotypes[22–24].

Two *Drosophila* miRNA genes have been identified by forward genetics. Screens for mutants defective in the regulation of programmed cell death and/or cell proliferation in the developing fly led to the identification of the *bantam* locus[25]. When this was cloned, it was found to encode a miRNA[26]. Similar genetic screens identified a locus affecting cell death and fat storage in the fly, which when cloned, was found to correspond to the *mir-14* gene[27]. The *bantam* mutant phenotype included increased frequency of apoptosis. So genes with known roles in regulating the apoptotic pathway were prime candidates for *bantam* targets. The pro-apoptotic gene *hid* was discovered to contain sequences of partial complementarity to *bantam*, and *in vivo* tests supported a direct role for *bantam* in controlling Hid protein synthesis during development of the fly imaginal discs[26]. The targets for *mir-14* in the control of apoptosis have not been identified yet, but may include cell-death effectors other than *hid*[27].

An exciting advance in the identification of developmental roles for animal miRNAs came from recent studies of genetic pathways controlling the asymmetric specification of certain neuronal cell types in *C. elegans*[28–30]. Part of the worm's sensory discriminatory system, which allows worms to distinguish various attractive or repellent chemical stimuli in their environment, consists of two asymmetric chemosensory neurons: ASE left (ASEL) and ASE right (ASER). These two

Table 1 **Animal miRNA genes with genetically assigned functions**

miRNA	Animal	Function	Targets
lin-4	Ce	developmental timing[19]	*lin-14* (refs 19, 22) *lin-28* (ref. 24)
let-7	Ce	developmental timing[20]	*lin-41* (ref. 23) *hbl-1* (refs 48, 49)
lsy-6	Ce	neuronal cell fate[29]	*cog-1* (ref. 29)
mir-273	Ce	neuronal cell fate[30]	*die-1* (ref. 30)
bantam	Dm	cell death, proliferation[26]	*hid* (ref. 26)
mir-14	Dm	cell death, fat storage[27]	caspase?
miR-181	Mm	haematopoietic cell fate[33]	?

Ce, *C. elegans*; Dm, *D. melanogaster*; Mm, *M. musculus*.

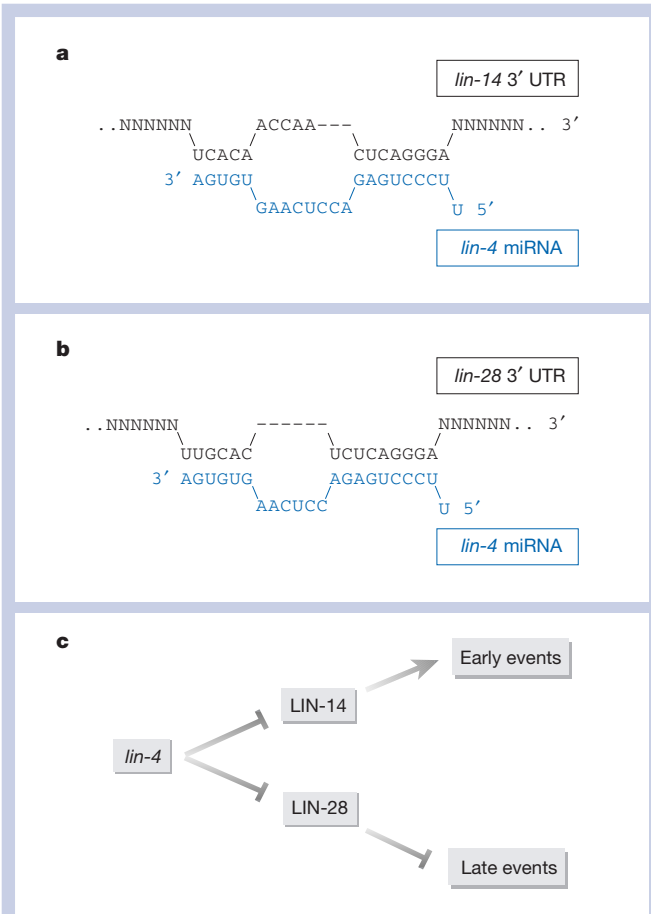

Figure 1 Examples of the imprecise base pairing of animal miRNAs with their targets. The *lin-4* miRNA is shown with its complementary sites in *lin-14* (**a**) and *lin-28* (**b**). There are several further complementary sites of imprecise base pairing in the 3′ UTR of *lin-14* (ref. 22; data not shown); only one site is predicted for *lin-4* in the *lin-28* 3′ UTR[24]. **c**, During larval development of *C. elegans*, *lin-4* coordinates the downregulation of LIN-14 and LIN-28 protein concentrations, which in turn regulates the expression of stage-specific developmental events.

neurons detect different compounds, and together, allow the animal to respond selectively to stimuli. This behavioural specificity is in part rooted in differences in gene expression between ASEL and ASER. ASEL expresses one chemoreceptor gene, *gcy-7*, and ASER expresses another, *gcy-5* (Fig. 3). Genes encoding transcription factors (including *cog-1*, *lim-6* and *ceh-36*) that control the specification of ASEL and ASER asymmetry were identified by screens for mutants with either two ASEL cells or two ASER cells (as judged by *gcy-7::GFP* (where GFP is green fluorescent protein) or *gcy-5::GFP* expression[28]). When one gene identified in these screens, *lsy-6*, was cloned, it was found to be unusual in that it does not encode a protein, but instead produces a novel miRNA[29] that regulates the synthesis of a downstream transcription factor COG-1. The predicted *lsy-6* transcript is a hairpin RNA with partial complementarity to a sequence in the 3′ untranslated region (UTR) of *cog-1* mRNA: these sequences can mediate *lsy-6*-dependent repression of *cog-1* in transgenic worms[29].

The discovery of the *lsy-6* miRNA is remarkable in several respects. First, *lsy-6* probably could not be identified except by forward genetics. The *lsy-6* miRNA was not detected by cDNA cloning, and had not emerged from computational predictions of worm miRNAs[11,31], presumably because its relatively low contribution to worm total RNA meant its expression was not confirmed by northern-blot hybridization. Because *lsy-6* functions in a specific cell type, and a loss-of-function mutation results in a very subtle behavioural phenotype, the discovery of this particular miRNA gene required a genetic screen that used ASEL- or ASER-specific GFP markers to directly assay for cell-fate transformations in mutants. Second, the difficulty of detecting such a scarce RNA as *lsy-6* by hybridization compelled Johnson and Hobert[29] to use unusual genetic experiments to prove that the *lsy-6* transcript forms a miRNA hairpin structure. In their approach, compensatory mutations were introduced into the *lsy-6* locus, and the mutant genes were then tested for function *in vivo*. The wild-type *lsy-6* sequence containing the putative miRNA hairpin structure rescued the *lsy-6* mutant phenotype, but a single point mutation designed to disrupt the predicted hairpin and prevent processing of the mature miRNA eliminated rescuing activity. Combining the original mutation with a compensatory mutation that was predicted to restore the hairpin structure restored rescuing activity. This provided strong genetic evidence that *lsy-6* functions as a hairpin RNA that is probably processed to a miRNA.

miRNAs found by genomics and reverse genetics

The discovery of *lsy-6* by forward genetics suggests that other interesting miRNA genes remain to be identified in *C. elegans* and other animals. An alternative to forward genetics for identifying rare miRNAs with developmental roles is to start with computationally-predicted miRNA genes and test for function by reverse genetics (Fig. 2). Remarkably, the latter approach has identified a second miRNA, *mir-273*, in the same pathway as *lsy*-6 (ref. 30). *mir-273* is also a rare miRNA in the worm, and so was not found by cDNA cloning. Instead, *mir-273* was predicted computationally in a screen for phylogenetically conserved hairpin-forming sequences[31]. Computational methods for finding miRNAs are the ideal complement to cDNA cloning, which works best for relatively abundant miRNAs. However, a limitation of computational approaches is that predicted noncoding RNAs must be validated by an assay that can confirm expression of the RNA. Indeed, the number of worm miRNA candidates predicted computationally has far exceeded the number whose expression can be confirmed[11,31]. A polymerase chain reaction with reverse transcription (RT–PCR)-based assay had to be used to detect the rare *mir-273* RNA in worm total RNA samples[31].

The function of *mir-273* in worm neural development was discovered by the molecular characterization of the same pathway as that which involves *lsy-6* genes regulating ASER and ASEL sensory neuron asymmetry in the worm head. In addition to the *lsy-6* miRNA gene, and the *cog-1* homeobox target of *lsy-6*, several other genes were identified by mutations that alter the specification of ASEL and ASER. Among these was *die-1*, a gene encoding a zinc-finger transcription factor. The *die-1* promoter is active in both neurons, but far more DIE-1 protein (assayed using GFP-tagged DIE-1 expressed from a transgene) accumulates in ASEL than in ASER[30]. Apparently, the *die-1* gene is regulated post-transcriptionally, and expressed specifically in ASEL. No candidate regulatory gene upstream of *die-1* was identified by mutation, but *die-1* mRNA was found to contain sequences that are complementary to parts of *mir-273*. Evidence that *mir-273* could be the repressor of *die-1* in ASER came from the use of transgenes expressing GFP from *mir-273* upstream sequences. In this way, *mir-273* regulatory elements that drove expression specifically in ASER were identified. No *mir-273* loss-of-function mutations were available, but ectopic expression of *mir-273* in ASEL, where its promoter is normally inactive, resulted in the transformation of ASEL into ASER[30].

The identification of two miRNAs in the ASEL/ASER pathway suggests that many developmental regulatory pathways, like the heterochronic and the ASEL/ASER pathways, involve several translational-regulatory steps involving miRNAs. Other miRNAs with very restricted expression patterns undoubtedly remain to be identified by forward genetics and carefully designed phenotypic assays, or by computational prediction followed by more sensitive expression assays.

cDNA cloning of rare, developmentally expressed miRNAs is aided by starting with RNA samples from material enriched for specific organs or cell types. For example, a set of six clustered mouse miRNAs

(*miR-290* to *miR-295*) are expressed specifically in embryonic stem (ES) cells[32] and therefore, would not easily be cloned from differentiated tissues. Similarly, *miR-181* was identified by cloning from selected adult organs[17,33], and the role of *miR-181* in haematopoiesis development was indicated by its substantial enrichment in cDNA libraries made from mouse bone marrow and thymus[33]. Consistent with its being cloned from these tissues, *miR-181* is expressed chiefly in bone-marrow B cells, and in thymus (presumably in T cells)[33].

Although *miR-181* loss-of-function mutations have not yet been isolated, Chen *et al.*[33] used a retroviral-based vector and an overexpression strategy to test whether *miR-181* specifically affects B-cell or T-cell development. Cells were extracted from mouse bone marrow, sorted by cell-type markers, and then infected with virus that had been engineered to overexpress *miR-181*, or other control miRNAs. Overexpression of *miR-181* in bone-marrow haematopoietic progenitor cells increased the numbers of B cells produced *in vitro* and *in vivo* and resulted in decreased numbers of $CD8^+$ T cells[33]. Although overexpression experiments must be interpreted with care when trying to infer the normal function of a gene, the Chen *et al.*[33] results suggest a role for *miR-181* in the normal development of B-cell and T-cell lineages in the mouse. *miR-181* may be a component of the trigger that directs bone-marrow progenitors along the B-cell differentiation pathway. For thymic T cells, *miR-181* seems to have a modulatory role in population homeostasis, perhaps as a sensor for signals that regulate the number of T cells. One prediction from these findings is that knockout of *miR-181* in haematopoietic lineages will lead to fewer B cells and an excess of thymic T cells.

miRNA gene clusters

A prominent characteristic of animal miRNAs is that their genes are often organized in tandem, and are closely clustered on the genome[10,32]. In many cases, such clustered miRNAs are probably processed from the same polycistronic precursor transcript (a single mRNA molecule produced from the transcription of several tandemly arranged genes). When clustered miRNAs are of similar sequence, the cohort of gene products may contribute additively to the regulation of a set of mRNA targets. Clusters can also contain miRNAs of distinct sequences, suggesting that in these cases distinct miRNAs are coordinately deployed towards their various targets. For example, in insects and in vertebrate genomes, *let-7* and *miR-125* (orthologues of *C. elegans let-7* and *lin-4*, respectively) are closely clustered in tandem. *miR-125* and *let-7* are developmentally regulated together in flies[34,35], but in *C. elegans*, *lin-4* and *let-7* are unlinked, and are expressed sequentially[36,20]. The clustered arrangement of *lin-4* and *let-7* in insects and vertebrates could represent an ancestral, conserved functional association between these two temporal regulators. A mammalian miRNA gene-cluster of particular interest encodes the six closely related genes for *miR-290* to *miR-295* in mouse[32]. These are expressed specifically in embryonic stem (ES) cells. Although it is not yet clear whether the human genome contains an orthologous cluster of ES-cell-enriched miRNAs, these findings suggest that translational repression of gene expression by miRNAs contributes to the maintenance of stem-cell potency[32].

The properties of another cluster of miRNAs, on mouse distal chromosome 12 (and conserved at human 14q32), imply other interesting functions for miRNAs[37]. This locus includes at least two miRNA genes (*miR-127* and *miR-136*) that are imprinted: they are expressed exclusively from the maternal chromosome. A retroviral transposon-like gene on the complementary strand to *miR-127* and *miR-136* is expressed only from the paternal chromosome. Consequently, *miR-127* and *miR-136* miRNAs are precisely complementary to the retroviral transposon-like mRNA. This arrangement suggests that *miR-127* and *miR-136* could regulate the transposon by means of an RNAi pathway[37]. Although the function of this distal 12/14q32 cluster of miRNAs is unclear at present, the locus is of interest, if only because many other imprinted genes are known to have roles in early development. Moreover, a high concentration of other hairpin-forming miRNA candidates in the sequences surrounding *miR-127* and *miR-136* (ref. 37) suggests that many more mammalian miRNAs are yet to be discovered.

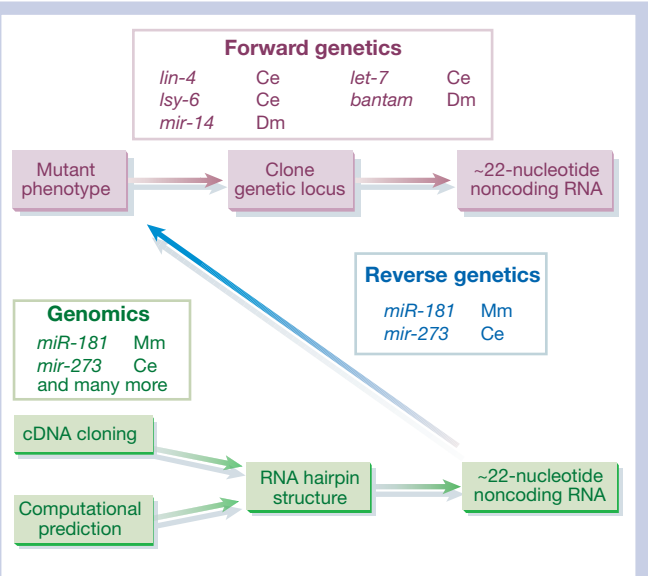

Figure 2 Approaches to miRNA gene discovery and the functional characterization of miRNA genes. Forward genetics approaches to the study of developmental timing in *C. elegans* identified *lin-4* (ref. 19) and *let-7* (ref. 20); and genetic analysis of the specification of *C. elegans* neuronal cell type identified *lsy-6* (ref. 29). Genetic analysis of mutations affecting programmed cell death in *Drosophila* led to the cloning of *bantam*[25,26] and *mir-14* (ref. 27) miRNA genes. Examples of miRNAs that were identified by genomics, and whose functions were subsequently analysed using reverse genetics are mouse *miR-181* (ref. 33; cDNA cloning from mouse) and *mir-273* (refs 30, 31; *C. elegans* computational genomics).

Prediction and validation of miRNA targets

The imprecise base pairing between the typical animal miRNA and a target mRNA suggests that any given miRNA can bind to a broad spectrum of different mRNAs, and so possess an enormous regulatory potential[38]. This raises questions and concerns for systems biologists attempting to use computational approaches to identify miRNA target genes. Of all the predicted candidate targets of a given miRNA, which ones are authentic targets *in vivo*? The development of reliable methods for the computational prediction of animal miRNA targets necessarily involves an iterative process of algorithm design and refinement, guided by *in vivo* tests of the *in silico* predictions. So far, several published papers make good progress towards this end, although more work remains to be done.

Three groups have conducted computational screens for *Drosophila* miRNA targets on a genomic[39,40] or subgenomic scale[41], and two screens for vertebrate miRNA targets have been reported[42,43]. All these approaches are based on the same fundamental assumption, which is that miRNA targets in general can be modelled on the base pairing between *lin-4* and *let-7* of *C. elegans* and their genetically-validated mRNA targets[19,20,22–24]. The salient characteristics of this base pairing are: (1) the location of the miRNA complementary elements in 3′ untranslated regions (UTRs) of mRNA targets, (2) the concentration of base pairing in a 'seed' or 'nucleus' of continuous Watson–Crick base pairing in the 5′-proximal half of the miRNA, and (3) the phylogenetic conservation of the complementary sequences in UTRs of orthologous genes. There is good evidence in support of the general applicability of these assumptions. First, for all the genetically-validated animal miRNA target genes, functional miRNA complementary sites lie within their 3′ UTR sequences[19,20,22–24,26,29,30]. Second, there are several lines of evidence,

apart from the topology of the known *C. elegans lin-4* and *let-7* sites, that point to the importance of base pairing in the 5′ half of the miRNA[1]. Of particular note are recent mutational studies using a variety of different miRNAs that support the primacy of critical base pairs and base-pairing patterns involving the first nine nucleotides of the miRNA[43–45]. Finally, the criterion of phylogenetic conservation was found to be particularly important for reducing what is assumed to be the vast number of false-positive partial-complementary elements predicted for a given miRNA[39–43].

Despite the fundamentally similar approaches used for the published screens for miRNA targets, the results are not particularly congruent. This is probably because the approaches differ in certain important details, including: (1) the particular method of defining phylogenetic conservation, (2) the precise manner of modelling sites based on the verified cases, and (3) how thermodynamic and statistical factors are applied to score and rank predicted sites. As it is essentially impossible at this early stage to know which specific target-prediction method is more accurate, for now, it is probably wise to treat their results as complementary data sets.

Stark *et al.*[39] report the 100 highest-scoring targets for each of 73 known fly miRNAs, and Enright *et al.*[40] report the top 10 most statistically significant predicted targets for each fly miRNA. Of the 730 predicted targets reported by Enright *et al.*[40] about 15% were also flagged by Stark *et al.*[39] The predicted targets identified by both groups arguably represent the most reliable predictions from a biological standpoint. For example, both Enright *et al.*[40] and Stark *et al.*[39] identified genes of the Notch cell-surface receptor pathway as targets of *miR-7*. A third *Drosophila* miRNA target search did not survey all fly genes for miRNA targets, but focused on a selected set of protein genes that encode well-characterized regulators of embryonic patterning[41]. From among these patterning genes, 39 predicted targets of miRNAs were identified. About 20% of these targets were associated with the same miRNAs as those identified by Stark *et al.*[39], but none were among the top 10 lists from Enright *et al.*[40].

Lewis *et al.*[42] identified orthologous genes from human, mouse and rat, which show conserved predicted sites for the same miRNA. A subset of these genes were identified as predicted targets for the same miRNA in mammals, and in the fish *Fugu rubripes*[42]. Kiriakidou *et al.*[43] reported several hundred predicted targets for mammalian miRNAs, and among these, two (initiation factor IF-2 targeted by *miR-20* and the protein ATAXIN targeted by *mir-101*) were also predicted by Lewis *et al.*[42] One possible reason for the lack of overlap between the two vertebrate target screens is the difference in how the two groups dealt with target genes that are predicted to have just a single site for a given miRNA. Lewis *et al.*[42] showed that the occurrence of multiple conserved sites for a given miRNA in the same UTR was a strong predictor of that gene being a statistically significant target. In contrast, genes with a single predicted high-scoring site could not, with confidence, be counted as potential miRNA targets. This was also the case in the insect target searches[39–41]. However, Kiriakidou *et al.* made progress towards reliable prediction of single-site targets[43]. Structure–function studies of a model miRNA complementary site and its cognate miRNA were conducted to derive rules for the characteristics of functional miRNA-target sites[43]. These rules were incorporated into a computational algorithm, and the resulting rates of false-positive predictions (as judged by comparison to a population of randomized miRNA sequences) were low enough to permit statistically significant single-site predictions[43].

Increasing the statistical significance of single-site predictions is important, as there is evidence that multiple sites for the same miRNA are not necessarily required for targeting. For example, there is a single *lin-4* complementary site in the *C. elegans lin-28* 3′ UTR[24]. In fact, Kiriakidou *et al.* used the phylogenetically conserved *let-7b* site in the 3′ UTR of human *lin-28* (ref. 46) as a test case for structure–function analysis, and showed that the site can mediate translational repression in human and mouse cultured cells[43]. These sites seem to be absent from the insect *lin-28* (ref. 46), consistent with evolutionary flexibility in miRNA–target-gene relationships. *let-7* sites in mammalian *lin-28* were not identified by Lewis *et al.*[42], probably because the mouse and human *lin-28* 3′ UTR contains too few repeats of these sites.

In vivo tests of some of the insect and vertebrate miRNA target predictions have produced promising results that tend to support the predictions. By fusing the UTR from predicted targets to a reporter construct, the regulatory capacity of the UTR-containing miRNA complementary elements could be assessed in transgenic flies[39–40] or in animal cells[42,43]. A significant fraction of the predicted UTR targets tested by these 'sensor' assays seemed to confer significant repression on the reporter, compared with a control UTR, in approximate agreement with the estimated frequencies of false positives[39,40,42,43]. A problem with using these reporter tests to validate sites was that it was not always possible to control for the specificity of the miRNA, which would ideally involve comparing the reporter expression levels in the presence and absence of the miRNA.

It is worth noting that even when a predicted site seems to be validated by reporter studies, that site may not normally function *in vivo* (that is, if the miRNA and the mRNA are not normally co-expressed in the same cell). So, to derive a reliable picture of the likely miRNA pathways, it will be necessary to correlate the expression profiles of miRNAs with those of their potential target genes.

How might the current methods for computational prediction of miRNA targets be improved? Further *in vivo* analysis of predicted miRNA–target structures should lead to more refined rules of functional engagement, such as those derived by Kiriakidou *et al.*[43]. If the short interfering RNA (siRNA)-transfection method[43,45] is to become a method of choice for these studies, it is important to ensure that a miRNA introduced exogenously faithfully engages the normal endogenous miRNA-mediated translational repression[8,9]. So far, the target prediction algorithms do not take into account the potential structure of the UTR sequences at or surrounding potential complementary sites. Using transgenic worms, Vella *et al.*[44] tested modified versions of the *C. elegans lin-41* 3′ UTR, which is targeted by *let-7*, and found that sequences surrounding the miRNA sites are important for efficient *let-7*-mediated repression of *lin-41* translation. This finding reinforces the potential importance of incorporating into target prediction algorithms parameters that capture salient structural characteristics of the 3′ UTR.

In future searches for animal miRNA targets, 5′ UTR sequences should be included along with 3′ UTRs , if only to critically test current assumptions about the primacy of 3′ UTRs. Likewise, because plant miRNAs can recognize complementary elements within a coding sequence, future animal target searches should probably address this possibility directly. As discussed above, the assumption that miRNA–target binding is dominated by contiguous Watson–Crick base-pairing involving the 5′-proximal miRNA nucleotides seems to be reasonably well justified on theoretical and experimental grounds. However, there are likely to be important exceptions. For example, a site with nearly perfect complementarity to *miR-196a* is conserved in the 3′ UTR of the *HOXB8* gene of mammals, fish and frog[47]. But the site contains a conserved G•U base pair that breaks the 5′ Watson–Crick 'seed' helix. Remarkably, the G•U base pair is the only imperfection in the *miR-196a*::*HOXB8* match, and it seems that this essentially continuous helix triggers cleavage of *HOXB8* mRNA *in vivo*, instead of translational repression[47]. This result shows that in animals, as in plants, a miRNA–target match of near-perfect complementarity, even one containing G•U base-pairs in the miRNA 5′ proximal block, can cause mRNA destruction.

The evolution of computational algorithms for animal miRNA target prediction will eventually lead to a scheme that is supported by experimental validation, that produces low rates of false-positive and false-negative predictions, and that reliably identifies functional single sites. Accurate rules for functional miRNA–target structures should permit the use of methods that are less dependent on conservation of the miRNA across wide phylogenetic distances. This is

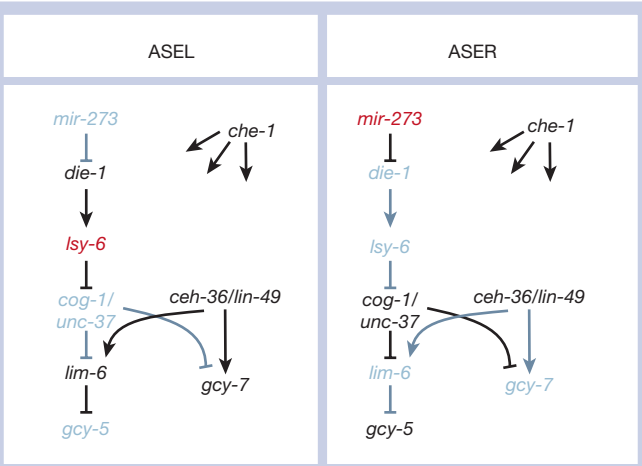

Figure 3 The roles of miRNAs *lsy-6* and *mir-273* in the pathway specifying the sensory neuron cell fates ASEL and ASER[28–30]. The ASE cell type is specified in both ASEL and ASER by the expression of *che-1*, *unc-37* and *ceh-36/lin-49*. Distinct chemosensory functionalities of ASEL and ASER are defined by the transcription of the *gcy-7* or *gcy-5* chemoreceptors respectively. The asymmetric expression of *gcy-7* and *gcy-5* is specified by differential expression of upstream transcription factors, including *die-1*, *cog-1* and *lim-6*. *die-1* is translationally repressed in ASER by the *mir-273* miRNA[30], and *cog-1* is translationally repressed in ASEL by *lsy-6* miRNA[29]. (Figure from Chang *et al.*[30].)

important for miRNAs such as those of the mouse ES-cell cluster, for which well-conserved human orthologues are not apparent. Some important miRNAs and/or miRNA targets may be less well conserved phylogenetically, and could be best identified by comparison of more closely related genomes, such as human and other primates.

The biological repertoire of miRNAs

What can be gleaned from the first set of miRNA/target searches regarding miRNA biology? Very few of the predicted vertebrate miRNA target genes were predicted to be targets for the orthologous miRNAs in insects. This suggests that although the sequence of a miRNA (and hence the sequence of its complementary sites) can be conserved across wide phylogenetic distances, the particular target genes (in which the binding sites reside) may be evolutionarily flexible. If binding sites (or blocks of UTR sequences containing the sites) can transpose between genes during evolution, then the biological pathways controlled by particular miRNAs would be evolutionarily plastic.

Are there any apparent trends in the types of gene that seem to be regulated by miRNAs? In *C. elegans*, the heterochronic genes targeted by *lin-4* or *let-7* encode nuclear proteins (LIN-14, HBL-1)[19,48,49] and apparent RNA binding proteins (LIN-28, LIN-41)[23,24,44]. *lsy-6* and *mir-273* regulate transcription factors[29,30]. Similarly, the lists of predicted insect miRNA targets seem to be enriched in genes encoding transcription factors[39–41], but also include genes with diverse functions that are not directly related to gene expression. For example, the predicted targets of *miR-277* in *Drosophila* include a striking enrichment of genes in the biochemical pathway for the catabolism of leucine, isoleucine and valine[44]. This result strongly suggests that *miR-277* could regulate this biochemical pathway at several points.

The control of cell fate is clearly a common theme for the activity of miRNAs[1]. The heterochronic regulators *lin-4* and *let-7*, and the cell-death regulator *bantam*, control cell-fate choices for diverse cell types — 'early' versus 'late' in the case of the *lin-4*/*let-7* pathway, and proliferation versus death for the *bantam* pathway. *miR-181*, *lsy-6* and *mir-273* control more specific cell-fate choices for specific cell types. Similarly, *miR-181* may have distinct activities in the B-cell and T-cell lineages of mouse haematopoiesis, although the targets in these cells are not obvious from the published target searches.

Outlook

The genetic analysis of miRNA genes in model organisms is beginning to put into place the pieces of a mosaic that will eventually show us the range of functions that miRNAs have in the control of animal development and physiology. The computational prediction of miRNA–target interactions must work in parallel with genetics to identify miRNA-regulated pathways. These computational predictions will probably steadily improve in accuracy, as the existing algorithms are refined through an iterative process of *in silico* prediction and *in vivo* experimentation. Many outstanding questions about miRNAs remain. What affects the accessibility and efficacy of a miRNA at a UTR? What dictates whether an animal miRNA represses translation or directs mRNA cleavage? What is the nature and significance of functional cooperation and redundancy among miRNA genes? To what extent do distinct miRNAs act in combination on the same targets? The analysis of functional redundancy or epistatic relationships among miRNAs should be aided by the development of both genetic and nongenetic methods for the inhibition of particular miRNAs (singly or in combination). For example, recent publication of methods for the inhibition of miRNAs *in vivo* using complementary oligonucleotides[50,51] should encourage the application of gene 'knockdown' approaches for the efficient analysis of miRNA function in cultured cells or intact animals.

The genetic analysis of miRNA function is an exciting challenge: the 'miRNA milieu' in a metazoan cell is likely to hold enormous potential for subtle and complex genetic regulatory interactions involving dozens of miRNAs and their numerous targets[38]. Classical forward genetics continues to demonstrate its power in the miRNA arena, but much of what miRNAs accomplish may be both complex and subtle, demanding the creative application of genomics and reverse genetic approaches. □

doi:10.1038/nature02871

1. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116,** 281–297 (2004).
2. Hutvágner, G. & Zamore, P. D. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297,** 2056–2060 (2002).
3. Zeng, Y., Wagner, E. J. & Cullen, B. R. Both natural and designed microRNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol. Cell* **9,** 1327–1333 (2002).
4. Zeng, Y., Yi, R. & Cullen, B. R. MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc. Natl Acad. Sci. USA* **100,** 9779–9784 (2003).
5. Doench, J. G., Peterson, C. P. & Sharp, P. A. siRNAs can function as miRNAs. *Genes Dev.* **17,** 438–442 (2003).
6. Rhoades, M. W. *et al.* Prediction of plant microRNA targets. *Cell* **110,** 513–520 (2002).
7. Tang, G., Reinhart, B. J., Bartel, D. P. & Zamore, P. D. A biochemical framework for RNA silencing in plants. *Genes Dev.* **17,** 49–63 (2003).
8. Olsen, P. H. & Ambros, V. The *lin-4* regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev. Biol.* **216,** 671–680 (1999).
9. Seggerson, K., Tang, L. & Moss, E. G. Two genetic circuits repress the *Caenorhabditis elegans* heterochronic gene *lin-28* after translation initiation. *Dev. Biol.* **243,** 215–225 (2002).
10. Lau, N. C., Lim, L. P., Weinstein, E. G. & Bartel, D. P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294,** 858–862 (2001).
11. Lee, R. C. & Ambros, V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294,** 862–864 (2001).
12. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **294,** 853–858 (2001).
13. Ambros, V., Lee, R. C., Lavanway, A., Williams, P. T., & Jewell, D. MicroRNAs and other tiny endogenous RNAs in *C. elegans*. *Curr. Biol.* **13,** 807–818 (2003).
14. Lai, E. C., Tomancak, P., Williams, R. W. & Rubin, G. M. Computational identification of *Drosophila* microRNA genes. *Genome Biol.* **4**(R42), 1–20 (2003).
15. Lim, L. P. *et al.* The microRNAs of *Caenorhabditis elegans*. *Genes Dev.* **17,** 991–1008 (2003).
16. Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W. & Tuschl, T. Identification of tissue-specific microRNAs from mouse. *Curr. Biol.* **12,** 735–739 (2002).
17. Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A. & Tuschl, T. New microRNAs from mouse and human. *RNA* **9,** 175–179 (2003).
18. Lim, L. P., Glasner, M. E., Yekta, S., Burge, C. B. & Bartel, D. P. Vertebrate microRNA genes. *Science* **299,** 1540 (2003).
19. Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75,** 843–854 (1993).
20. Reinhart, B. J. *et al.* The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403,** 901–906 (2000).
21. Ambros, V. *et al.* A uniform system for microRNA annotation. *RNA* **9,** 277–279 (2003).
22. 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* **75,** 855–862 (1993).

23. Slack, F. J. *et al.* 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* **5,** 659–669 (2000).
24. Moss, E. G., Lee, R. C. & Ambros, V. The cold shock domain protein LIN-28 controls developmental timing in *C. elegans* and is regulated by the *lin-4* RNA. *Cell* **88,** 637–646 (1997).
25. Hipfner, D. R., Weigmann, K. & Cohen, S. M. The *bantam* gene regulates *Drosophila* growth. *Genetics* **161,** 1527–1537 (2002).
26. 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).
27. Xu, P., Vernooy, S. Y., Guo, M. & Hay, B. A. The *Drosophila* microRNA *mir-14* suppresses cell death and is required for normal fat metabolism. *Curr. Biol.* **13,** 790–795 (2003).
28. Chang, S., Johnston, R. J. Jr & Hobert, O. A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of *C. elegans. Genes Dev.* **17,** 2123–2137 (2003).
29. Johnston, R. J. & Hobert, O. A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans. Nature* **426,** 845–849 (2003).
30. Chang, S., Johnston, R. J. Jr, Frøkjær-Jensen, C., Lockery, S. & Hobert, O. MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. *Nature* **430,** 785–798 (2004).
31. Grad, Y. *et al.* Computational and experimental identification of *C. elegans* microRNAs. *Mol. Cell* **11,** 1253–1263 (2003).
32. Houbaviy, H. B., Murray, M. F. & Sharp, P. A. Embryonic stem-cell-specific microRNAs. *Dev. Cell* **5,** 351–358 (2003).
33. Chen, C. Z., Li, L., Lodish, H. F. & Bartel, D. P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* **303,** 83–86 (2004).
34. Sempere, L. F., Sokol, N. S., Dubrovsky, E. B., Berger, E. M. & Ambros, V. Temporal regulation of microRNA expression in *Drosophila melanogaster* mediated by hormonal signals and broad-complex gene activity. *Dev. Biol.* **259,** 9–18 (2003).
35. Bashirullah, A. *et al.* Coordinate regulation of small temporal RNAs at the onset of *Drosophila* metamorphosis. *Dev. Biol.* **259,** 1–8 (2003).
36. Feinbaum, R. & Ambros, V. The timing of *lin-4* RNA accumulation controls the timing of postembryonic developmental events in *Caenorhabditis elegans. Dev. Biol.* **210,** 87–95 (1999).
37. Seitz, H. *et al.* Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene. *Nature Genet.* **34,** 261–262 (2003).
38. Bartel, D. P. & Chen, C.-Z. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nature Rev. Genet.* **5,** 396–401 (2004).
39. Stark, A., Brennecke, J., Russell, R. B. & Cohen, S. M. Identification of *Drosophila* microRNA targets. *PLoS Biol.* **1,** E60 (2003).
40. Enright, A. J. *et al.* MicroRNA targets in *Drosophila. Genome Biol.* **5,** R1 (2003).
41. Rajewskya, N. & Socci, N. D. Computational identification of microRNA targets. *Dev. Biol.* **267,** 529–535 (2004).
42. Lewis, B. P., Shih, I., Jones-Rhoades, M. W., Bartel, D. P. & Burge, C. B. Prediction of mammalian microRNA targets. *Cell* **115,** 787–798 (2003).
43. Kiriakidou, M. *et al.* A combined computational-experimental approach predicts human microRNA targets. *Genes Dev.* **18,** 1165–1178 (2004).
44. Vella, M. C., Choi, E. Y., Lin, S. Y., Reinert, K. & Slack, F. J. The *C. elegans* microRNA *let-7* binds to imperfect *let-7* complementary sites from the *lin-41* 3′ UTR. *Genes Dev.* **18,** 32–37 (2004).
45. Doench, J. G. & Sharp, P. A. Specificity of microRNA target selection in translational repression. *Genes Dev.* **18,** 504–511 (2004).
46. Moss, E. G. & Tang, L. Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. *Dev. Biol.* **258,** 432–442 (2003).
47. Yekta, S., Shih, I.-h. & Bartel, D. P. MicroRNA-directed cleavage of *HOXB8* mRNA. *Science*, **304,** 594–596 (2004).
48. Abrahante, J. E. *et al.* The *Caenorhabditis elegans* hunchback-like gene *lin-57/hbl-1* controls developmental time and is regulated by microRNAs. *Dev. Cell* **4,** 625–637 (2003).
49. Lin, S. Y. *et al.* The *C. elegans* hunchback homolog, *hbl-1,* controls temporal patterning and is a probable microRNA target. *Dev. Cell* **4,** 639–650 (2003).
50. Hutvágner, G., Simard, M. J., Mello, C. C. & Zamore, P. D. Sequence-specific inhibition of small RNA function. *PLoS Biol.* **2,** E98 (2004).
51. Meister, G., Landthaler, M., Dorsett, Y. & Tuschl, T. Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. *RNA* **10,** 544–550 (2004).

Acknowledgements We thank members of the Ambros laboratory and numerous other colleagues for stimulating discussions. V.A. is supported by NIH grants.

Competing interests statement The author declares that he has no competing financial interests.