MicroRNAs: Genomics, Biogenesis, Mechanism, and Function

Review

David P. Bartel[1,2,*]
[1]Whitehead Institute for Biomedical Research
9 Cambridge Center
Cambridge, Massachusetts 02142
[2]Department of Biology
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

MicroRNAs (miRNAs) are endogenous ~22 nt RNAs that can play important regulatory roles in animals and plants by targeting mRNAs for cleavage or translational repression. Although they escaped notice until relatively recently, miRNAs comprise one of the more abundant classes of gene regulatory molecules in multicellular organisms and likely influence the output of many protein-coding genes.

In an investigation inspiring for both its perseverance and its scientific insight, Victor Ambros and colleagues, Rosalind Lee and Rhonda Feinbaum, discovered that *lin-4*, a gene known to control the timing of *C. elegans* larval development, does not code for a protein but instead produces a pair of small RNAs (Lee et al., 1993). One RNA is approximately 22 nt in length, and the other is approximately 61 nt; the longer one was predicted to fold into a stem loop proposed to be the precursor of the shorter one. The Ambros and Ruvkun labs then noticed that these *lin-4* RNAs had antisense complementarity to multiple sites in the 3′ UTR of the *lin-14* gene (Lee et al., 1993; Wightman et al., 1993). This complementarity fell in a region of the 3′ UTR previously proposed to mediate the repression of *lin-14* by the *lin-4* gene product (Wightman et al., 1991). The Ruvkun lab went on to demonstrate the importance of these complementary sites for regulation of *lin-14* by *lin-4*, showing also that this regulation substantially reduces the amount of LIN-14 protein without noticeable change in levels of *lin-14* mRNA. Together, these discoveries supported a model in which the *lin-4* RNAs pair to the *lin-14* 3′ UTR to specify translational repression of the *lin-14* message as part of the regulatory pathway that triggers the transition from cell divisions of the first larval stage to those of the second (Lee et al., 1993; Wightman et al., 1993).

The shorter *lin-4* RNA is now recognized as the founding member of an abundant class of tiny regulatory RNAs called microRNAs or miRNAs (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). The breadth and importance of miRNA-directed gene regulation are coming into focus as more miRNAs and their regulatory targets and functions are discovered. Recently discovered miRNA functions include control of cell proliferation, cell death, and fat metabolism in flies (Brennecke et al., 2003; Xu et al., 2003), neuronal patterning in nematodes (Johnston and Hobert, 2003), mod-

ulation of hematopoietic lineage differentiation in mammals (Chen et al., 2004), and control of leaf and flower development in plants (Aukerman and Sakai, 2003; Chen, 2003; Emery et al., 2003; Palatnik et al., 2003). Computational approaches for finding messages controlled by miRNAs indicate that these examples represent a very small fraction of the total (Rhoades et al., 2002; Enright et al., 2003; Lewis et al., 2003; Stark et al., 2003).

This review highlights what has been learned about miRNAs in the decade since the report of the *lin-4* RNA and its regulation of *lin-14*. The major topics discussed are miRNA genomics, miRNA biogenesis, miRNA regulatory mechanisms, and the roles of miRNAs in gene regulatory pathways.

Genomics: The miRNA Genes

For seven years after the discovery of the *lin-4* RNA, the genomics of this type of tiny regulatory RNA appeared simple: there was no evidence for *lin-4*-like RNAs beyond nematodes and no sign of any similar noncoding RNAs within nematodes. This all changed upon the discovery that *let-7*, another gene in the *C. elegans* heterochronic pathway, encoded a second ~22 nt regulatory RNA. The *let-7* RNA acts to promote the transition from late-larval to adult cell fates in the same way that the *lin-4* RNA acts earlier in development to promote the progression from the first larval stage to the second (Reinhart et al., 2000; Slack et al., 2000). Furthermore, homologs of the *let-7* gene were soon identified in the human and fly genomes, and *let-7* RNA itself was detected in human, *Drosophila*, and eleven other bilateral animals (Pasquinelli et al., 2000).

Because of their common roles in controlling the timing of developmental transitions, the *lin-4* and *let-7* RNAs were dubbed small temporal RNAs (stRNAs), with anticipation that additional regulatory RNAs of this type would be discovered (Pasquinelli et al., 2000). Indeed, less than one year later, three labs cloning small RNAs from flies, worms, and human cells reported a total of over one hundred additional genes for tiny noncoding RNAs, approximately 20 new genes in *Drosophila*, approximately 30 in human, and approximately 60 in worms (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). The RNA products of these genes resembled the *lin-4* and *let-7* stRNAs in that they were ~22 nt endogenously expressed RNAs, potentially processed from one arm of a stem loop precursor (Figure 1), and they were generally conserved in evolution—some quite broadly, others only in more closely related species such as *C. elegans* and *C. briggsae*. But unlike *lin-4* and *let-7* RNAs, many of the newly identified ~22 nt RNAs were not expressed in distinct stages of development and instead were more likely to be expressed in particular cell types. Thus the term microRNA was used to refer to the stRNAs and all the other tiny RNAs with similar features but unknown functions (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). Intensified cloning efforts have revealed numer-
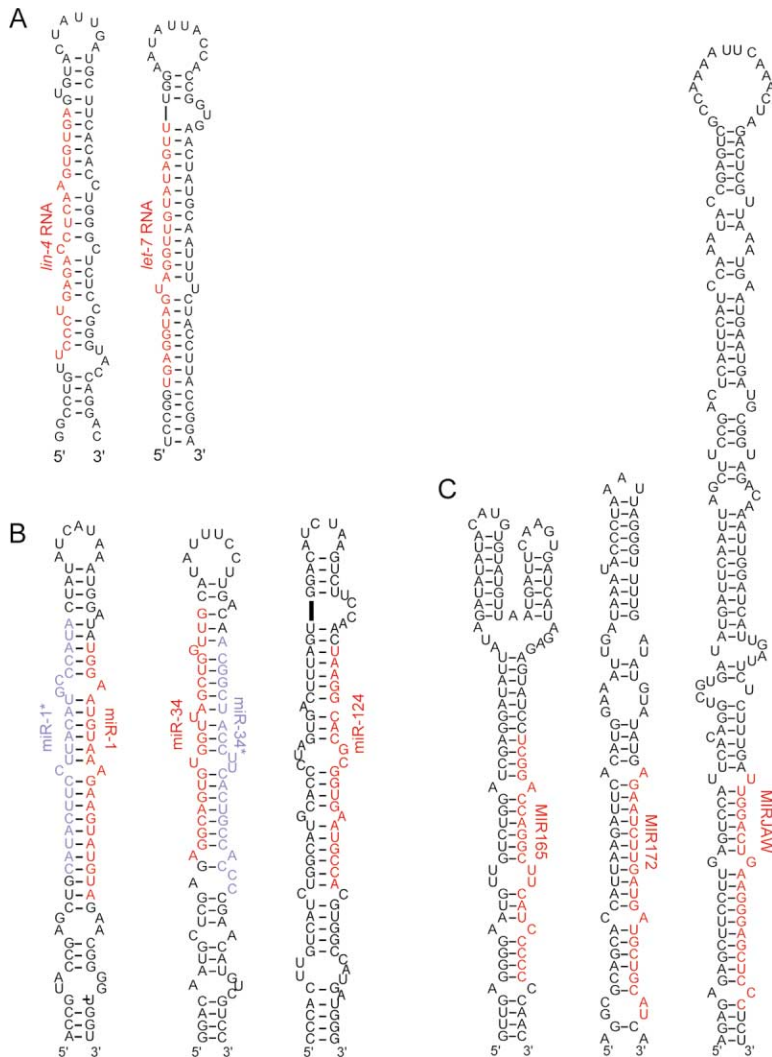
*Correspondence: dbartel@wi.mit.edu

A

B

C

ous additional miRNA genes in mammals, fish, worms, and flies (Lagos-Quintana et al., 2002, 2003; Mourelatos et al., 2002; Ambros et al., 2003b; Aravin et al., 2003; Dostie et al., 2003; Houbaviy et al., 2003; Kim et al., 2003; Lim et al., 2003a, 2003b; Michael et al., 2003). A registry has been set up to catalog the miRNAs and facilitate the naming of newly identified genes (Griffiths-Jones, 2004).

Like *C. elegans lin-4* and *let-7*, most miRNA genes come from regions of the genome quite distant from previously annotated genes, implying that they derive from independent transcription units (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). Nonetheless, a sizable minority (e.g., about a quarter of the human miRNA genes) are in the introns of pre-mRNAs. These are preferentially in the same orientation as the predicted mRNAs, suggesting that most of these miRNAs are not transcribed from their own promoters but are instead processed from the introns, as seen also for many snoRNAs (Aravin et al., 2003; Lagos-Quintana et al., 2003; Lai et al., 2003; Lim et al., 2003a). This arrangement provides a convenient mechanism for the coordinated expression of a miRNA and a protein. Regu-

latory scenarios are easy to imagine in which such coordinate expression could be useful, which would explain the conserved relationships between miRNAs and host mRNAs. A striking example of this conservation involves *mir-7*, found in the intron of *hnRNP K* in both insects and mammals (Aravin et al., 2003).

Other miRNA genes are clustered in the genome with an arrangement and expression pattern implying transcription as a multi-cistronic primary transcript (Lagos-Quintana et al., 2001; Lau et al., 2001). Although the majority of worm and human miRNA genes are isolated and not clustered (Lim et al., 2003a, 2003b), over half of the known *Drosophila* miRNAs are clustered (Aravin et al., 2003). The miRNAs within a genomic cluster are often, though not always, related to each other; and related miRNAs are sometimes but not always clustered (Lagos-Quintana et al., 2001; Lau et al., 2001). Orthologs of *C. elegans lin-4* and *let-7* are clustered in the fly and human genomes and are coexpressed, sometimes from the same primary transcript, leading to the idea that the genomic separation of *lin-4* from *let-7* in nematodes might be unique to the worm lineage (Aravin et al., 2003; Bashirullah et al., 2003; Sempere et al., 2003). This exam-

ple illustrates the possibility that even in cases where clustered genes have no apparent sequence homology, they may share functional relationships.

Some of the more interesting genomic locations of miRNA genes include those in the Hox clusters. The *mir-10* gene lies in the Antennapedia complex of insects and in the orthologous locations in two Hox clusters of mammals, whereas the *mir-iab-4* gene is within the insect Bithorax cluster (Aravin et al., 2003; Lagos-Quintana et al., 2003). In light of the roles of other genes of the Hox clusters, the Hox miRNAs are especially good candidates for having interesting functions in animal development. Other interesting loci include the *mir-15a-mir-16* cluster, which falls within a region of human chromosome 13 thought to harbor a tumor suppressor gene because it is the site of the most common structural aberrations in both mantle cell lymphoma and B cell chronic lymphocytic leukemia (Lagos-Quintana et al., 2001; Calin et al., 2002).

Nearly all of the cloned miRNAs are conserved in closely related animals, such as human and mouse, or *C. elegans* and *C. briggsae* (Lagos-Quintana et al., 2003; Lim et al., 2003a, 2003b). This statement remains true even when ignoring evolutionary conservation as a criterion for classifying clones as miRNAs. Many are also conserved more broadly among the animal lineages (Ambros et al., 2003b; Aravin et al., 2003; Lagos-Quintana et al., 2003; Lim et al., 2003a). For instance, more than a third of the *C. elegans* miRNAs have easily recognized homologs among the human miRNAs (Lim et al., 2003a). When comparing distant lineages, considerable expansion or contraction of gene families is apparent, the most striking example being the *let-7* family, which has four identified members in *C. elegans* and at least 15 in human, but only one in *Drosophila* (Pasquinelli et al., 2000; Aravin et al., 2003; Lai et al., 2003; Lim et al., 2003a).

Genomics: miRNA Expression

Many miRNAs have intriguing expression patterns. For example, paralogs and orthologs of the *C. elegans lin-4* and *let-7* RNAs have stage-specific expression in development as if they, too, function as stRNAs (Pasquinelli et al., 2000; Lau et al., 2001; Lagos-Quintana et al., 2002; Bashirullah et al., 2003; Lim et al., 2003a). Other interesting examples include miR-1, which is primarily found in the mammalian heart (Lee and Ambros, 2001; Lagos-Quintana et al., 2002); miR-122, which is primarily in the liver (Lagos-Quintana et al., 2002); miR-223, which is primarily in the granulocytes and macrophages of mouse bone marrow (Chen et al., 2004); miRNAs of the *mir-35–mir-42* cluster, which are preferentially in the *C. elegans* embryo (Lau et al., 2001); and those of the *mir-290–mir-295* cluster, which are expressed in mouse embryonic stem cells but not in differentiated cells (Houbaviy et al., 2003). Expression array technology has been adapted to examine miRNAs and has revealed distinct expression patterns in different developmental stages or regions of the mammalian brain (Krichevsky et al., 2003). With all the different genes and expression patterns, it is reasonable to propose that every metazoan cell type at each developmental stage might have a distinct miRNA expression profile—providing ample op-

portunity for "micromanaging" the output of the transcriptome.

Another remarkable aspect of miRNA expression is the sheer abundance of certain miRNAs in the cells. For example, miR-2, miR-52, and miR-58 are each present on average at more than 50,000 molecules per adult worm cell—a greater abundance than the U6 snRNA of the spliceosome (Lim et al., 2003a). Whether this high expression is attributable to very robust transcription or to slow decay is not yet known. Some miRNAs are expressed at much lower levels. For instance, miR-124 is present in the adult worm on average at 800 molecules per cell (Lim et al., 2003a). This lower average level (though still higher than that of the typical mRNA) might be due to low expression in many cells or high expression in just a few cells. The finding that the mouse ortholog of miR-124 is nearly exclusively expressed in the brain supports the latter explanation (Lagos-Quintana et al., 2002).

Genomics: Computational Approaches and Gene Number

There has been some speculation as to why miRNAs were not discovered earlier; the answer is clearly not that they are rare. MicroRNAs and their associated proteins appear to be one of the more abundant ribonucleoprotein complexes in the cell. Nonetheless, miRNAs whose expression is restricted to nonabundant cell types or specific environmental conditions could still be missed in cloning efforts. Thus, computational approaches have been developed to complement experimental approaches to miRNA gene identification. From early on, homology searches have revealed orthologs and paralogs of known miRNA genes (Pasquinelli et al., 2000; Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). Another simple approach has been to search the vicinity of known miRNA genes for other stem loops that might represent additional genes of a genomic cluster (Lau et al., 2001; Aravin et al., 2003; Seitz et al., 2003; Ohler et al., 2004). This strategy is important because some of the most rapidly evolving miRNA genes are present as tandem arrays within operon-like clusters, and the divergent sequences of these genes make them relatively difficult to spot using the more general approaches.

Gene-finding approaches that do not depend on homology or proximity to known genes have also been developed and applied to entire genomes (Ambros et al., 2003b; Grad et al., 2003; Lai et al., 2003; Lim et al., 2003a). They typically start by identifying conserved genomic segments that both fall outside of predicted protein-coding regions and potentially could form stem loops and then score these candidate miRNA stem loops for the patterns of conservation and pairing that characterize known miRNAs genes. So far, the two most sensitive computational scoring tools are MiRscan, which has been systematically applied to nematode and vertebrate candidates (Lim et al., 2003a, 2003b), and miRseeker, which has been systematically applied to insect candidates (Lai et al., 2003). Both MiRscan and miRseeker have identified dozens of genes that were subsequently (or concurrently) verified experimentally. Because of their relatively high sensitivity, MiRscan and miRseeker

have also enabled reasonably firm estimates of the number of miRNA genes in the genomes of human (200–255 miRNA genes; Lim et al., 2003b), *C. elegans* (103–120 genes; Lim et al., 2003a; Ohler et al., 2004), and *Drosophila* (96–124 genes; Lai et al., 2003). In each species, these numbers represent nearly 1% of the predicted genes in the genome, a fraction similar to that of other large gene families with regulatory roles, such as the homeodomain transcription-factor family.

These estimates imply that the majority of miRNA genes have now been found in the mammalian and nematode lineages—particularly in *C. elegans*, where approximately 100 miRNA genes have been identified. (This tally is conservative in that it excludes some reported genes that appear to be questionable [Ohler et al., 2004].) In *Drosophila*, 77 genes, representing 71 unique miRNAs, have been reliably identified (Aravin et al., 2003; Lai et al., 2003), and in humans, approximately 175 genes, representing approximately 145 unique miRNAs, have either been validated in human cells or identified based on their homology to genes validated in mouse or zebrafish (miRNA Registry, release 3.0; Griffiths-Jones, 2004). When considering the number of miRNAs remaining to be identified or validated in these species, it is important to remember that gene number estimates by MiRscan and miRseeker rest on the assumption that the stem loops of the rare, difficult-to-clone miRNAs will show patterns of conservation and pairing resembling those of the abundant, easily cloned miRNAs. This assumption appears to hold for *C. elegans*, for which there was a reassuring lack of correlation between the number of times an miRNA was cloned and its MiRscan score (Lim et al., 2003a).

If instead a disproportionate number of difficult-to-clone miRNAs are also difficult to identify computationally, then estimates of the number of miRNA genes in the genome will be too low. This might be the situation in humans—perhaps because the vertebrate genomes used in the analysis are more highly diverged. Most of the first 109 miRNAs cloned from mammals have readily identifiable homologs in the genome of pufferfish (*Fugu ripens*), which enabled MiRscan analysis to identify 81 (74%) of these genes by scoring stem loops conserved in human, mouse, and fish (Lim et al., 2003b). Extrapolating from this sensitivity and the number of additional candidates with scores matching the known miRNAs, an upper bound on the number of human miRNA genes was calculated to be 255 (Lim et al., 2003b). However, more recently identified mammalian miRNA genes appear relatively less likely to be conserved in fish, particularly those genes cloned from embryonic stem cells and mammalian brain and the 14 miRNA candidates residing in a large imprinted cluster (Houbaviy et al., 2003; Kim et al., 2003; Seitz et al., 2003). These recent data suggest that the more difficult-to-clone mammalian miRNAs are less likely to be conserved in fish and thus less likely to have been identified computationally, which implies that a confident upper bound on the number of human genes is difficult to determine using analyses that extended to fish and that 255 is too low a value for this upper bound—although it still might exceed the actual number of human miRNA genes.

Genomics: miRNAs in Plants

Cloning of small RNAs from plants has also revealed miRNAs, although the multitude of other 21 to 24 nt RNAs found in plants sometimes complicated their initial classification (Llave et al., 2002a; Mette et al., 2002; Park et al., 2002; Reinhart et al., 2002). Like the metazoan miRNAs, the plant miRNAs (1) are endogenously expressed ∼22 nt RNAs potentially processed from one arm of foldback precursors, (2) are generally conserved in evolution, and (3) come from regions of the genome distinct from previously annotated genes (Reinhart et al., 2002). To date, 20 unique *Arabidopsis* miRNAs have been reported; a few are closely related to each other, and thus the reported genes represent 15 distinct miRNA families. (Bartel and Bartel, 2003; Palatnik et al., 2003). Because some could be derived from multiple genomic loci, the 20 miRNAs could represent more than 40 *Arabidopsis* genes. The homology searches based on the cloned genes also reveal numerous potential paralogs with a point substitution or two in the predicted miRNA. Additional gene families are likely to be found when the cloning of small plant RNAs is scaled up and computational gene-finding methods are extended to plants. It appears that, as in animals, a substantial fraction of the gene regulatory molecules in plants could be RNA rather than protein.

The discovery of miRNAs in both plants and animals suggests that this class of noncoding RNAs has been modulating gene expression since at least the last common ancestor of these lineages (Reinhart et al., 2002). Nonetheless, plant and animal miRNAs differ in some aspects, which appear to be related to differences in their biogenesis. The most notable differences are in the miRNA stem loops; the plant predicted foldbacks are much more variable in size and typically larger than those of animals (Figure 1; for a more comprehensive look at plant miRNA predicted stem loops, see online supplemental material of Reinhart et al., 2002). More subtle differences include somewhat more pairing between the miRNA and the other arm of the stem loop in plants compared to animals, a tighter distribution of plant miRNA lengths that centers on 21 nt rather than the 22–23 nt lengths most often seen in animals, and perhaps a stronger preference for a U at the 5′ terminus of the plant miRNAs (Lau et al., 2001; Reinhart et al., 2002; Bartel and Bartel, 2003). These differences, together with the absence of reports that particular miRNA genes are conserved between plants and animals, leave open the prospect that miRNA genes arose independently in each of these multicellular lineages, after their last common ancestor (which is thought to have been unicellular). Even in this scenario of dual origins, the presence of miRNAs in all plant and animal species examined thus far suggests early origins in both lineages, perhaps preceding and facilitating the developmental patterning needed for multicellular body plans.

Biogenesis: miRNA Transcription

A 693 bp genomic fragment rescues the *lin-4* deficiency, implying that all the elements required for the regulation and initiation of transcription are located in this short fragment (Lee et al., 1993). However, little is known regarding these transcriptional processes for *lin-4* or any

other miRNA gene. Some miRNAs residing in introns are likely to share their regulatory elements and primary transcript with their pre-mRNA host genes. For the remaining miRNA genes, presumably transcribed from their own promoters, no primary transcripts have been fully defined. Nonetheless, these primary miRNA transcripts, called pri-miRNAs (Lee et al., 2002), are generally thought to be much longer than the conserved stem loops currently used to define miRNA genes, as suggested by the following: (1) the idea that clustered miRNA stem loops are transcribed from a single primary transcript (Lagos-Quintana et al., 2001; Lau et al., 2001), (2) matches between miRNAs and lengthy ESTs in the databases (Lagos-Quintana et al., 2002; Aukerman and Sakai, 2003), (3) RT-PCR experiments amplifying large fragments of the pri-miRNAs (Lee et al., 2002; Aravin et al., 2003).

The two candidate RNA polymerases for pri-miRNA transcription are pol II and pol III. Pol II produces the mRNAs and some noncoding RNAs, including the small nucleolar RNAs (snoRNAs) and four of the small nuclear RNAs (snRNAs) of the spliceosome, whereas pol III produces some of the shorter noncoding RNAs, including tRNAs, 5S ribosomal RNA, and the U6 snRNA. The miRNAs processed from the introns of protein-coding host genes are undoubtedly transcribed by pol II. The following observations provide indirect evidence that many of the other miRNAs also are pol II products, even though most of the metazoan miRNA genes do not have the classical signals for polyadenylylation (Ohler et al., 2004): (1) The pri-miRNAs can be quite long, more than one 1 kb, which is longer than typical pol III transcripts. (2) These presumed pri-miRNAs often have internal runs of uridine residues, which would be expected to prematurely terminate pol III transcription. (3) Many miRNAs are differentially expressed during development, as is observed often for pol II but not pol III products. (4) Fusions that place the open reading frame of a reporter protein downstream from the 5′ portion of miRNA genes lead to robust reporter protein expression, suggesting that miRNA primary transcripts are capped pol II transcripts. Examples of such fusions include artificial reporter constructs designed to investigate the regulation of miRNA expression (Johnson et al., 2003; Johnston and Hobert, 2003) and a natural chromosome translocation linked to an aggressive B cell leukemia, in which a truncated *MYC* gene is fused to the 5′ portion of *mir-142* (Gauwerky et al., 1989; Lagos-Quintana et al., 2002). Although these observations indicate that many miRNAs are pol II transcripts, others might still be pol III transcripts, just as most but not all snRNAs are pol II products. Ectopic expression of miR-142 and other miRNAs from a pol III promoter produces efficiently and precisely processed miRNAs that function in vivo (Chen et al., 2004), indicating that there is no obligate link between the identity of the polymerase and downstream miRNA processing or function.

Biogenesis: miRNA Maturation

The current model for maturation of the mammalian miRNAs is shown in Figure 2B. The first step is the nuclear cleavage of the pri-miRNA, which liberates a ~60–70 nt stem loop intermediate, known as the miRNA precursor, or the pre-miRNA (Lee et al., 2002; Zeng and Cullen, 2003). This processing is performed by the Drosha RNase III endonuclease, which cleaves both strands of the stem at sites near the base of the primary stem loop (Lee et al., 2003) (Figure 2B, step 2). Drosha cleaves the RNA duplex with a staggered cut typical of RNase III endonucleases, and thus the base of the pre-miRNA stem loop has a 5′ phosphate and ~2 nt 3′ overhang (Basyuk et al., 2003; Lee et al., 2003). This pre-miRNA is actively transported from the nucleus to the cytoplasm by Ran-GTP and the export receptor Exportin-5 (Yi et al., 2003; Lund et al., 2004) (Figure 2B, step 3).

The nuclear cut by Drosha defines one end of the mature miRNA. The other end is processed in the cytoplasm by the enzyme Dicer (Lee et al., 2003). Dicer, also an RNase III endonuclease, was first recognized for its role in generating the small interfering RNAs (siRNAs) that mediate RNA interference (RNAi) (Bernstein et al., 2001) and was later shown to play a role in miRNA maturation (Grishok et al., 2001; Hutvágner et al., 2001; Ketting et al., 2001). According to the current model of miRNA maturation, Dicer performs an activity in metazoan miRNA maturation similar to that which it performs when chopping up double-stranded RNA during RNAi: It first recognizes the double-stranded portion of the pre-miRNA, perhaps with particular affinity for a 5′ phosphate and 3′ overhang at the base of the stem loop. Then, at about two helical turns away from the base of the stem loop, it cuts both strands of the duplex. This cleavage by Dicer lops off the terminal base pairs and loop of the pre-miRNA, leaving the 5′ phosphate and ~2 nt 3′ overhang characteristic of an RNase III and producing an siRNA-like imperfect duplex that comprises the mature miRNA and similar-sized fragment derived from the opposing arm of the pre-miRNA (Figure 2B, step 4).

The fragments from the opposing arm, called the miRNA* sequences (Lau et al., 2001), are found in libraries of cloned miRNAs but typically at much lower frequency than are the miRNAs (Lagos-Quintana et al., 2002; Aravin et al., 2003; Lim et al., 2003a). For example, in an effort that identified over 3400 clones representing 80 *C. elegans* miRNAs, only 38 clones representing 14 miRNAs* were found (Lim et al., 2003a). This approximately 100-fold difference in cloning frequency indicates that the miRNA:miRNA* duplex is generally short-lived compared to the miRNA single strand.

According to the current model, the specificity of the initial cleavage mediated by Drosha determines the correct register of cleavage within the miRNA precursor and thus defines both mature ends of the miRNA (Lee et al., 2003). This idea that Drosha, not Dicer, imparts the specificity is appealing because studies have shown that generic double-stranded RNA is refractory to Drosha cleavage and that Dicer progressively chops up an RNA double strand, irrespective of its sequence (Zamore et al., 2000; Bernstein et al., 2001; Elbashir et al., 2001a; Zhang et al., 2002). The determinants of Drosha recognition are largely undefined but include the secondary structure at the base of the primary stem loop as well as some elements flanking the stem loop but generally within 125 nt of the miRNA (Lee et al., 2003; Chen et al., 2004).

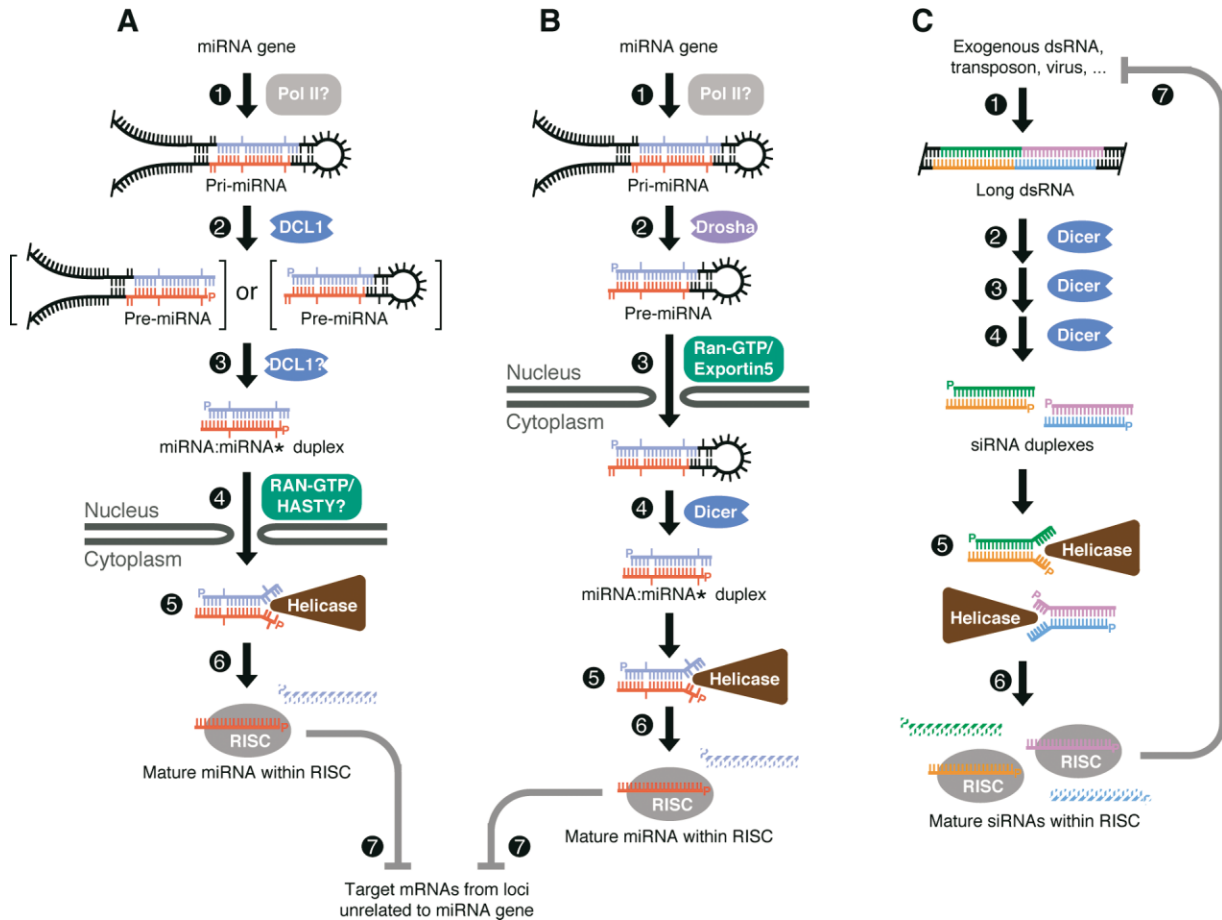This stepwise scenario for miRNA maturation is based

Figure 2. The Biogenesis of miRNAs and siRNAs

(A) The biogenesis of a plant miRNA (steps 1–6; see text for details) and its hetero-silencing of loci unrelated to that from which it originated (step 7). The pre-miRNA intermediates (bracketed), thought to be very short-lived, have not been isolated in plants. The miRNA (red) is incorporated into the RISC (step 6), whereas the miRNA* (blue) is degraded (hatched segment). A monophosphate (P) marks the 5′ terminus of each fragment.

(B) The biogenesis of a metazoan miRNA (steps 1–6; see text for details) and its hetero-silencing of loci unrelated to that from which it originated (step 7).

(C) The biogenesis of animal siRNAs (steps 1–6; see text for details) and their auto-silencing of the same (or similar) loci from which they originated (step 7).

primarily on the investigation of mammalian Drosha and Dicer function (Lee et al., 2002, 2003). The notion that it applies to other metazoan species is supported by the identity of the long form of the *C. elegans lin-4* RNA, which appears to be an excellent match (within the resolution of nuclease mapping) to that expected for the *lin-4* pre-miRNA (Lee et al., 1993). Furthermore, presumed pre-miRNAs for numerous miRNAs can be detected on Northern blots, and when examined in the context of reduced Dicer activity, these pre-miRNAs invariably increase in abundance, as would be expected if Dicer was responsible for their processing (Grishok et al., 2001; Hutvágner et al., 2001; Ketting et al., 2001; Lee and Ambros, 2001; Lim et al., 2003a). Finally, the general existence of the miRNA:miRNA* duplex is supported by the cloning of numerous miRNAs* in nematodes and flies, although for most miRNA genes, an experimentally identified miRNA* has not yet been reported.

The cloning of a few miRNAs* in plants also points to a transient miRNA:miRNA* duplex (Reinhart et al., 2002).

However, the biogenesis of this duplex appears to differ in plants (Figure 2A). Most notably, pre-miRNAs have not been compellingly detected in plants—not even in plants with crippled DCL1, a Dicer-like protein known to assist in miRNA maturation (Reinhart et al., 2002). The lack of pre-miRNA in these *dcl1-9* plants (formerly known as *caf-1* plants), together with the apparent nuclear localization of the DCL1 protein (Papp et al., 2003), suggests that DCL1 provides the Drosha functionality in plants, making the first cut that sets the register for miRNA maturation (Figure 2A, step 2). DCL1 (or another enzyme yet to be identified) then makes the second cut, which corresponds to metazoan Dicer cleavage, before the miRNA leaves the nucleus (Figure 2A, step 3). A coupled second cut in the nucleus would explain why pre-miRNA-like RNAs do not accumulate to detectable levels in plants. It would also explain why ectopic nuclear but not cytoplasmic expression of P19, a plant viral protein that inhibits silencing by sequestering siRNA duplexes, prevents miRNA accumulation (Papp et al.,

2003). Perhaps HASTY, the plant ortholog of Exportin-5, is responsible for exporting the miRNA:miRNA* duplex from the nucleus, which would explain the pleiotropic developmental phenotypes of *hasty* mutants (Bollman et al., 2003; Yi et al., 2003; Lund et al., 2004) (Figure 2A, step 4).

Biogenesis: RISC Assembly
Following cleavage and nucleocytoplasmic export, the miRNA pathway of plants and animals appears to be biochemically indistinguishable from the central steps of RNA silencing pathways known as posttranscriptional gene silencing (PTGS) in plants, quelling in fungi, and RNAi in animals. Indeed, understanding miRNA biogenesis and function has been greatly facilitated by analogy and contrast to the siRNAs of RNAi, and vice versa. In light of these biochemical connections, the discovery of *lin-4* and its regulation of *lin-14* can be considered in hindsight as the first characterization of an RNAi-like phenomenon in animals.

To illustrate the commonality between miRNAs and siRNAs, the RNAi pathway is briefly outlined here (and depicted in Figure 2C). The pathway begins with long double-stranded RNA, either a bimolecular duplex or an extended hairpin, that either is artificially introduced into the cell or animal during a gene knockdown experiment (Fire et al., 1998) or is naturally generated—from sense and antisense genomic transcripts, or perhaps from the activity of a cellular RNA-dependent RNA polymerase (found in plants, fungi, and nematodes, but not flies or mammals) or as an intermediate of viral replication (Cogoni and Macino, 1999; Ketting et al., 1999; Dalmay et al., 2000; Mourrain et al., 2000; Smardon et al., 2000; Aravin et al., 2001, 2003; Li et al., 2002). The double-stranded RNA is processed by Dicer into many ~22 nt siRNAs (Hamilton and Baulcombe, 1999; Hammond et al., 2000; Parrish et al., 2000; Zamore et al., 2000; Grishok et al., 2001; Ketting et al., 2001; Knight and Bass, 2001) (Figure 2C, steps 2–4). Although these siRNAs are initially short double-stranded species with 5′ phosphates and 2 nt 3′ overhangs characteristic of RNase III cleavage products, they eventually become incorporated as single-stranded RNAs into a ribonucleoprotein complex, known as the RNA-induced silencing complex (RISC) (Hammond et al., 2000; Elbashir et al., 2001a, 2001b; Nykänen et al., 2001; Martinez et al., 2002; Schwarz et al., 2002) (Figure 2C, step 6). The RISC identifies target messages based on perfect (or nearly perfect) complementarity between the siRNA and the mRNA, and then the endonuclease of the RISC cleaves the mRNA at a site near the middle of the siRNA complementarity, measuring from the 5′ end of the siRNA and cutting between the nucleotides pairing to residues 10 and 11 of the siRNA (Elbashir et al., 2001a, 2001b). Similar pathways have been proposed for gene silencing in plants and fungi (Hamilton and Baulcombe, 1999; Vance and Vaucheret, 2001; Pickford et al., 2002).

The RISC has been purified from fly and human cells and in both cases contains a member of the Argonaute protein family, which is thought to be a core component of the complex (Hammond et al., 2001; Hutvágner and Zamore, 2002; Martinez et al., 2002). This fits nicely with previous genetic data showing that Argonaute proteins RDE-1, QDE2, and AGO1 are crucial for RNAi and analogous processes in worms, fungi, and plants, respectively (Tabara et al., 1999; Catalanotto et al., 2000; Fagard et al., 2000). Argonaute and its homologs are approximately 100 kDa proteins that are sometimes called PPD proteins because they all share the PAZ and PIWI domains (Cerutti et al., 2000). The PAZ domain (first recognized in Piwi, Argonaute, and Zwille/Pinhead proteins) has a stable fold when isolated from the rest of the protein, which has a β barrel core that together with a side appendage appears to bind weakly to single-stranded RNAs at least 5 nt in length and also to double-stranded RNA (Lingel et al., 2003; Song et al., 2003; Yan et al., 2003). This dual binding ability suggests that the Argonaute protein could be directly associated with the siRNA before and after it recognizes the mRNA target.

Other RISC-associated proteins include the suspected RNA binding proteins VIG and Fragile X-related protein and the nuclease Tudor-SN, none of which have defined roles in the RISC (Caudy et al., 2002, 2003; Ishizuka et al., 2002). These proteins do not copurify with RISC in all purification schemes and their stoichiometry in RISC has not been established. Perhaps they are also core components of the RISC that do not remain associated during some purification methods. Alternatively, they could be accessory factors that modify the specificity or function of the core complex. The notion that RISC comes in different subtypes is already supported by the number of Argonaute family members found in different species, ranging up to 24 in *C. elegans*, and the preferential genetic or biochemical association of different family members with different types of silencing RNAs (Grishok et al., 2001; Caudy et al., 2002; Zilberman et al., 2003). The RISC endonuclease, known as Slicer, has not been identified, suggesting that it might be present in sub-stoichiometric amounts and only recruited after the other components of RISC have found a suitable match to the siRNA. Another possibility is that one of the identified RISC components provides the Slicer activity by means of an unrecognized nuclease domain.

MicroRNAs were first reported to reside in the miRNA ribonucleoprotein complex (miRNP), which in humans includes the proteins eIF2C2, the helicase Gemin3, and Gemin4 (Mourelatos et al., 2002). eIF2C2 is a human Argonaute homolog and was later found to be a constituent of the human siRNA-programmed RISC (Martinez et al., 2002). Furthermore, the human *let-7* miRNA is associated with eIF2C2 and capable of specifying cleavage of an artificial target with perfect complementarity to the miRNA (Hutvágner and Zamore, 2002). Thus, the miRNP possesses the salient properties that define the RISC (Hutvágner and Zamore, 2002), and although it might later be shown to represent a particular subtype of RISC, it is referred to as a RISC in this review. This perspective is further supported by the demonstration that plant miRNAs can direct cleavage of their natural targets (Llave et al., 2002b; Tang et al., 2003) and that siRNAs originally designed to specify cleavage can also mediate translational repression (Doench et al., 2003; Zeng et al., 2003).

When the miRNA strand of the miRNA:miRNA* duplex is loaded into the RISC, the miRNA* appears to be peeled away and degraded. What then is the mechanism
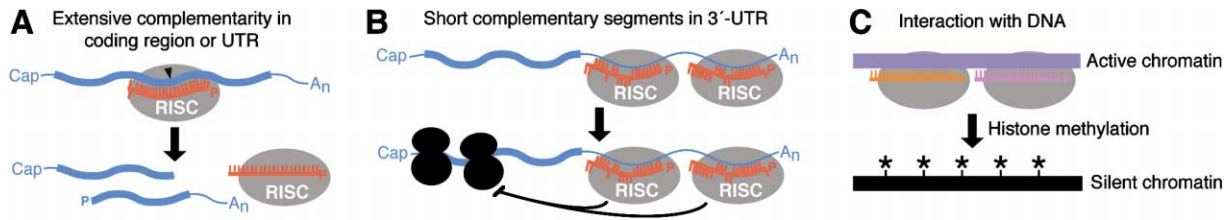
Figure 3. The Actions of Small Silencing RNAs

(A) Messenger RNA cleavage specified by a miRNA or siRNA. Black arrowhead indicates site of cleavage.

(B) Translational repression specified by miRNAs or siRNAs.

(C) Transcriptional silencing, thought to be specified by heterochromatic siRNAs.

for choosing which of the two strands enters the RISC? The answer largely lies in the relative stability of the two ends of the duplex: for both siRNA and miRNA duplexes, the strand that enters the RISC is nearly always the one whose 5′ end is less tightly paired (Khvorova et al., 2003; Schwarz et al., 2003). This observation suggests that a helicase-like enzyme (yet to be identified) samples the ends of the duplex multiple times—usually releasing the end before beginning to productively unwind the duplex but occasionally unwinding the duplex, resulting in a strong bias for productive unwinding at the easier end (Khvorova et al., 2003; Schwarz et al., 2003) (Figures 2A–2C, steps 5). This elegant rule for predicting which strand of the duplex will enter the RISC was initially formulated based on observations and experiments in animal systems, but it also applies to plant siRNAs (Khvorova et al., 2003) and plant miRNAs. Its predictive value for the vast majority of plant and animal miRNAs strongly implies the existence of the miRNA:miRNA* duplex as a transient intermediate in the biogenesis of all miRNAs, even those for which a miRNA* has not yet been cloned. For a few vertebrate and insect genes, both strands of the miRNA duplex accumulate at frequencies suggesting that both enter the RISC, raising the prospect that either or both might be functional (Lagos-Quintana et al., 2002; Krichevsky et al., 2003; Schwarz et al., 2003). These rare cases can be reconciled with the asymmetric loading of the RISC because the ends of these duplexes have nearly equivalent stabilities at their ends; for each RISC assembled, the helicase loads only one strand of each duplex but chooses each strand with similar frequency (Schwarz et al., 2003).

Mechanism: mRNA Cleavage

MicroRNAs can direct the RISC to downregulate gene expression by either of two posttranscriptional mechanisms: mRNA cleavage or translational repression (Figures 3A and 3B). According to the prevailing model, the choice of posttranscriptional mechanisms is not determined by whether the small silencing RNA originated as an siRNA or a miRNA but instead is determined by the identity of the target: Once incorporated into a cytoplasmic RISC, the miRNA will specify cleavage if the mRNA has sufficient complementarity to the miRNA, or it will repress productive translation if the mRNA does not have sufficient complementarity to be cleaved but does have a suitable constellation of miRNA complementary sites (Hutvágner and Zamore, 2002; Zeng et al., 2002, 2003; Doench et al., 2003). Although this model

is generally supported by experimental tests, highly functional siRNAs and metazoan miRNAs have sequence-composition differences centering at positions 12 and 13, which might point to inherent differential sequence preferences for the two respective modes of repression (Khvorova et al., 2003). Furthermore, a perplexing observation has come from the study of a plant miRNA, miR172, which appears to regulate *APETALA2* via translational repression despite the near-perfect complementarity between the miRNA and its single complementary site in the *APETALA2* ORF (Aukerman and Sakai, 2003; Chen, 2003).

When a miRNA guides cleavage, the cut is at precisely the same site as that seen for siRNA-guided cleavage, i.e., between the nucleotides pairing to residues 10 and 11 of the miRNA (Elbashir et al., 2001a; Hutvágner and Zamore, 2002; Llave et al., 2002b; Kasschau et al., 2003). The register of cleavage does not change when the miRNA is not perfectly paired to the target at its 5′ terminus (Kasschau et al., 2003; Palatnik et al., 2003). Therefore, the cut site appears to be determined relative to miRNA residues, not miRNA:target base pairs. After cleavage of the mRNA, the miRNA remains intact and can guide the recognition and destruction of additional messages (Hutvágner and Zamore, 2002; Tang et al., 2003).

Mechanism: Translational Repression

From the beginning, it was proposed that *lin-4* RNA specifies the translational repression of *C. elegans lin-14* mRNA. This is the simplest interpretation of the observation that *lin-4* RNA expression coincides with a drop in LIN-14 protein without a change in *lin-14* mRNA (Wightman et al., 1993). The surprise came later, when it was shown that the polysome profile of *lin-14* mRNA at the first larval stage is indistinguishable from that at later larval stages, when LIN-14 protein levels have dropped (Olsen and Ambros, 1999). The same is true for *lin-28* mRNA, another message targeted by *lin-4* RNA (Seggerson et al., 2002). Two possibilities were put forward to explain these results (Olsen and Ambros, 1999). The *lin-4* RNA might repress translation at a step after translation initiation, in a manner that does not perceivably alter the density of the ribosomes on the message, e.g., by the slowing or stalling of all the ribosomes on the message. An alternative possibility is that translation continues at the same rate but is nonproductive because the newly synthesized polypeptide is specifically degraded. In this review, both of these mechanistic possi-

bilities are lumped together as translational repression, as is common practice, even though in the second possibility polypeptide synthesis per se is not repressed. A better mechanistic understanding of *lin-4*-specified translational repression awaits the development of an in vitro system that faithfully recapitulates *lin-4* regulation of its targets.

Extending the analysis of polysome profiles beyond *C. elegans lin-4* regulation will be important for learning whether the postinitiation mechanism applies more generally to translational repression mediated by other miRNAs. Indeed, evidence for translational repression of any metazoan miRNA targets other than those of *lin-4* is scant because the fate of the messenger RNA during miRNA-mediated regulation has not yet been monitored for these non-*lin-4* targets. Nonetheless, several indirect lines of evidence support the notion that metazoan miRNAs other than *lin-4* RNA typically mediate translational repression rather than mRNA cleavage: First, other metazoan miRNAs, as well as siRNAs, can repress the expression of heterologous reporter transcripts without decreasing mRNA levels, if these messages contain either the natural miRNA complementary sites from the miRNA target (Brennecke et al., 2003) or multiple artificial complementary sites that have bulges or mismatches at their center when paired to the miRNA, such that the pattern of base pairing resembles that found between the *let-7* RNA and its natural complementary sites in the *C. elegans*, *lin-41* 3′ UTR (Zeng et al., 2002, 2003; Doench et al., 2003). Second, the *let-7*-programmed RISC endogenous to human cells does not cleave an RNA fragment containing the *let-7* complementary sites found in *C. elegans lin-41* (Hutvágner and Zamore, 2002). Third, there is a difference between plants and animals with regard to the extent of complementarity between the miRNAs and mRNAs (Rhoades et al., 2002). Because near-perfect complementarity is thought to be required for RISC-mediated cleavage but not translational repression, the lower degree of complementarity seen in animals suggests that translational repression is more prevalent in animals than in plants. Nonetheless, it would be premature to conclude that more metazoan miRNA regulatory targets are translationally inhibited than are cleaved. Surprisingly little complementarity appears to be needed to specify detectable RISC-mediated cleavage in mammalian cells (Jackson et al., 2003), suggesting that it will not be long before natural examples of miRNA-directed mRNA cleavage will be reported in animals.

The cooperative action of multiple RISCs appears to provide the most efficient translational inhibition (Doench et al., 2003). This explains the presence of multiple miRNA complementary sites in most genetically identified targets of metazoan miRNAs (Lee et al., 1993; Wightman et al., 1993; Reinhart et al., 2000; Abrahante et al., 2003; Lin et al., 2003). The computationally identified metazoan targets also have multiple sites, but this pattern is uninformative because the presence of multiple sites was a criterion for their identification (Brennecke et al., 2003; Lewis et al., 2003; Stark et al., 2003). Although only a small fraction of the miRNA-mRNA regulatory pairs are known in any animal, there are already instances in which different miRNA species have been proposed to regulate the same targets (Reinhart et al., 2000; Abrahante et al., 2003; Lin et al., 2003). These examples, and the analogy to other biological regulatory systems, most notably transcriptional regulation, have led to the general expectation that as the list of known metazoan miRNA:mRNA regulatory interactions becomes more comprehensive, combinatorial control will be seen to be common, if not the norm.

The complementary sites for the known metazoan targets reside in the 3′ UTRs. This bias might reflect a mechanistic preference, perhaps enabling the bound complexes to avoid the mRNA-clearing activity of the ribosome. After all, numerous other examples of eukaryotic translation regulation are mediated through 3′ UTR elements (Kuersten and Goodwin, 2003). Alternatively, it might reflect a bias in the way that metazoan miRNA targets and complementary sites are discovered: The *lin-4:lin-14* precedent might have directed subsequent searches to the 3′ UTRs, and conserved complementary sites are easier to distinguish in the UTRs, away from the confounding sequence conservation of the ORFs. The reported siRNA-mediated translational repression from a single imperfect complementary site in the ORF of a mammalian reporter construct (Saxena et al., 2003) illustrates why it would be premature to conclude that most metazoan miRNA regulation is mediated through multiple complementary sites in the 3′ UTRs.

Among the dozens of miRNA-target relationships that have been examined, there has been no evidence for miRNAs directing upregulation of gene expression. These findings are consistent with the idea that miRNAs are all acting within a silencing complex, namely the RISC. Even if miRNAs are limited to functioning within RISC complexes, there is still the prospect that some miRNAs might specify more than just posttranscriptional repression; some might target DNA for transcriptional silencing (Figure 3C). Argonaute proteins and siRNAs are associated with DNA methylation and silencing in plants (Mette et al., 2000; Hamilton et al., 2002; Zilberman et al., 2003), heterochromatin formation in fungi (Hall et al., 2002; Reinhart and Bartel, 2002; Volpe et al., 2002), and DNA rearrangements in ciliates (Mochizuki et al., 2002). Each of these examples suggests the existence of a nuclear RISC-like complex. If miRNAs are not involved in DNA silencing, it will be interesting to learn how they avoid entering the nuclear RISC, particularly in plants, where processing appears to be completed in the nucleus.

Mechanism: Target Recognition

The importance of complementarity to the 5′ portion of metazoan miRNAs has been suspected since the observation that the *lin-14* UTR has "core elements" of complementarity to the 5′ region of the *lin-4* miRNA (Wightman et al., 1993). More recent observations support this idea: (1) Residues 2–8 of several invertebrate miRNAs are perfectly complementary to 3′ UTR elements previously shown to mediate posttranscriptional repression (Lai, 2002). (2) Within the miRNA complementary sites of the first validated targets of invertebrate miRNAs, mRNA residues that pair (sometimes imperfectly) to residues 2–8 of the miRNA are perfectly conserved in orthologous messages of other species, and a contiguous helix of at least six basepairs is nearly always seen in

this region (Stark et al., 2003). (3) Residues 2–8 of the miRNA are the most conserved among homologous metazoan miRNAs (Lewis et al., 2003; Lim et al., 2003a). (4) When predicting targets of mammalian miRNAs, requiring perfect pairing to the heptamer spanning residues 2–8 of the miRNA is much more productive than is requiring pairing to any other heptamer of the miRNA (Lewis et al., 2003). Pairing to this 5′ core region also appears to disproportionally govern the specificity of siRNA-mediated mRNA cleavage (Jackson et al., 2003; Pusch et al., 2003), and the same is true for a plant miRNA that mediates mRNA cleavage (Reinhart, Mallory, Tang, Zamore, Barton, D.B., unpublished).

Why is complementarity to the 5′ end of the small RNA universally important, regardless of the mechanism of gene regulation? One possibility is that the RISC presents only this core region to nucleate pairing to the mRNAs. Presentation of these ∼7 nucleotides prearranged in the geometry of an A-form helix would preferentially enhance the affinity with matched mRNA segments. Presentation of a preformed helical segment of this length would be a reasonable compromise between the topological difficulties associated with longer prearranged helical geometry and the drop in initial binding specificity that would result from a shorter core. In this scenario, mismatches with the core region inhibit initial target recognition and thus prevent cleavage or translational repression regardless of the degree of complementarity elsewhere in the complementary site. If there is sufficient additional pairing after the remainder of the miRNA is allowed to participate, cleavage ensues. However, core pairing supplemented by just a few flanking pairs appears to be sufficient to mediate translational repression in cooperation with other RISCs bound to the message (Lewis et al., 2003). Interestingly, the ability of the Argonaute PAZ domain to bind both double- and single-stranded RNAs (Lingel et al., 2003; Song et al., 2003; Yan et al., 2003), mentioned earlier, would make it a suitable candidate for presenting the core and stabilizing the core pairing.

Mechanism: Distinctions between miRNAs and siRNAs

Because miRNAs and endogenous siRNAs have a shared central biogenesis (Figures 2B and 2C, steps 4–6) and can perform interchangeable biochemical functions (Figures 3A and 3B), these two classes of silencing RNAs cannot be distinguished by either their chemical composition or mechanism of action. Nonetheless, important distinctions can be made, particularly in regard to their origin, evolutionary conservation, and the types of genes that they silence (Figured 2B and 2C, steps 1–3 and 7; Bartel and Bartel, 2003): First, miRNAs derive from genomic loci distinct from other recognized genes, whereas siRNAs often derive from mRNAs, transposons, viruses, or heterochromatic DNA (Figure 2, steps 1). Second, miRNAs are processed from transcripts that can form local RNA hairpin structures, whereas siRNAs are processed from long bimolecular RNA duplexes or extended hairpins (Figure 2, steps 2). Third, a single miRNA:miRNA* duplex is generated from each miRNA hairpin precursor molecule, whereas a multitude of siRNA duplexes are generated from each siRNA precur-

sor molecule, leading to many different siRNAs accumulating from both strands of this extended dsRNA (Figure 2). Fourth, miRNA sequences are nearly always conserved in related organisms, whereas endogenous siRNA sequences are rarely conserved. These types of differences are the basis of practical guidelines for distinguishing and annotating newly discovered miRNAs and endogenous siRNAs (Ambros et al., 2003a).

Although much remains to be learned about the biological targets of miRNAs and endogenous siRNAs, a fifth distinction can be made between these two classes of silencing RNAs: endogenous siRNAs typically specify "auto-silencing," in that they specify the silencing of the same locus (or very similar loci) from which they originate, whereas miRNAs specify "hetero-silencing," in that they are produced from genes that specify the silencing of very different genes (Figure 2, steps 7). Natural examples of auto-silencing include the silencing of viruses, transposons, and the heterochromatic outer repeats of centromeres. Another example is the *Drosophila Su(Ste)* repeats, which generate siRNAs that silence the Su(Ste) repeats themselves as well as the very similar *Stellate* genes (Aravin et al., 2001). At first glance, miR-127 and miR-136 might seem to be exceptions to this principle because they originate from the antisense strand of their presumptive target, the *Rtl1* mRNA (Seitz et al., 2003). However, because these genes lie in an imprinted locus, in which the miRNAs are expressed from the maternal chromosome and the *Rtl1* mRNA is expressed from the paternal chromosome, these miRNAs can still be thought of as specifying hetero-silencing. This fifth distinction explains the greater sequence conservation seen for miRNAs. To the extent that the siRNAs come from the same loci that they target, a mutational event that changes the sequence of the siRNA would also change the sequence of its regulatory target, and siRNA regulation would be preserved—an unusual case of maintaining an important function without selective pressure for conserving the sequence. In contrast, a mutation in a miRNA would rarely be accompanied by simultaneous compensatory changes at the loci of its targets, and thus selection pressure would preserve the miRNA sequence.

With these distinctions between the miRNAs and the endogenous siRNAs in mind, it is perhaps worth considering how to classify the small RNAs that arise from constructs introduced into cells for the purpose of gene knockdown experiments. Small RNAs processed from the extended double-stranded regions of long, inverted repeats are clearly siRNAs. At the other extreme are approximately 22 nt RNAs processed from pre-miRNA-like stem loops. For metazoan cases in which these stem loops include the determinants for the sequential processing by Drosha then Dicer, classification is again simple; these would be artificial miRNAs. However, classification is less clear for RNAs deriving from the short hairpin constructs typically used for knockdowns in mammalian cells (Dykxhoorn et al., 2003), whose processing is unlikely to involve Drosha and even might not involve Dicer.

Function: Regulatory Roles of miRNAs
The most pressing question to arise from the discovery of the hundreds of different miRNAs is, what are all these

Table 1. MicroRNAs and Their Functions: Examples for which Phenotypic Consequences of Disrupted or Ectopic miRNA Regulation Are Known

miRNA	Target Gene(s)	Biological Role of miRNA/Target Gene	Refs
Nematodes			
lin-4 RNA	*Ce lin-14* probable transcription factor	Timing of early larval developmental transitions	1,2
	Ce lin-28 cold shock domain protein	Timing of early larval developmental transitions	3
let-7 RNA	*Ce lin-41* probable RNA-binding protein	Timing of late larval developmental transitions	4,5
	Ce hbl-1 transcription factor	Timing of late larval developmental transitions	6,7
lsy-6 RNA	*Ce cog-1* transcription factor	Left/right asymmetry of chemoreceptor expression	8
Insects			
bantam miRNA	*Dm hid* pro-apoptotic protein	Apoptosis and growth control during development	9
miR-14	unknown	Apoptosis and fat metabolism	10
Mammals			
miR-181	unknown	Hematopoietic differentiation	11
Plants			
miR165/166	*At REV* and relelated transcription factors	Axial meristem initit ion and leaf development	12-14
miR172	*At AP2* and related transcription factors	Flower development; timing transition to flowering	15-18
miR-JAW	*At TCP4* and releated transcription factors	Leaf development, embryonic patterning	19
miR159	*At MYB33* and related transcription factors	Leaf development	12,15,19

Species abbreviations: *Caenorhabditis elegans*, *Ce*; *Drosophila melanogaster*, *Dm*; *Arabidopsis thaliana*, *At*.

1 (Lee et al., 1993); 2 (Wightman et al., 1993); 3 (Moss et al., 1997); 4 (Reinhart et al., 2000); 5 (Slack et al., 2000); 6 (Abrahante et al., 2003); 7 (Lin et al., 2003); 8 (Johnston and Hobert, 2003); 9 (Brennecke et al., 2003); 10 (Xu et al., 2003); 11 (Chen et al., 2004); 12 (Rhoades et al., 2002); 13 (Tang et al., 2003); 14 (Emery et al., 2003); 15 (Park et al., 2002); 16 (Kasschau et al., 2003); 17 (Chen, 2003); 18 (Aukerman and Sakai, 2003); 19 (Palatnik et al., 2003)

tiny noncoding RNAs doing? For *lin-4*, *let-7*, and several other miRNAs identified by forward genetics, crucial clues to their function and regulatory targets came even before their status as noncoding RNA genes was discovered (Meneely and Herman, 1979; Chalfie et al., 1981; Ambros, 1989; Weigel et al., 2000; Hipfner et al., 2002; Aukerman and Sakai, 2003; Brennecke et al., 2003; Johnston and Hobert, 2003; Xu et al., 2003). These and other miRNAs that have reported functions based on in vivo experimentation are listed in Table 1. For some of these cases, function was determined by the phenotypic consequences of a mutated miRNA or an altered miRNA complementary site, either of which can disrupt miRNA regulation. In other cases, function was inferred from the effects of mutations or transgenic constructs that lead to ectopic expression of the miRNA.

For the vast majority of miRNAs, the phenotypic consequences of disrupted or altered miRNA regulation are not known. However, computational approaches are being developed to find the regulatory targets of the miRNAs, providing clues to miRNA function based on the known roles of these targets (Rhoades et al., 2002; Enright et al., 2003; Lewis et al., 2003; Stark et al., 2003). Computationally predicted targets supported by subsequent experiments or independent phylogenetic evidence are listed in Table 2. The experiments supporting the identity of these targets typically fall into two classes. In cases where the miRNA is thought to specify mRNA cleavage, the cleavage products can be reverse-transcribed, cloned, and sequenced; a preponderance of sequences that end precisely at the predicted site of cleavage provides experimental validation that this mRNA is a cleavage target of the complementary miRNA (Llave et al., 2002b; Kasschau et al., 2003; Xie et al., 2003). To enable detection of both translational repression and mRNA cleavage, heterologous reporter assays can be used in which the miRNA complementary sites are fused to a reporter gene and expression is examined relative to control constructs, or in the presence and

absence of the miRNA (Lewis et al., 2003; Stark et al., 2003). Caution is warranted when interpreting reporter assays that involve multimerization of the miRNA complementary site(s) because such an assay succeeded in validating a miRNA complementary site that was mistakenly taken from a gene that was unrelated to the intended target but similarly annotated (Kawasaki and Taira, 2003a, 2003b). A positive result in the heterologous reporter assay indicates that determinants needed for miRNA regulation are indeed present within the mRNA fragment fused to the reporter, which together with evolutionary conservation of both the miRNA and its complementary sites can provide reasonable evidence of a regulatory relationship. Of course, such a hypothesis is considerably strengthened with evidence of coincident expression of the miRNA and its target in the animal or plant, or experiments that examine the effects of manipulating the miRNA or its complementary site in its native in vivo context.

Function: Roles of Plant miRNAs
In plants, miRNAs have a propensity to pair to mRNAs with near-perfect complementarity, enabling convincing targets to be readily predicted for most known plant miRNAs (Rhoades et al., 2002; Bartel and Bartel, 2003). Evolutionary conservation of the miRNA:mRNA pairing in *Arabidopsis* and rice, together with experimental evidence showing that miRNAs can direct cleavage of targeted mRNAs, supports the validity of these predictions (Llave et al., 2002a; Rhoades et al., 2002; Kasschau et al., 2003; Tang et al., 2003). The known plant miRNAs have a remarkable penchant for targeting transcription factor gene families, particularly those with known or suspected roles in developmental patterning or cell differentiation (Rhoades et al., 2002; Tables 1 and 2). This explains the pleiotropic developmental phenotypes of plants mutant in *DCL1(CAF)* and *HEN1*, genes known to influence miRNA accumulation, and *AGO1*, a gene that might be involved in miRNA function (Bohmert et al.,

Table 2. MicroRNAs and Their Functions: Examples for which Strong Evidence for the Validity of the Regulatory Target Has Been Reported but the Phenotypic Consequences of Disrupted or Ectopic miRNA Regulation Are Not Yet Known

miRNA	Target Gene(s)	Biological Role of Target Gene(s)	Refs
Insects			
miR-7	*Dm HLHm3* basic HLH transcriptional repressor	Interprets Notch-mediated decisions in neuronal development	1,2
	Dm hairy basic HLH transcriptional repressor	Interprets Notch-mediated decisions in neuronal development	2
	Dm m4 Brd family protein	Interprets Notch-mediated decisions in neuronal development	2
miR-14 family	*Dm grim* antagonist of caspase inhibitor	Promotes apoptosis	2
	Dm reaper antagonist of caspase inhibitor	Promotes apoptosis	2
	Dm sickle antagonist of caspase inhibitor	Promotes apoptosis	2
Mammals			
miR-1	*Hs* Brain-derived neurotrophic factor (*BDNF*)	Growth factor; neuronal development	3
	Hs Glucose-6-phosphate dehydrogenase (*G6PD*)	Oxidative stress resistance	3
miR-19a	*Hs* PtdIns(3,4,5)P3 phosphatase (*PTEN*)	Tumor suppressor gene	3
miR-23a	*Hs* Stromal cell-derived factor 1 (*SDF-1*)	Growth & localization of hematopoietic progenitor cells	3
	Hs BRN-3b POU-domain transcription factor	Nueronal development	3
miR-26a	*Hs SMAD-1* transcriptional co-modulator	Regulates TGF-dependent gene expression	3
miR-34	*Hs Delta1* transmembrane protein	Activates Notch during cell-fate decisions	3
	Hs Notch1 transmembrane receptor for Delta	Cell-fate decisions during development	3
miR-101	*Hs ENX-1* polycomb gene	Proliferation of hemotpoeitic cells and other gene regulation	3
	Hs N-MYC basic HLH transcription factor	Proto-oncogene; cell differentiation & proliferation	3
miR-130	*Hs* Macrophage colony stimulating factor-1 *(MCSF)*	Mononuclear phagocytic lineage regulation	3
Plants			
miR170/171	*At SCL6-III, -IV* & related transcription factors	Related to genes for root radial patterning	4-7
miR156/157	*At SPL2* & related transcription factors	Related to genes for floral meristem identity	6,8
miR160	*At ARF10, ARF17* & related transcription factors	Related to genes for auxin response & development	6,8
miR167	*At ARF8 & ARF6* transcription factors	Related to genes for auxin response & development	6,8,9
miR164	*At CUC1, CUC2* & related transcription factors	Shoot apical meristem formation & organ separation	6,8
miR169	*At CBF-HAP2* DNA-binding proteins	unknown	6
miR162	*At DCL1* Dicer-like RNase III	miRNA biogenesis	10,11

The metazoan regulatory targets listed were predicted computationally then supported experimentally. The plant regulatory targets listed were predicted computationally then supported with independent phylogenetic and/or experimental evidence. Species abbreviations: *Drosophila melanogaster, Dm*; human, *Hs*; *Arabidopsis thaliana, At*.
1 (Lai, 2002); 2 (Stark et al., 2003); 3 (Lewis et al., 2003); 4 (Llave et al., 2002a); 5 (Reinhart et al., 2002); 6 (Rhoades et al., 2002); 7 (Llave et al., 2002b); 8 (Kasschau et al., 2003); 9 (Park et al., 2002); 10 (Xie et al., 2003); 11 (Bartel and Bartel, 2003)

1998; Jacobsen et al., 1999; Park et al., 2002; Reinhart et al., 2002; Schauer et al., 2002). Of the few predicted plant targets that are not transcription factors, two are *DCL1* and *AGO1*, suggesting a negative feedback mechanism that controls the expression of these genes with known or suspected roles in miRNA biogenesis and function (Rhoades et al., 2002; Bartel and Bartel, 2003; Xie et al., 2003).

Why are there so many targets of the plant miRNAs transcription factors that have been implicated in the control of plant development? The model put forward to answer this question proposes that many plant miRNAs function during cellular differentiation by mediating the degradation of key regulatory gene transcripts in specific daughter cell lineages (Rhoades et al., 2002; Figure 4). For example, during differentiation, certain genes specifying a less differentiated state might need to be turned off. This can be achieved by repressing transcription; however, a gene is not fully off until its message stops making protein. Thus, to more quickly stop expression of such a gene, the differentiating cell can deploy a miRNA that specifies the cleavage of that mRNA. The active clearing of the lingering regulatory messages (or of new messages generated by continued transcription) could enable rapid daughter cell differentiation without having to depend on regulatory genes having constitutively unstable messages. In this respect, miRNA regulation would be analogous to ubiquitin-dependent protein degradation, except that specific mRNAs, rather than proteins, are targeted for degradation.

This model concurs with the observation that a mutation disrupting the miRNA complementary site of *PHB* mRNA leads to a more expansive distribution of the message, as if it were no longer being cleared from cells expressing the miRNA (McConnell et al., 2001; Rhoades et al., 2002). It also explains why so many of the initially identified target genes specify formation and identity of meristem, i.e., plant stem cells (Tables 1 and 2)—these are precisely the genes that would need to be turned off during early differentiation. The model also would apply to scenarios later in differentiation or to cases where the daughter cell is choosing among two or more differentiated states, which would explain the targeting of the other transcripts that have regulatory roles later in development. One point of caution in trying to deduce the general roles of plant miRNAs is that the known set of plant miRNAs is enriched in the more abundant miRNAs of plant tissues and organs and thus might not be representative. For example, miRNAs specifying an undifferentiated state would have been less likely to be cloned because most cells of plant organs are typically differentiated.

Function: Roles of Animal miRNAs

Computational methods have recently been developed to identify the targets of *Drosophila* and mammalian miRNAs (Enright et al., 2003; Lewis et al., 2003; Stark et al., 2003). These methods search for multiple conserved regions of miRNA complementarity within 3′ UTRs. Iden-

Precursor cell expressing
transcription factor mRNAs

❶

❷

❸

Daughter cells with distinct
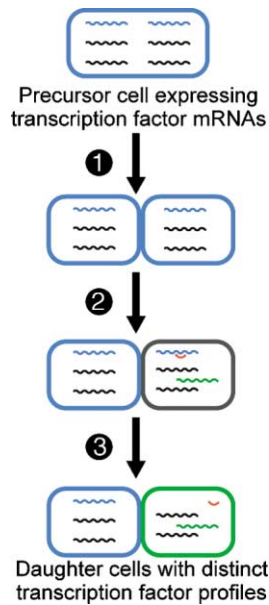transcription factor profiles

Figure 4. Working Model for the Roles of miRNAs that Target the Messages of Transcription Factors during Plant Development

Following cell division, the daughter cells inherit mRNAs from the precursor cell (step 1). A differentiating daughter cell (cell on right) expresses new transcription factor messages (green) as well as a miRNA (red) complementary to messages that must be cleared (blue) in order for the cell to progress to the differentiated state (step 2). The miRNA directs the cleavage of target messages, preventing prolonged or inappropriate expression of the transcriptional regulator, thus enabling the rapid differentiation of the daughter cell (step 3). (Figure redrawn from Rhoades et al., 2002, copyrighted by Cell press, used with permission.)

tifying targets in animals has been a more difficult task than in plants because in animals there are far fewer mRNAs with near-perfect complementarity to miRNAs. This makes the analysis noisier—much more prone to false positives. Furthermore, evolutionary conservation was used as a criterion for target identification in animals, and thus it could not be used as a means to independently validate the targets. Nonetheless, the experimental support achieved for a majority of the predictions tested is encouraging (Table 2), and there are compelling reasons to take seriously the remaining untested predictions. For example, in one of the fly studies, there were striking clusters of functionally related genes among the top predictions (Stark et al., 2003). The most notable examples were Notch target genes for miR-7, proapoptotic genes for miR-2, and a set of enzymes involved in branched-chain amino acid degradation for miR-277. In the mammalian study, over 400 regulatory targets were predicted when using parameter cutoffs that gave a signal-to-noise ratio of 3.2:1 (Lewis et al., 2003). This signal:noise ratio was seen only when restricting the miRNAs to those most conserved among mammals and fish, and only when demanding perfect complementarity to the most conserved portion of miRNAs (the 7 nt core segment comprising residues 2–8 of the miRNAs), observations that would be exceedingly difficult to explain if most of the identified messages were not relevant targets of the miRNAs.

The ability to identify hundreds of miRNA targets with confidence that most of the predicted targets are authentic enables the analysis of the types of genes most commonly targeted by mammalian miRNAs (Lewis et al., 2003). As in plants, the predicted targets are significantly enriched in genes involved in transcriptional regulation, suggesting that the model proposed for the roles of many plant miRNAs (Figure 4) could also be operating in animals. Nonetheless, this enrichment for transcriptional regulators is far less pronounced in mammals, and only a minority of the predicted mammalian targets are involved in development. The predicted targets represent a surprisingly broad diversity of molecular functions and biological processes. Thus, in contrast to the plant miRNAs, most mammalian miRNAs do not appear to be primarily involved at the upper levels of the gene regulatory cascades but instead appear to be operating at many levels to regulate the expression of a diverse set of genes, many of which do not go on to directly influence the expression of other genes (Lewis et al., 2003).

Function: The Question of Specificity

Although current lists of predicted miRNA targets provide insights and hypotheses for thousands of follow-up experiments, they could be far from comprehensive. For example, in the animal studies, the computational methods used evolutionary conservation to distinguish miRNA target sites from the multitude of 3′ UTR segments that otherwise would score equally well with regard to the quality and stability of base pairing (Lewis et al., 2003; Stark et al., 2003). The cell, on the other hand, cannot use the filter of evolutionary conservation to choose among the possibilities. Does this mean that many of these other mRNAs would in fact be targeted if expressed in the same cells as the cognate miRNAs? Perhaps not—perhaps miRNA base pairing is not the only major determinant of specificity. Proteins or mRNA structure could restrict miRNP accessibility to the UTRs. But if this were generally true, siRNA knockdown experiments might be expected to have a much lower success rate. Proteins or mRNA structure could also facilitate recognition of the authentic mRNA targets by means of elements in the mRNAs that have thus far escaped detection. One candidate for such a protein is the Fragile X-related protein, a *Drosophila* RISC component that is related to proteins known to bind specific mRNAs (Caudy et al., 2002; Ishizuka et al., 2002).

The alternative idea—that the quality and stability of base pairing is in fact the primary determinant of specificity—should also be considered. After all, this complementarity requirement includes a 7 nt perfect or near-perfect core match near the 5′ terminus of the miRNA (Lai, 2002; Lewis et al., 2003; Stark et al., 2003), which by itself would represent a degree of specificity comparable to that of the DNA sites recognized by many transcription factors. Pairing outside the 7 nt core site, although perhaps less important than once thought, provides means of conferring added specificity. Just as chromatin structure limits the possibilities for transcription-factor binding, the restricted set of genes transcribed in each cell limits which genes of the genome will be under miRNA control in that cell. And in

the same way that the cooperative action of multiple transcription factors increases the specificity of their control, the cooperative action of homotypic and heterotypic miRNA:UTR interactions would provide an additional mechanism to increase specificity of miRNA control. Despite these mechanisms for increasing the regulatory specificity, the notion that target-site recognition is primarily determined by multiple instances of 7 nt core complementarity would imply that miRNAs influence the expression of a remarkably large number of different mRNAs (Lewis et al., 2003).

The "many targets" hypothesis is embraced and partially rationalized in a proposal that the miRNA milieu, unique to each cell type, provides important context for the evolution of all mRNA sequences and is productively used to dampen the utilization of thousands of mRNAs (D.B. and C.-Z. Chen, unpublished). For mRNAs that should not be expressed in a particular cell type, miRNAs reduce protein production to inconsequential levels. The result is equivalent to a discrete off switch, and thus these messages, which include targets of Table 1, can be thought of as "switch targets." In addition to these classical targets, at least three other categories of mRNAs can be imagined: For messages called "tuning targets," miRNAs could adjust protein output in a manner that allows for customized expression in different cell types yet a more uniform level within each cell type. Other mRNAs could be simply bystanders, "neutral targets," for which downregulation by miRNAs is tolerated or is negated by feedback processes. Finally, when thinking about the effects of the miRNA milieu on the evolution of mRNA sequences, it is also useful to consider "antitargets," messages under selective pressure to avoid fortuitous complementarity to the multitude of miRNAs in the cells where they are expressed, either because such complementarity would inappropriately dampen their expression or because it would titrate the miRNAs away from their proper targets.

While molecular biologists will have their hands full identifying and characterizing additional instances where miRNAs are playing the classical role of discrete gene regulatory switches, computational and systems biologists will have to contend with the prospect that a substantial fraction of all animal mRNAs could have their precise level of expression defined by miRNA regulation. To the extent that the miRNAs direct translational repression rather than mRNA cleavage, this regulation will be invisible to the most powerful tool of the systems biologist, microarray analysis of mRNA levels. Nonetheless, in only two years since the abundance of miRNA genes was reported, there has been rapid progress in cataloging the miRNA genes, determining their expression patterns, and identifying their regulatory targets, providing hope that the goal of accurately integrating their function into models of metazoan gene regulatory circuitry can one day be realized.

Acknowledgments

References

Abrahante, J.E., Daul, A.L., Li, M., Volk, M.L., Tennessen, J.M., Miller, E.A., and Rougvie, A.E. (2003). The Caenorhabditis elegans hunchback-like gene lin-57/hbl-1 controls developmental time and is regulated by microRNAs. Dev. Cell 4, 625–637.

Ambros, V. (1989). A hierarchy of regulatory genes controls a larva-to-adult developmental switch in C. elegans. Cell 57, 49–57.

Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S., Marshall, M., et al. (2003a). A uniform system for microRNA annotation. RNA 9, 277–279.

Ambros, V., Lee, R.C., Lavanway, A., Williams, P.T., and Jewell, D. (2003b). MicroRNAs and other tiny endogenous RNAs in C. elegans. Curr. Biol. 13, 807–818.

Aravin, A.A., Naumova, N.M., Tulin, A.V., Vagin, V.V., Rozovsky, Y.M., and Gvozdev, V.A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in Drosophila melanogaster germline. Curr. Biol. 11, 1017–1027.

Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., and Tuschl, T. (2003). The small RNA profile during Drosophila melanogaster development. Dev. Cell 5, 337–350.

Aukerman, M.J., and Sakai, H. (2003). Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. Plant Cell 10, 10.

Bartel, B., and Bartel, D.P. (2003). MicroRNAs: At the root of plant development? Plant Physiol. 132, 709–717.

Bashirullah, A., Pasquinelli, A.E., Kiger, A.A., Perrimon, N., Ruvkun, G., and Thummel, C.S. (2003). Coordinate regulation of small temporal RNAs at the onset of Drosophila metamorphosis. Dev. Biol. 259, 1–8.

Basyuk, E., Suavet, F., Doglio, A., Bordonne, R., and Bertrand, E. (2003). Human let-7 stem-loop precursors harbor features of RNase III cleavage products. Nucleic Acids Res. 31, 6593–6597.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409, 295–296.

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., and Benning, C. (1998). AGO1 defines a novel locus of Arabidopsis controlling leaf development. EMBO J. 17, 170–180.

Bollman, K.M., Aukerman, M.J., Park, M.Y., Hunter, C., Berardini, T.Z., and Poethig, R.S. (2003). HASTY, the Arabidopsis ortholog of exportin 5/MSN5, regulates phase change and morphogenesis. Development 130, 1493–1504.

Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B., and Cohen, S.M. (2003). bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. Cell 113, 25–36.

Calin, G.A., Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., et al. (2002). Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc. Natl. Acad. Sci. USA 99, 15524–15529.

Catalanotto, C., Azzalin, G., Macino, G., and Cogoni, C. (2000). Gene silencing in worms and fungi. Nature 404, 245.

Caudy, A.A., Myers, M., Hannon, G.J., and Hammond, S.M. (2002). Fragile X-related protein and VIG associate with the RNA interference machinery. Genes Dev. 16, 2491–2496.

Caudy, A.A., Ketting, R.F., Hammond, S.M., Denli, A.M., Bathoorn, A.M., Tops, B.B., Silva, J.M., Myers, M.M., Hannon, G.J., and Plasterk, R.H. (2003). A micrococcal nuclease homologue in RNAi effector complexes. Nature 425, 411–414.

Cerutti, L., Mian, N., and Bateman, A. (2000). Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain. Trends Biochem. Sci. 25, 481–482.

Chalfie, M., Horvitz, H.R., and Sulston, J.E. (1981). Mutations that lead to reiterations in the cell lineages of C. elegans. Cell 24, 59–69.

Chen, X. (2003). A MicroRNA as a Translational Repressor of APET-

ALA2 in Arabidopsis Flower Development. Science. Published online September 11, 2003. 10.1126/science.1088060.

Chen, C.Z., Li, L., Lodish, H.F., and Bartel, D.P. (2004). MicroRNAs modulate hematopoietic lineage differentiation. Science 303, 83–86.

Cogoni, C., and Macino, G. (1999). Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. Nature 399, 166–169.

Dalmay, T., Hamilton, A., Rudd, S., Angell, S., and Baulcombe, D.C. (2000). An RNA-dependent RNA polymerase in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. Cell 101, 543–553.

Doench, J.G., Peterson, C.P., and Sharp, P.A. (2003). siRNAs can function as miRNAs. Genes Dev. 17, 438–442.

Dostie, J., Mourelatos, Z., Yang, M., Sharma, A., and Dreyfuss, G. (2003). Numerous microRNPs in neuronal cells containing novel microRNAs. RNA 9, 631–632.

Dykxhoorn, D.M., Novina, C.D., and Sharp, P.A. (2003). Killing the messenger: short RNAs that silence gene expression. Nat. Rev. Mol. Cell Biol. 4, 457–467.

Elbashir, S.M., Leneckel, W., and Tuschl, T. (2001a). RNA interference is mediated by 21- and 22- nucleotide RNAs. Genes Dev. 15, 188–200.

Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. (2001b). Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. EMBO J. 20, 6877–6888.

Emery, J.F., Floyd, S.K., Alvarez, J., Eshed, Y., Hawker, N.P., Izhaki, A., Baum, S.F., and Bowman, J.L. (2003). Radial patterning of arabidopsis shoots by class III HD-ZIP and KANADI genes. Curr. Biol. 13, 1768–1774.

Enright, A.J., John, B., Gaul, U., Tuschl, T., Sander, C., and Marks, D.S. (2003). MicroRNA targets in Drosophila. Genome Biol. 5, R1.

Fagard, M., Boutet, S., Morel, J.B., Bellini, C., and Vaucheret, H. (2000). AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. Proc. Natl. Acad. Sci. USA 97, 11650–11654.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806–811.

Gauwerky, C.E., Huebner, K., Isobe, M., Nowell, P.C., and Croce, C.M. (1989). Activation of MYC in a masked t(8;17) translocation results in an aggressive B-cell leukemia. Proc. Natl. Acad. Sci. USA 86, 8867–8871.

Grad, Y., Aach, J., Hayes, G.D., Reinhart, B.J., Church, G.M., Ruvkun, G., and Kim, J. (2003). Computational and experimental identification of C. elegans microRNAs. Mol. Cell 11, 1253–1263.

Griffiths-Jones, S. (2004). The microRNA registry. Nucleic Acids Res. 32, D109–D111.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell 106, 23–34.

Hall, I.M., Shankaranarayana, G.D., Noma, K., Ayoub, N., Cohen, A., and Grewal, S.I. (2002). Establishment and maintenance of a heterochromatin domain. Science 297, 2232–2237.

Hamilton, A.J., and Baulcombe, D.C. (1999). A novel species of small antisense RNA in posttranscriptional gene silencing. Science 286, 950–952.

Hamilton, A., Voinnet, O., Chappell, L., and Baulcombe, D. (2002). Two classes of short interfering RNA in RNA silencing. Embo J. 21, 4671–4679.

Hammond, S.C., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates posttranscriptional gene silencing in Drosophila cells. Nature 404, 293–296.

Hammond, S.M., Boettcher, S., Caudy, A.C., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. Science 293, 1146–1150.

Hipfner, D.R., Weigmann, K., and Cohen, S.M. (2002). The bantam gene regulates Drosophila growth. Genetics 161, 1527–1537.

Houbaviy, H.B., Murray, M.F., and Sharp, P.A. (2003). Embryonic stem cell-specific microRNAs. Dev. Cell 5, 351–358.

Hutvágner, G., and Zamore, P.D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. Science 297, 2056–2060.

Hutvágner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science 293, 834–838.

Ishizuka, A., Siomi, M.C., and Siomi, H. (2002). A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. Genes Dev. 16, 2497–2508.

Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S.V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P.S. (2003). Expression profiling reveals off-target gene regulation by RNAi. Nat. Biotechnol. 21, 635–637.

Jacobsen, S.E., Running, M.P., and Meyerowitz, E.M. (1999). Disruption of an RNA helicase/RNAseIII gene in Arabidopsis causes unregulated cell division in floral meristems. Development 126, 5231–5243.

Johnson, S.M., Lin, S.Y., and Slack, F.J. (2003). The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. Dev. Biol. 259, 364–379.

Johnston, R.J., and Hobert, O. (2003). A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans. Nature 426, 845–849.

Kasschau, K.D., Xie, Z., Allen, E., Llave, C., Chapman, E.J., Krizan, K.A., and Carrington, J.C. (2003). P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA function. Dev. Cell 4, 205–217.

Kawasaki, H., and Taira, K. (2003a). Hes1 is a target of microRNA-23 during retinoic-acid-induced neuronal differentiation of NT2 cells. Nature 423, 838–842.

Kawasaki, H., and Taira, K. (2003b). retraction: Hes1 is a target of microRNA-23 during retinoic-acid-induced neuronal differentiation of NT2 cells. Nature 426, 100.

Ketting, R.F., Haverkamp, T.H., vanLuenen, H.G., and Plasterk, R.H. (1999). Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner sydrome helicase and RNaseD. Cell 99, 133–141.

Ketting, R.F., Fischer, S.E.J., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H.A. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev. 15, 2654–2659.

Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. Cell 115, 209–216.

Kim, J., Krichevsky, A., Grad, Y., Hayes, G.D., Kosik, K.S., Church, G.M., and Ruvkun, G. (2003). Identification of many microRNAs that copurify with polyribosomes in mammalian neurons. Proc. Natl. Acad. Sci. USA. Published online December 22, 2003, 10.1073/pnas.2333854100.

Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development, in caenorhabditis elegens. Science 293, 2269–2271.

Krichevsky, A.M., King, K.S., Donahue, C.P., Khrapko, K., and Kosik, K.S. (2003). A microRNA array reveals extensive regulation of microRNAs during brain development. RNA 9, 1274–1281.

Kuersten, S., and Goodwin, E.B. (2003). The power of the 3′ UTR: translational control and development. Nat. Rev. Genet. 4, 626–637.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science 294, 853–858.

Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., and Tuschl, T. (2002). Identification of tissue-specific microRNAs from mouse. Curr. Biol. 12, 735–739.

Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A., and

Tuschl, T. (2003). New microRNAs from mouse and human. RNA 9, 175–179.

Lai, E.C. (2002). MicroRNAs are complementary to 3′UTR motifs that mediate negative post-transcriptional regulation. Nat. Genet. 30, 363–364.

Lai, E.C., Tomancak, P., Williams, R.W., and Rubin, G.M. (2003). Computational identification of Drosophila microRNA genes. Genome Biol 4:R42, 1–20.

Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science 294, 858–862.

Lee, R.C., and Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. Science 294, 862–864.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75, 843–854.

Lee, Y., Jeon, K., Lee, J.T., Kim, S., and Kim, V.N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. EMBO J. 21, 4663–4670.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V.N. (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature 425, 415–419.

Lewis, B.P., Shih, I., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. (2003). Prediction of mammalian microRNA targets. Cell 115, 787–798.

Li, H., Li, W.X., and Ding, S.W. (2002). Induction and suppression of RNA silencing by an animal virus. Science 296, 1319–1321.

Lim, L.P., Lau, N.C., Weinstein, E.G., Abdelhakim, A., Yekta, S., Rhoades, M.W., Burge, C.B., and Bartel, D.P. (2003a). The microRNAs of Caenorhabditis elegans. Genes Dev. 17, 991–1008.

Lim, L.P., Glasner, M.E., Yekta, S., Burge, C.B., and Bartel, D.P. (2003b). Vertebrate microRNA genes. Science 299, 1540.

Lin, S.Y., Johnson, S.M., Abraham, M., Vella, M.C., Pasquinelli, A., Gamberi, C., Gottlieb, E., and Slack, F.J. (2003). The C. elegans hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. Dev. Cell 4, 639–650.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2003). Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. Nature 426, 465–469.

Llave, C., Kasschau, K.D., Rector, M.A., and Carrington, J.C. (2002a). Endogenous and silencing-associated small RNAs in plants. Plant Cell 14, 1605–1619.

Llave, C., Xie, Z., Kasschau, K.D., and Carrington, J.C. (2002b). Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. Science 297, 2053–2056.

Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science 303, 95–98.

Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell 110, 563–574.

McConnell, J.R., Emery, J., Eshed, Y., Bao, N., Bowman, J., and Barton, M.K. (2001). Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots. Nature 411, 709–713.

Meneely, P.M., and Herman, R.K. (1979). Lethals, steriles and deficiencies in a region of the X chromosome of Caenorhabditis elegans. Genetics 92, 99–115.

Mette, M.F., Aufsatz, W., Winden, J.v.d., Matzke, M.A., and Matzke, A.J. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. EMBO J. 19, 5194–5201.

Mette, M.F., van der Winden, J., Matzke, M., and Matzke, A.J. (2002). Short RNAs can identify new candidate transposable element families in Arabidopsis. Plant Physiol. 130, 6–9.

Michael, M.Z., O'Connor, S.M., van Holst Pellekaan, N.G., Young, G.P., and James, R.J. (2003). Reduced accumulation of specific microRNAs in colorectal neoplasia. Mol. Cancer Res. 1, 882–91.

Mochizuki, K., Fine, N.A., Fujisawa, T., and Gorovsky, M.A. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. Cell 110, 689–699.

Moss, E.G., Lee, R.C., and Ambros, V. (1997). 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.

Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Genes Dev. 16, 720–728.

Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.B., Jouette, D., Lacombe, A.M., Nikic, S., Picault, N., et al. (2000). Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell 101, 533–542.

Nykänen, A., Haley, B., and Zamore, P.D. (2001). ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell 107, 309–321.

Ohler, U., Yekta, S., Lim, L.P., Bartel, D.P., and Burge, C.B. (2004). Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification. RNA, in press.

Olsen, P.H., and Ambros, V. (1999). 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.

Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J.C., and Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. Nature 425, 257–263.

Papp, I., Mette, M.F., Aufsatz, W., Daxinger, L., Schauer, S.E., Ray, A., van der Winden, J., Matzke, M., and Matzke, A.J. (2003). Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors. Plant Physiol. 132, 1382–1390.

Park, W., Li, J., Song, R., Messing, J., and Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Curr. Biol. 12, 1484–1495.

Parrish, S., Fleenor, J., Xu, S., Mello, C., and Fire, A. (2000). Functional anatomy of a dsRNA trigger: differential requirement for the two trigger strands in RNA interference. Mol. Cell 6, 1077–1087.

Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M., Maller, B., Srinivasan, A., Fishman, M., Hayward, D., Ball, E., et al. (2000). Conservation across animal phylogeny of the sequence and temporal regulation of the 21 nucleotide let-7 heterochronic regulatory RNA. Nature 408, 86–89.

Pickford, A.S., Catalanotto, C., Cogoni, C., and Macino, G. (2002). Quelling in Neurospora crassa. Adv. Genet. 46, 277–303.

Pusch, O., Boden, D., Silbermann, R., Lee, F., Tucker, L., and Ramratnam, B. (2003). Nucleotide sequence homology requirements of HIV-1-specific short hairpin RNA. Nucleic Acids Res. 31, 6444–6449.

Reinhart, B.J., and Bartel, D.P. (2002). Small RNAs correspond to centromere heterochromatic repeats. Science 297, 1831.

Reinhart, B.J., Slack, F.J., Basson, M., Bettinger, J.C., Pasquinelli, A.E., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. (2000). The 21 nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 403, 901–906.

Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B., and Bartel, D.P. (2002). MicroRNAs in plants. Genes Dev. 16, 1616–1626.

Rhoades, M.W., Reinhart, B.J., Lim, L.P., Burge, C.B., Bartel, B., and Bartel, D.P. (2002). Prediction of plant microRNA targets. Cell 110, 513–520.

Saxena, S., Jonsson, Z.O., and Dutta, A. (2003). Small RNAs with imperfect match to endogenous mRNA repress translation. Implications for off-target activity of small inhibitory RNA in mammalian cells. J. Biol. Chem. 278, 44312–44319.

Schauer, S.E., Jacobsen, S.E., Meinke, D.W., and Ray, A. (2002). DICER-LIKE1: blind men and elephants in Arabidopsis development. Trends Plant Sci. 7, 487–491.

Schwarz, D.S., Hutvágner, G., Haley, B., and Zamore, P.D. (2002). Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. Mol. Cell 10, 537–548.

Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore,

P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. Cell *115*, 199–208.

Seggerson, K., Tang, L., and Moss, E.G. (2002). Two genetic circuits repress the Caenorhabditis elegans heterochronic gene lin-28 after translation initiation. Dev. Biol. *243*, 215–225.

Seitz, H., Youngson, N., Lin, S.P., Dalbert, S., Paulsen, M., Bachellerie, J.P., Ferguson-Smith, A.C., and Cavaille, J. (2003). Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene. Nat. Genet. *34*, 261–262.

Sempere, L.F., Sokol, N.S., Dubrovsky, E.B., Berger, E.M., and Ambros, V. (2003). Temporal regulation of microRNA expression in Drosophila melanogaster mediated by hormonal signals and broad-Complex gene activity. Dev. Biol. *259*, 9–18.

Slack, F.J., Basson, M., Liu, Z., Ambros, V., Horvitz, H.R., and Ruvkun, G. (2000). 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.

Smardon, A., Spoerke, J.M., Stacey, S.C., Klein, M.E., Mackin, N., and Maine, E.M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr. Biol. *10*, 169–178.

Song, J.J., Liu, J., Tolia, N.H., Schneiderman, J., Smith, S.K., Martienssen, R.A., Hannon, G.J., and Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. Nat. Struct. Biol. *10*, 1026–1032.

Stark, A., Brennecke, J., Russell, R.B., and Cohen, S.M. (2003). Identification of Drosophila microRNA targets. PLOS Biol. *1*, E60.

Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. Cell *99*, 123–132.

Tang, G., Reinhart, B.J., Bartel, D.P., and Zamore, P.D. (2003). A biochemical framework for RNA silencing in plants. Genes Dev. *17*, 49–63.

Vance, V., and Vaucheret, H. (2001). RNA silencing in plants - defense and counterdefense. Science *292*, 2277–2280.

Volpe, T., Kidner, C., Hall, I., Teng, G., Grewal, S., and Martienssen, R. (2002). Heterochromatic silencing and histone H3 lysine 9 methylation are regulated by RNA interference. Science *297*, 1833–1837.

Weigel, D., Ahn, J.H., Blazquez, M.A., Borevitz, J.O., Christensen, S.K., Fankhauser, C., Ferrandiz, C., Kardailsky, I., Malancharuvil, E.J., Neff, M.M., et al. (2000). Activation tagging in Arabidopsis. Plant Physiol. *122*, 1003–1013.

Wightman, B., Burglin, T.R., Gatto, J., Arasu, P., and Ruvkun, G. (1991). Negative regulatory sequences in the lin-14 3′-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development. Genes Dev. *5*, 1813–1824.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. Cell *75*, 855–862.

Xie, Z., Kasschau, K.D., and Carrington, J.C. (2003). Negative feedback regulation of Dicer-like1 in Arabidopsis by microRNA-guided mRNA degradation. Curr. Biol. *13*, 784–789.

Xu, P., Vernooy, S.Y., Guo, M., and Hay, B.A. (2003). The Drosophila microRNA mir-14 suppresses cell death and is required for normal fat metabolism. Curr. Biol. *13*, 790–795.

Yan, K.S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M.M. (2003). Structure and conserved RNA binding of the PAZ domain. Nature *426*, 468–474.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev. *17*, 3011–3016.

Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell *101*, 25–33.

Zeng, Y., and Cullen, B.R. (2003). Sequence requirements for micro RNA processing and function in human cells. RNA *9*, 112–123.

Zeng, Y., Wagner, E.J., and Cullen, B.R. (2002). Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. Mol. Cell *9*, 1327–1333.

Zeng, Y., Yi, R., and Cullen, B.R. (2003). MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. Proc. Natl. Acad. Sci. USA *100*, 9779–9784.

Zhang, H., Kolb, F.A., Brondani, V., Billy, E., and Filipowicz, W. (2002). Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. EMBO J. *21*, 5875–5885.

Zilberman, D., Cao, X., and Jacobsen, S.E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science *299*, 716–719.