MicroRNAs: SMALL RNAs WITH A BIG ROLE IN GENE REGULATION

Lin He and Gregory J. Hannon

MicroRNAs are a family of small, non-coding RNAs that regulate gene expression in a sequence-specific manner. The two founding members of the microRNA family were originally identified in *Caenorhabditis elegans* as genes that were required for the timed regulation of developmental events. Since then, hundreds of microRNAs have been identified in almost all metazoan genomes, including worms, flies, plants and mammals. MicroRNAs have diverse expression patterns and might regulate various developmental and physiological processes. Their discovery adds a new dimension to our understanding of complex gene regulatory networks.

RNA INTERFERENCE (RNAi). A form of post-transcriptional gene silencing, in which dsRNA induces degradation of the homologous mRNA, mimicking the effect of the reduction, or loss, of gene activity.

Non-coding RNAs participate in a surprisingly diverse collection of regulatory events, ranging from copy-number control in bacteria[1] to X-chromosome inactivation in mammals[2]. MicroRNAs (miRNAs) are a family of 21–25-nucleotide small RNAs that, at least for those few that have characterized targets, negatively regulate gene expression at the post-transcriptional level[3–5]. Members of the miRNA family were initially discovered as small temporal RNAs (stRNAs) that regulate developmental transitions in *Caenorhabditis elegans*[6]. Over the past few years, it has become clear that stRNAs were the prototypes of a large family of small RNAs, miRNAs, that now claim hundreds of members in worms, flies, plants and mammals. The functions of miRNAs are not limited to the regulation of developmentally timed events. Instead, they have diverse expression patterns and probably regulate many aspects of development and physiology[3,4,7–9]. Although the mechanisms through which miRNAs regulate their target genes are largely unknown, the finding that at least some miRNAs feed into the RNA INTERFERENCE (RNAi) pathway has provided a starting point in our journey to understand the biological roles of miRNAs.

In this review, we revisit the history of miRNAs and summarize recent findings in miRNA biogenesis, translational repression and biological function. We conclude by highlighting the continuing genome-wide efforts to identify novel miRNAs and to predict their targets.

The discovery of miRNAs

The founding member of the miRNA family, *lin-4*, was identified in *C. elegans* through a genetic screen for defects in the temporal control of post-embryonic development[10,11]. In *C. elegans*, cell lineages have distinct characteristics during 4 different larval stages (L1–L4). Mutations in *lin-4* disrupt the temporal regulation of larval development, causing L1 (the first larval stage)-specific cell-division patterns to reiterate at later developmental stages[10]. Opposite developmental phenotypes — omission of the L1 cell fates and premature development into the L2 stage — are observed in worms that are deficient for *lin-14* (REF. 12). Even before the molecular identification of *lin-4* and *lin-14*, these loci were placed in the same regulatory pathway on the basis of their opposing phenotypes and antagonistic genetic interactions[11]. Most genes identified from mutagenesis screens are protein-coding, but *lin-4* encodes a 22-nucleotide non-coding RNA that is partially complementary to 7 conserved sites located in the 3′-untranslated region (UTR) of the *lin-14* gene (FIG. 1b)[13,14]. *lin-14* encodes a nuclear protein, downregulation of which at the end of the first larval stage initiates the developmental progression into the second larval stage[13,15]. The negative regulation of LIN-14 protein expression requires an intact 3′ UTR of its mRNA[14], as well as a functional *lin-4* gene[13]. These genetic interactions inspired a series of molecular and biochemical studies demonstrating that

Cold Spring Harbor Laboratory, Watson School of Biological Sciences, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA. Correspondence to G.J.H. e-mail: hannon@cshl.edu

the direct, but imprecise, base pairing between *lin-4* and the *lin-14* 3′ UTR was essential for the ability of *lin-4* to control LIN-14 expression through the regulation of protein synthesis[16–18]. Through an analogous mechanism, *lin-4* also negatively regulates the translation of *lin-28*, a cold-shock-domain protein that initiates the developmental transition between the L2 and L3 stages[19]. Compared with *lin-14*, *lin-28* has fewer *lin-4* binding sites, which might lead to its translational repression being delayed following *lin-4* expression owing to less efficient *lin-4* binding[6,19].

The discovery of *lin-4* and its target-specific translational inhibition hinted at a new mechanism of gene regulation during development. In 2000, almost 7 years after the initial identification of *lin-4*, the second miRNA, *let-7*, was discovered, also using forward genetics in worms. *let-7* encodes a temporally regulated 21-nucleotide small RNA that controls the developmental transition from the L4 stage into the adult stage[20–22]. Similar to *lin-4*, *let-7* performs its function by binding to the 3′ UTR of *lin-41* and *hbl-1* (*lin-57*), and inhibiting their translation[20–24]. The identification of *let-7* not only provided another vivid example of developmental regulation by small RNAs, but also raised the possibility that such RNAs might be present in species other than nematodes. Unlike *lin-4*, the orthologues of which in flies and mammals initially escaped bioinformatic searches, and were only recognized recently[25,26], both *let-7* and *lin-41* are evolutionarily conserved throughout metazoans, with homologues that were readily detected in molluscs, sea urchins, flies, mice and humans[27]. This extensive conservation strongly indicated a more general role of small RNAs in developmental regulation, as supported by the recent characterization of miRNA functions in many metazoan organisms.

miRNAs and siRNAs — what's the difference

Hundreds of miRNAs have now been identified in various organisms, and the RNA structure and regulatory mechanisms that have been characterized in *lin-4* and *let-7* still provide unique molecular signatures as to what defines miRNAs. miRNAs are generally 21–25 nucleotide, non-coding RNAs that are derived from larger precursors that form imperfect stem-loop structures (FIG. 1a)[4,5]. The mature miRNA is most often derived from one arm of the precursor hairpin, and is released from the primary transcript through stepwise processing by two ribonuclease-III (RNase III) enzymes[28,29]. At least in animals, most miRNAs bind to the target-3′ UTR with imperfect complementarity and function as translational repressors (see below for a discussion of plant miRNAs)[4].

Almost coincident with the discovery of the second miRNA, *let-7*, small RNAs were also characterized as components of a seemingly separate biological process, RNA interference (RNAi). RNAi is an evolutionarily conserved, sequence-specific gene-silencing mechanism that is induced by exposure to dsRNA[30]. In many systems, including worms, plants and flies, the stimulus that was used to initiate RNAi was the introduction of a dsRNA (the trigger) of ~500 bp. The trigger is ultimately processed *in vivo* into small dsRNAs of ~21–25 bp in length, designated as small interfering RNAs (siRNAs)[31,32]. It is now clear that one strand of the siRNA duplex is selectively incorporated into an effector complex (the RNA-induced silencing complex; RISC). The RISC directs the cleavage of complementary mRNA targets, a process that is also known as post-transcriptional gene silencing (PTGS) (FIG. 2)[33]. The evolutionarily conserved RNAi response to exogenous dsRNA might reflect an endogenous defense mechanism against virus infection or parasitic nucleic acids[30]. Indeed, mutations of the RNAi components greatly compromise virus resistance in plants, indicating that PTGS might normally mediate the destruction of the viral RNAs[34]. In addition, siRNAs can also regulate the expression of target transcripts at the transcriptional level, at least in some organisms. Not only can siRNAs induce sequence-specific promoter methylation in plants[35,36], but they are also crucial for heterochromatin formation in fission yeast[37,38], and transposon silencing in worms[39,40].

Fundamentally, siRNAs and miRNAs are similar in terms of their molecular characteristics, biogenesis and effector functions (see below for details). So, the current distinctions between these two species might be arbitrary, and might simply reflect the different paths through which they were originally discovered. miRNAs and siRNAs share a common RNase-III processing enzyme, Dicer, and closely related effector complexes, RISCs, for post-transcriptional repression (FIG. 2). In fact, much of our current knowledge of the biochemistry of miRNAs stems from what we know about siRNAs and the RNAi pathway.
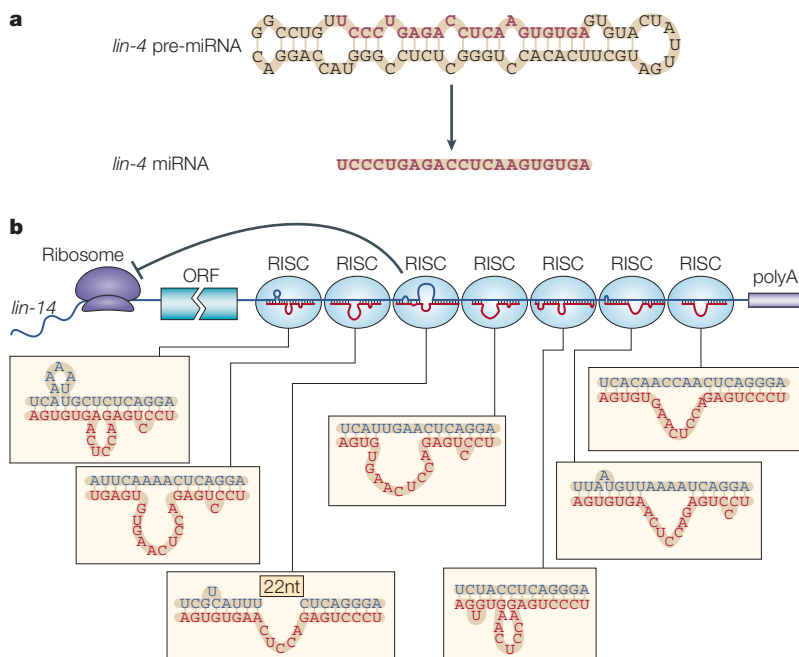
Figure 1 | **The molecular hallmarks of *lin-4*, the founding member of the microRNA family. a** | The precursor structure and mature microRNA (miRNA) sequence of *lin-4*. **b** | Sequence complementarity between *lin-4* (red) and the 3′-untranslated region (UTR) of *lin-14* mRNA (blue). *lin-4* is partially complementary to 7 sites in the *lin-14* 3′ UTR; its binding to these sites of complementarity brings about repression of LIN-14 protein synthesis[13,18]. RISC, RNA-induced silencing complex.
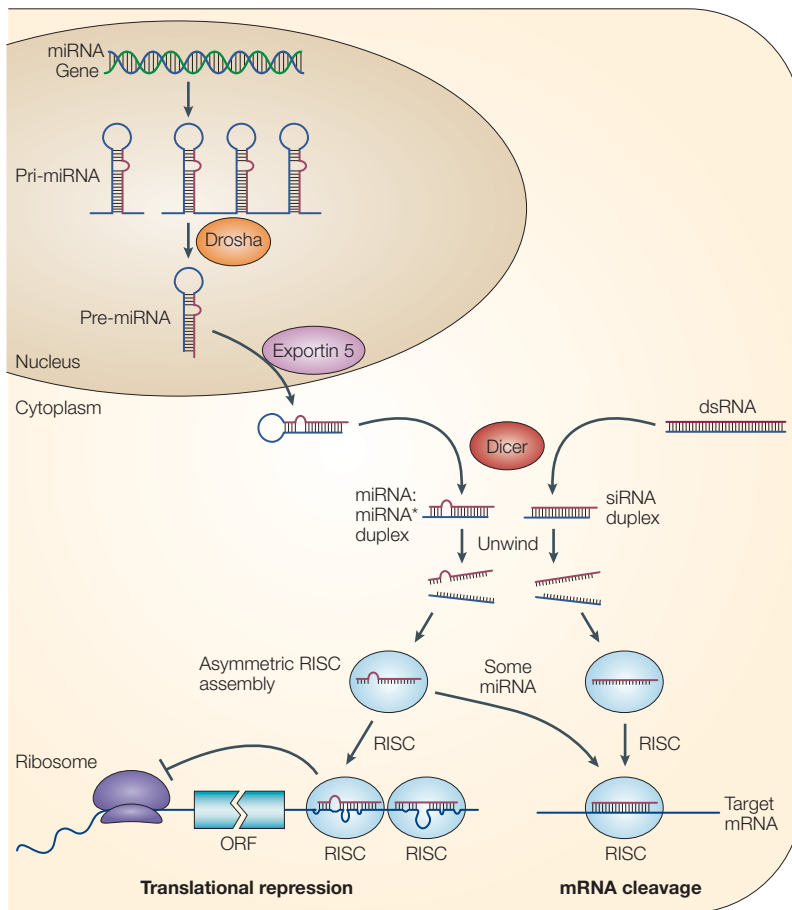
Figure 2 | **The current model for the biogenesis and post-transcriptional suppression of microRNAs and small interfering RNAs.** The nascent pri-microRNA (pri-miRNA) transcripts are first processed into ~70-nucleotide pre-miRNAs by Drosha inside the nucleus. Pre-miRNAs are transported to the cytoplasm by Exportin 5 and are processed into miRNA:miRNA* duplexes by Dicer. Dicer also processes long dsRNA molecules into small interfering RNA (siRNA) duplexes. Only one strand of the miRNA:miRNA* duplex or the siRNA duplex is preferentially assembled into the RNA-induced silencing complex (RISC) , which subsequently acts on its target by translational repression or mRNA cleavage, depending, at least in part, on the level of complementarity between the small RNA and its target. ORF, open reading frame.
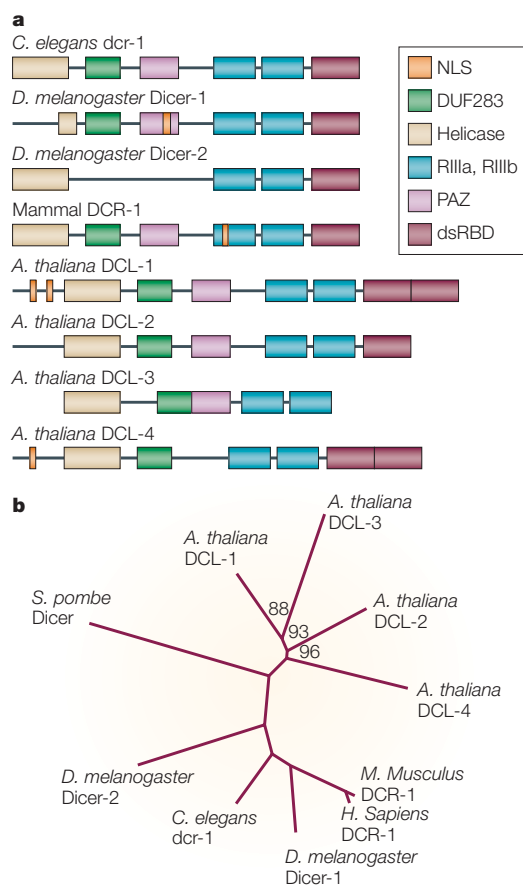
The artificial nature of current distinctions between siRNA and miRNA is highlighted by numerous recent findings. The extent of complementarity between the siRNA/miRNA and its target can determine the mechanism of silencing. For example, with the exception of *miR-172*, which acts as a translational repressor, all characterized plant miRNAs anneal to their targets with nearly complete complementarity at a single site, either in the coding region or in the UTRs, therefore condemning their target mRNAs to destruction by cleavage and degradation[41–43]. A similar situation has also been described for one mammalian miRNA, *miR-196*: the nearly perfect base pairing between *miR-196* and *Hoxb8* directs cleavage of *Hoxb8* mRNA both in mouse embryos and in cell culture[44]. Conversely, when siRNAs pair with their targets imperfectly, siRNAs can trigger translational repression rather than mRNA cleavage in mammalian tissue culture[45]. So, we are left with the question of whether miRNAs are truly different from siRNAs or whether our current understanding fails to functionally distinguish these two species under physiological conditions.

Both miRNAs and siRNAs depend on Dicer for their maturation[46,47], and both have been shown to be part of similar RISCs (FIG. 2)[48,49]. However, the effector complexes have only been studied for a few miRNAs, and in no case has there been biochemical data to confirm that most miRNA-containing complexes have been accounted for[50,51]. So, as we get beyond the superficial similarities of the structure and functions of these small RNAs, we must now begin to focus on the details that distinguish the modes of action of siRNAs and miRNAs *in vivo* to understand their true biological functions.

Biogenesis of miRNAs

Two processing events lead to mature miRNA formation in animals. In the first, the nascent miRNA transcripts (pri-miRNA) are processed into ~70-nucleotide precursors (pre-miRNA); in the second event that follows, this precursor is cleaved to generate ~21–25-nucleotide mature miRNAs[29]. Little is known about the transcriptional regulation of pri-miRNAs, except that certain pri-miRNAs are located within introns of host genes, including both protein-coding genes and non-coding genes, and might therefore be transcriptionally regulated through their host-gene promoters[52]. In addition, certain miRNAs are clustered in polycistronic transcripts, indicating that these miRNAs are coordinately regulated during development[52].

The sequential cleavages of miRNA maturation are catalysed by two RNase-III enzymes, Drosha and Dicer (FIG. 2)[28,47]. Both are dsRNA-specific endonucleases that generate 2-nucleotide-long 3′ overhangs at the cleavage site. Drosha is predominantly localized in the nucleus and contains two tandem RNase-III domains, a dsRNA binding domain and an amino-terminal segment of unknown function[28]. Regardless of the diverse primary sequences and structures of pri-miRNAs, Drosha cleaves these into ~70-bp pre-miRNAs that consist of an imperfect stem-loop structure (FIG. 2)[28]. Although the precise mechanisms that Drosha uses to discriminate miRNA precursors remain unknown, several studies

On the other hand, miRNAs differ from siRNAs in their molecular origins and, in many of the cases that have been characterized so far, in their mode of target recognition. miRNAs are produced as a distinct species from a specific precursor that is encoded in the genome. The structure of the primary miRNA transcript and the recognition of this precursor by a nuclear processing machinery probably determines the sequence and structure of mature miRNAs[4]. By contrast, siRNAs are sampled more randomly from long dsRNAs that can be introduced exogenously or produced from bi-directionally transcribed endogenous RNAs that anneal to form dsRNA[30]. In many cases, miRNAs bind to the target 3′ UTRs through imperfect complementarity at multiple sites, and therefore negatively regulate target expression at the translational level. By contrast, siRNAs often form a perfect duplex with their targets at only one site, and therefore direct the cleavage of the target mRNAs at the site of complementarity (FIG. 2).

a

C. elegans dcr-1

D. melanogaster Dicer-1

D. melanogaster Dicer-2

Mammal DCR-1

A. thaliana DCL-1

A. thaliana DCL-2

A. thaliana DCL-3

A. thaliana DCL-4

NLS	
DUF283	
Helicase	
RIIIa, RIIIb	
PAZ	
dsRBD	

b

S. pombe Dicer

A. thaliana DCL-1

A. thaliana DCL-3

A. thaliana DCL-2

A. thaliana DCL-4

88
93
96

D. melanogaster Dicer-2

C. elegans dcr-1

M. Musculus DCR-1

H. Sapiens DCR-1

D. melanogaster Dicer-1

Figure 3 | **The structure and function of the Dicer family.**
a | The domain structure of Dicer homologues in worms, flies, mammals and plants. The Dicer homologues that function in the microRNA (miRNA) maturation pathway have the PAZ (Piwi–Argonaute–Zwille) domain. **b** | The phylogenetic tree of the Dicer protein family. Multiple sequence alignment was done using ClustalW (version 1.80). PHYLIP95 (see online links box) was used to do BOOTSTRAP SAMPLING, protein-distance calculation (using Dayhoff PAM distance) and tree construction. Bootstrap percentages are indicated at each fork when the percentage is below 100. *A. thaliana*, *Arabidopsis thaliana*; *C. elegans*, *Caenorhabditis elegans*; *D. melanogaster*, *Drosophila melanogaster*; *H. sapiens*, *Homo sapiens*; *M. musculus*, *Mus musculus*; RBD, RNA binding domain; *S. pombe*, *Schizosaccharomyces pombe*.

miRNA*) that contains both the mature miRNA strand and its complementary strand (miRNA*)[47,55,56]. Dicer contains a putative helicase domain, a DUF283 domain, a PAZ (Piwi–Argonaute–Zwille) domain, two tandem RNase-III domains and a dsRNA-binding domain (dsRBD) (FIG. 3a)[46]. Recent structural analysis of the PAZ domain revealed a variant of the OB fold, a module that allows a low-affinity interaction with the 3′ end of ssRNAs[57–59]. This association also allows the PAZ domain to interact with dsRNAs that present 2-nucleo-tide 3′ overhangs, such as those that result from Drosha cleavage[51,57,58]. In addition, efficient Dicer cleavage also requires the presence of the overhang and a minimal stem length (D. Siolas and G.J.H., unpublished observations), indicating a model in which the Dicer PAZ domain might recognize the end of the Drosha cleavage product, and therefore position the site of the second RNase-III cleavage on the stem of the miRNA precursors[60]. Dicer cleavage generates mature miRNAs that range from 21 to 25 nucleotides, and such differences in size possibly result from the presence of bulges and mismatches on the pre-miRNA stem.

During Dicer processing, efficient cleavage of dsRNA requires dimerized RNase-III domains, because, on the basis of known RNAse-III structures, functional catalytic sites can only be formed at the interface of the RNase-III dimer[61]. As Dicer might only catalyse one cleavage event during miRNA processing, it is probable that Dicer functions as a monomer with its two RNase-III domains forming an intramolecular dimer and that it generates an active catalytic site for dsRNA cleavage[60]. Similar to Dicer, Drosha also contains two tandem RNase-III domains and carries out a single cleavage event to generate pre-miRNA. Given our lack of knowledge about the precise structural elements that guide Dicer and Drosha cleavage, it is difficult to speculate about their exact biochemical mechanisms. However, it is probable that Drosha and Dicer share closely related mechanisms for the processing of miRNAs.

Although plant miRNAs seem to be produced from long, primary transcripts, the maturation of miRNAs in plants must, by necessity, occur differently from miRNA maturation in animals, as no plant Drosha homologue has yet been found. However, the Dicer family in *Arabidopsis thaliana* has a level of complexity that has not been observed in other organisms studied so far (FIG. 3b). There are four Dicer homologues in *A. thaliana* — *DCL1*, *DCL2*, *DCL3* and *DCL4* — two of which (*DCL1* and *DCL4*) contain nuclear localization signals (FIG. 3a). Therefore, it seems possible that Drosha function might be carried out by one or more specialized Dicer enzymes in plants. The Dicer homologue *DCL1* contains two nuclear localization signals, and predominates the nuclei when expressed as a GFP-fusion protein in transient-transfection studies[62]. In plants that are deficient for *DCL1*, the production of mature miRNAs is reduced for all miRNAs examined, yet no accumulation is detected for the corresponding pre-miRNAs[62,63]. These findings imply that *DCL1* might catalyse both the Drosha- and Dicer-like cleavages for the maturation of some, if not most, miRNAs inside the nucleus.

BOOTSTRAP SAMPLING
As applied to molecular phylogenies, nucleotide or amino-acid sites are sampled randomly, with replacement, and a new tree is constructed. This is repeated many times and the frequency of appearance of a particular node among the bootstrap trees is viewed as a support (confidence) value for deciding on the significance of that node.

have addressed the features of pri-miRNAs that contribute to Drosha cleavage both *in vitro* and *in vivo*[28,53]. The efficiency of Drosha processing depends on the terminal loop size, the stem structure and the flanking sequence of the Drosha cleavage site, because shortening of the terminal loop, disruption of complementarity within the stem and removal or mutation of sequences that flank the Drosha cleavage site significantly decrease, if not abolish, the Drosha processing of pri-miRNAs[28,53].

After the initial cleavage by Drosha, pre-miRNAs are exported from the nucleus into the cytoplasm by Exportin 5 (Exp5), a Ran-GTP dependent nucleo/cytoplasmic cargo transporter (FIG. 2)[54]. Once inside the cytoplasm, these hairpin precursors are cleaved by Dicer into a small, imperfect dsRNA duplex (miRNA:

Consistent with this model, mature miRNAs seem to be produced within the nucleus in plants, because nuclear expression of P19, a viral protein that represses the accumulation of siRNAs and miRNAs, results in a significant reduction in mature miRNAs, whereas the cytoplasmic expression of P19 has no such effects[62]. It is not clear whether other Dicer homologues in plants are also involved in miRNA maturation. However, it has been speculated that the presence of multiple plant Dicer enzymes might explain the complexity of siRNA species in plants. Unlike siRNAs in animals, which are usually 21–22 nucleotides long, plants generate two classes of siRNA: 21–22- and ~25-bp-long siRNAs[64]. Given the different classes of small dsRNAs found in plants, it is possible that each Dicer homologue regulates separate cleavage events to generate specific species of dsRNA duplexes[65].

The functional specificity of different Dicer enzymes in organisms with multiple Dicer homologues has recently been indicated by a series of genetic and biochemical studies in *Drosophila melanogaster*[66,67]. Two Dicer homologues have been identified in flies: Dicer1 and Dicer2 (FIG. 3a)[67]. Deficiency in Dicer1 disrupts the processing of pre-miRNAs, whereas loss of Dicer2 affects the production of siRNAs, but not miRNA maturation[67]. These findings are consistent with the fact that the PAZ domain is only present in Dicer1, but absent in Dicer2 (FIG. 3a), given a model in which the PAZ domain recognizes the staggered ends of pre-miRNAs[67] and mediates their cleavage. Both Dicer1 and Dicer2 seem to function downstream of miRNA and siRNA production to facilitate the RISC-mediated gene silencing. Originally indicated by physical interactions between Dicer and RISC in *D. melanogaster*[68], this model has recently gained concrete biochemical support. Dicer2 forms a complex with R2D2, a dsRNA-binding protein, and the formation of the Dicer2/R2D2 complex with siRNAs enhances sequence-specific mRNA degradation that is mediated by the RISC complex[69,70]. In addition, Dicer1 deficiency affects siRNA-mediated gene silencing without disrupting siRNA production, indicating a possible role of Dicer1 in enhancing RNAi effector activity[67]. Therefore, the functions of Dicer might be broader than previously suspected, involving both the initiation and the effector steps of miRNA/siRNA-mediated gene silencing.

The effector complex for miRNAs shares core components with that of siRNAs, so much so that both are collectively referred to as the RNA-induced silencing complex (RISC). The target specificity, and probably also the functional efficiency, of a miRNA requires that the mature miRNA strand from the miRNA:miRNA* duplex be selectively incorporated into the RISC for target recognition[71,72]. The miRNA* strand, on the other hand, is probably degraded rapidly on its exclusion from the RISC, as the recovery rate of miRNA*s from endogenous tissues is ~100-fold lower than that of miRNAs[71]. As Dicer processes the pre-miRNA into the miRNA:miRNA* duplex, the stability of the 5′ ends of the two arms of the miRNA:miRNA* duplex is usually different. Although mature miRNAs can reside on either strand of the hairpin stem, it is almost always derived from the

strand with the less stable 5′ end compared with the miRNA* strand[71,72]. These findings indicate that the relative instability at the 5′ end of the mature miRNA might facilitate its preferential incorporation into the RISC. The selective assembly of the mature miRNA into the RISC probably reflects the relative ease of unwinding from one end of the miRNA: miRNA* duplex (FIG. 2)[71,72]. Therefore, the thermodynamic properties of the miRNA precursor determine the asymmetrical RISC assembly, and therefore, the target specificity for post-transcriptional inhibition. However, in rare cases in which miRNA and miRNA* have similar 5′-end stability, each arm of the miRNA precursor is predicted to be assembled into the RISC at similar frequencies. This prediction has been confirmed by similar recovery rates for such miRNAs and miRNA*s from endogenous tissues[71]. This thermodynamic model also applies to the asymmetrical assembly of the siRNA duplex, in which the strand of siRNA with the less stable 5′ end is preferentially assembled into the RISC complex to target mRNA cleavage[71,72]. Altogether, there seems to be a common thermodynamic mechanism that regulates the asymmetric assembly of siRNA or miRNA from the dsRNA duplexes, which safeguards specificity towards corresponding targets.

Post-transcriptional repression by miRNAs

The precise molecular mechanisms that underlie post-transcriptional repression by miRNAs still remain largely unknown. One of the best-studied examples is *lin-4*, which negatively regulates its target, *lin-14*, by repressing its translation[16]. Base pairing between *lin-4* and *lin-14* has proved to be crucial for their interaction *in vivo*, as mutations that affect their complementarity compromise or abolish this negative regulation[13]. Interestingly, *lin-4* only inhibits the synthesis of the LIN-14 protein but fails to affect the synthesis, polyadenylation state or abundance of *lin-14* mRNA[16]. Furthermore, the initiation of *lin-14* translation seems to occur normally in the presence of *lin-4,* because *lin-14* mRNAs are efficiently incorporated into polyribosomes regardless of *lin-4* expression[16]. Therefore, it is reasonable to speculate that the translational repression by *lin-4* occurs after translational initiation, probably during translational elongation and/or the subsequent release of the LIN-14 protein[16].

Translational repression of target genes is not specific to *lin-4*; in fact, it turns out to be the predominant mechanism by which miRNAs negatively regulate their targets throughout the animal kingdom. *let-7* and *Bantam* (which encodes an anti-apoptotic miRNA) bind to the 3′ UTRs of their targets and negatively regulate their translation in worms and flies, respectively[20,73]. In addition, *miR-30*, a mammalian miRNA, inhibits protein synthesis of a reporter gene that bears an artificial 3′ UTR with *miR-30* complementary sites[53].

Although most animal miRNAs repress target translation, one miRNA, *mir-196*, was found recently to direct mRNA cleavage of its target, *Hoxb8* (REF. 44). In plants, however, most miRNAs that have been studied so far mediate the destruction of their target mRNAs[74]. These

Table 1 | **Expression studies on mammalian microRNAs**

Expression pattern	microRNA	References
Tissue-specific expression patterns of mammalian microRNAs		
ES-cell specific	*miR-296*	86
Expressed in ES cells, but upreguated on differentiation	*miR-21* and *miR-22*	86
Expressed in both ES cells and various adult tissues	*miR-15a, miR-16, miR-19,b, miR-92, miR-93, miR-96, miR-130* and *miR-130b*	86
Enriched during mouse brain development	*miR-128, miR-19b, miR-9, miR-125b, miR-131, miR-178, miR-124a, miR-266* and *miR-103*	26,90
Enriched in adult brain	*miR-9*, miR-125a, miR-125b, miR-128, miR-132, miR-137, miR-139, miR-7, miR-9, miR-124a, miR-124b, miR-135, miR-153, miR-149, miR-183, miR-190* and *miR-219*	26
Enriched in lung	*miR-18, miR-19a, miR-24, miR-32, miR-130, miR-213, miR-20, miR -141, miR-193* and *miR-200b*	26
Enriched in spleen	*miR-99a, miR-127, miR-142-a, miR-142-s, miR-151, miR-189* and *miR-212*	26
Haemetopoietic tissues	*miR-181, miR-223* and *miR-142*	26
Enriched in liver	*miR-122a, miR-152, miR-194, miR-199* and *miR-215*	26
Enriched in heart	*miR-1b, miR-1d, miR-133, miR-206, miR-208* and *miR-143*	26
Enriched in kidney	*miR-30b, miR-30c, miR-18, miR-20, miR-24, miR-32, miR-141, miR-193* and *miR-200b*	26
Ubiquitously expressed	*miR-16, miR-26a, miR-27a, miR143a, miR-21, let-7a, miR-7b, miR-30b* and *miR-30c*	26
Abnormal microRNA expression during tumorigenesis		
Downregulated in chronic lymphocytic leukaemias	*miR-15* and *miR-16*	102
Downregulated in lung cancer cell lines	*miR-26a* and *miR-99a*	89
Downregulated in colon cancers	*miR143/miR-145* cluster	103
Upregulated in Burkitt lymphoma	*miR-155*	88

ES cells, embryonic stem cells.

plant miRNAs differ from animal miRNAs in that their base pairing with the corresponding targets is nearly perfect, and that their complementary sites are located throughout the transcribed regions of the target gene, instead of being limited to the 3′ UTRs. So far, only one plant miRNA, *miR-172,* has been shown to act as a translational repressor during *A. thaliana* flower development. Unlike most characterized animal miRNAs, the site of *miR-172* complementarity (which is high) within the target gene *APETALA 2* (*AP2*) was located in the coding region instead of the 3′ UTR[41]. An important caveat to these observations is that those plant miRNAs that act through RNA cleavage might be overrepresented in the current studies owing to the ease of their target identification through computational analysis (see below).

Despite progress in identifying protein and RNA components of the RISC, the biochemical mechanism by which this complex functions still remains unknown. Genetic screens combined with biochemical purification have pinpointed several important components. Argonaute (AGO) proteins belong to an evolutionarily conserved family that is defined by the presence of a PAZ domain and a Piwi domain[75]. AGO-family proteins have been consistently co-purified with RISC activity in many organisms[48], and mutations in AGO homologues have been associated with distinct developmental and/or RNAi phenotypes[75]. As a core component of the RISC, the AGO family has multiple homologues in each metazoan species (24 for *C. elegans*, 5 for *D. melanogaster* and 8 for mammals)[75]. Given the diverse mutant phenotypes and expression patterns of AGO homologues, RISCs might come in different flavours and act in a tissue-specific or a developmentally regulated manner.

In addition to AGO homologues, several other proteins have also been co-purified with the RISC. It is not clear, however, whether these RISC-associated proteins are core RISC components or whether they act as accessory proteins that provide functional specificity for the RISCs under different developmental and/or physiological contexts. The RNA-binding proteins VIG and Fragile X-related protein[49,68], and the nuclease Tudor-SN, have been co-purified with RISC activity from *D. melanogaster* S2 CELL extracts[76]. Such components also associate with siRNAs and miRNAs in mammalian cells and *C. elegans*[76]. Alternative RISCs must also exist, as the helicases Gemin3 and Gemin4 have been co-purified with miRNAs and the human Argonaute homologue eIF2C2 (AGO2)[50] from mammalian cells. These observations are consistent with the model according to which RISCs exist in different subtypes, perhaps specialized to mediate different post-transcriptional repression events *in vivo.*

Table 2 | **microRNAs and their targets: examples of microRNAs with experimentally validated functions/targets**

microRNA	Target gene(s)	Function	Mode of repression	References
Caenorhabditis elegans				
lin-4	lin-14, lin-28	Regulation of developmental transition between the first two larval stages, L1 and L2	Translational repression	13,18,19
let-7	lin-41, hbl-1	Regulation of developmental transition between the last larval stage (L4) and the adult stage	Translational repression	20–23
lsy-6	cog-1	Determination of left/right asymmetry of neuronal development	Unknown	9
Drosophila melanogaster				
Bantam	hid	Promotion of cell proliferation and suppression of apoptosis	Translational repression	73
miR-14	Unknown	Suppression of apoptosis and regulation of fat metabolism	Unknown	97
Mus musculus				
miR-181	Unknown	Promotion of haematopoietic differentiation towards the B-cell lineage	Unknown	99
miR-196	Hoxb8	Unknown	PTGS	44
Arabidopsis thaliana				
miR-165, miR-166	PHB, PHV and REV	Regulation of leaf morphogenesis	PTGS	104,105
miR-172	AP2	Regulation of flowering time and floral-organ identity	Translational repression	41
miR-JAW	TCP transcription factors	Regulation of leaf development and embryogenesis	PTGS	98
miR-39	SCL family proteins	Unknown	PTGS	42
miR-159	MYB33 family transcription factors	Regulation of leaf morphogenesis	PTGS	43,63,98
Zea mays				
miR-166	rld1	Regulation of leaf morphogenesis	PTGS	106

AP2, APETALA 2; Hid, head involution defective; Hoxb8, homeobox B8; PHB, PHABULOSA; PHV, PHAVOLUTA; PTGS, post-transcriptional gene silencing; REV, REVOLUTA; rld1, rolled leaf1; SCL, SCARECROW-LIKE; TCP, teosinte branched 1-cycloidea-PCF.

Genome-wide efforts for miRNA identification

The extension of miRNA-mediated regulation from being a curiosity in *C. elegans* to a potentially general mechanism for gene regulation began with the work of the Tuschl, Bartel and Ambros groups, who collectively identified more than 100 novel miRNAs by cloning and sequencing endogenous small RNAs of 21–25 bp long from worms, flies and mammals[77–79]. miRNAs that were isolated in these studies show the characteristics of Dicer cleavage, temporally and spatially regulated expression and, in many cases, evolutionary conservation, either throughout metazoans or within closely related species[77–79]. Similar to *lin-4* and *let-7*, these miRNAs mostly map to one arm of their predicted stem-loop precursors. The existence of miRNA precursors with stem-loop structure is also indicated by northern-blot detection of ~70-nucleotide precursors, and by cloning of the rare reverse-strand miRNA[77–79]. In addition to the continued cloning efforts, novel miRNAs have been isolated through their association with POLYSOMES[80] and ribonucleoprotein complexes[50,51]. So far, several hundred miRNAs have been identified in mammals, fish, flies, worms and plants, corresponding to ~1% of the protein-coding genes in each organism[81]. As the current miRNA cloning technology clearly favours highly expressed miRNAs, more exhaustive cloning efforts are needed that must include more tissue types and developmental stages. In addition to experimental approaches, bioinformatic predictions have helped to identify novel miRNAs in various organisms, mostly on the basis of pre-miRNA hairpin structures and sequence conservation throughout evolution[82–84]. With the growing number of experimentally confirmed miRNAs, refined bioinformatic approaches will undoubtedly increase in power for identifying miRNAs that have escaped experimental searches.

A comprehensive miRNA registry

The global efforts for miRNA cloning and characterization have led to the establishment of an important collection of miRNA data. The miRNA Registry (see online links box) contains up-to-date annotation for all published miRNAs[81]. Most database entries require experimental validation of mature miRNA expression and computational prediction of the corresponding hairpin precursor. In cases in which homologous miRNAs can be easily identified from closely related species, such as mouse and human, computationally predicted entries are also included in the registry to provide a more comprehensive list[81]. At the time of writing, the database contains 116 *C. elegans* miRNAs, 50 miRNAs from *C. briggsae*, 78 *D. melanogaster* miRNAs, 43 *A. thaliana* miRNAs, 28 miRNA from rice, 207 mouse miRNAs, 185 miRNAs from rat and 191 human miRNAs.

S2 CELL
A cell line that is isolated from dissociated *Drosophila melanogaster* embryos. The cell line is phagocytic, which might contribute to its susceptibility to RNAi.

POLYSOME
A functional unit of protein synthesis that consists of several ribosomes that are attached along the length of a single molecule of mRNA.

Following the identification of hundreds of miRNAs in various organisms, large-scale studies on miRNA-expression profiles were carried out in many model organisms using northern-blot analysis, microarrays and miRNA cloning[25,85–87]. miRNAs show dynamic temporal and spatial expression patterns, disruption of which is associated with developmental/physiological abnormalities (TABLE 1). For example, the *miR-155* precursor is enriched in Burkitt lymphoma tissues; the location of *miR-155* correlates with a specific chromosomal translocation[88]. *miR-26a* and *miR-99a* expression is reduced in human lung cancer cell lines, which correlates with their location in regions of loss of heterozygosity in lung tumours/cell lines[89]. *miR-9* and *miR-131* levels are reduced in Presenilin-1-deficient mice that show lethal developmental defects[90]. These findings indicate that miRNAs might have a general role in regulating gene expression in diverse developmental and physiological processes, and provide substantial hints that misregulation of miRNA function might contribute to human disease.

Functional characterization of miRNAs

Although the studies of *lin-4* and *let-7* shaped our understanding of miRNA molecular structures and functional mechanisms, their roles in temporal regulation of development only revealed one of many possible aspects of miRNA function. Mutations in Dicer homologues disrupt the biogenesis of miRNAs, and cause diverse developmental defects, including germline defects in *C. elegans*[91], abnormal embryogenesis in *A. thaliana*[63], developmental arrest in zebrafish[92] and depletion of stem cells in mice[93]. Mutations in AGO family proteins are also associated with pleiotropic

developmental phenotypes, such as premature MERISTEM differentiation in *A. thaliana zwille* mutants[94], defective embryogenesis and larval development in *C. elegans alg-1* and *alg-2* mutants[55] and decreased stem-cell self-renewal during oogenesis in *Drosophila piwi* mutants[95]. Because Dicer and AGO are essential components in miRNA and siRNA biogenesis and function, these defects might reflect the collective functions of multiple miRNAs and/or siRNAs that are expressed during early development.

So far, functional studies have only touched on a handful of miRNAs. Following the classic studies on the role of *lin-4* and *let-7* in worm developmental timing, functional characterization of miRNAs continues to benefit from the loss-of-function mutations of miRNA genes. The fly miRNA, *Bantam,* was originally characterized from a fly P-ELEMENT screen for genes that promote cell proliferation and suppress apoptosis during tissue growth[96]. Its sequence complementarity with the 3′ UTR and its functional antagonism of the pro-apoptotic *hid* (head involution defective) gene were the first clues for the translational repression of *hid* by *Bantam*[73,96]. In addition, fly *miR-14* was identified through a P-element screen for inhibitors of apoptotic cell death. Deficiency in *miR-14* enhances cell death that is induced by the cell-death activator, Reaper, and results in defective stress responses and fat metabolism[97]. Loss of *miR-14* also leads to an elevated level of Drice, an apoptotic effector caspase, indicating a direct or indirect repression of Drice by *miR-14.* Perhaps one of the most interesting miRNAs identified so far is *lsy-6*, a miRNA that is asymmetrically expressed in worm taste-receptor neurons to promote specific cell fate in left–right patterning. ASE left (ASEL) and ASE right (ASER) are two taste-receptor neurons, each with specific

MERISTEM
The undifferentiated tissue at the tips of stems and roots in which new cell division is concentrated.

P-ELEMENTS
A family of transposable elements that are widely used as the basis of tools for mutating and manipulating the *Drosophila* genome.

WING DISC
A sac-like structure of a mature third instar fly larva, which will give rise to the adult wing.

Box1 | **Bioinformatic predictions of microRNA targets**

The functional characterization of microRNAs (miRNAs) heavily relies on the identification of miRNA target genes. In plants, because certain miRNAs are nearly perfectly complementary to their targets, bioinformatic predictions of miRNA targets have proved to be relatively straightforward, and in fact, powerful[43]. Interestingly, many predicted miRNA targets are transcription factors, highlighting possible miRNA functions in regulating diverse developmental processes in plants[43]. However, the computational prediction of animal miRNA targets is more difficult, both because the miRNAs are only ~21–25 nucleotides long and because the miRNA-target pairings are not entirely complementary[100]. As a result, bioinformatic prediction has to rely on rules that are built on a few known miRNA-target interactions, which are generalized for genome-wide searches. Starting from 3 animal miRNAs, *lin-4, let-7* and *Bantam,* the targets of which have been experimentally validated, the Cohen group screened the fly genome for miRNA targets on the basis of the following 3 criteria: perfect complementarity or G•U pairing between the target 3′-untranslated region (UTR) and the first 8 nucleotides of miRNA; favourable structural and thermodynamic heteroduplex formation between miRNA and its putative targets; and evolutionary conservation of miRNA target sites between *Drosophila melanogaster* and *D. pseudoobscura*[101]. To experimentally validate potential targets, they fused a *GFP* reporter gene upstream of the predicted target 3′ UTR and examined *GFP* expression in WING DISC with and without overexpression of the corresponding miRNA[101]. Using this approach, they gained experimental support for the informatically predicted targets of three miRNAs, including Notch genes for *miR-7*, pro-apoptotic genes for *miR-2* and genes that encode metabolic enzymes for *miR-277* (REF. 101).

The Burge group developed the TargetScan algorithm for predicting vertebrate miRNA targets on the basis of miRNA-target complementarity (particularly at the 5′ region of the miRNA) and evolutionary conservation among vertebrates. In their first study, they experimentally validated 11 predicted targets out of 15 tested, using a HeLa cell reporter system that contains a luciferase reporter gene fused to the predicted target 3′-UTR fragments[100]. Among their predicted targets, 29% have unknown functions, whereas 71% have diverse functions, such as DNA binding, transcriptional regulation, signal-transducer and kinase activity[100]. Although the prediction success rate is difficult to determine on a genome-wide scale, these studies are an important step towards our understanding of miRNA function in animals.

expression patterns of chemoreceptor genes. The specific *lsy-6* expression in the ASEL neuron suppresses the expression of an *Nkx*-type homeobox gene, *cog-1*, thereby promoting the ASEL cell fate by specifying ASEL-specific expression of chemoreceptors.

The functional characterization of plant miRNAs has also benefited from genetic mutations in miRNA genes. The *miR-JAW* gene was originally identified as a gain-of-function mutation causing the uneven leaf curvature and shape[98]. The identification of *miR-JAW* targets — the *TCP* genes — was first indicated by their opposing biological functions during leaf morphogenesis and their sequence complementarity[98].

For most of the miRNAs for which genetic mutations are unavailable, functional characterization might have to begin with expression studies and/or bioinformatic predictions. For example, *miR-181* is enriched in B-lymphoid cells of mouse bone marrow; this unique expression pattern led to the discovery of its function in promoting haematopoietic differentiation towards the B-cell lineage[99]. *miR-39*, which accumulates predominantly in the INFLORESCENCE TISSUE of *A. thaliana*[42] is another example. Using the sequence complementarity and overlapping expression pattern as the target search criteria, Llave *et al.* predicted the *miR-39* targets as Scarecrow-like (SCL) family proteins and further confirmed their hypothesis through biochemical studies (TABLE 2)[42].

Most of the miRNAs that have been characterized so far seem to regulate aspects of development, including larval developmental transitions and neuronal development in *C. elegans*, growth control and apoptosis in *D. melanogaster*, haematopoietic differentiation in mammals and leaf development in *A. thaliana* (TABLE 2). But, only a handful of miRNAs have been carefully studied so far, and the diverse expression patterns of miRNAs and altered miRNA expression under certain physiological conditions, such as tumorigenesis, indicate a range of unknown functions that might extend beyond developmental regulation. In addition, bioinformatic predictions of miRNA targets (BOX 1) have revealed a diversity of regulatory pathways that might be subject to miRNA-mediated regulation. Overall, no specific indication has emerged that miRNA-mediated regulation is restricted to one biological process. Instead, the emerging picture is that miRNAs have the potential to regulate almost all aspects of cellular physiology.

Conclusions

Since the discovery of miRNAs as stRNAs in *C. elegans*[13,18], remarkable advances in the characterization of this gene family have not only demonstrated that these small, non-coding RNAs are a prevalent class of regulatory RNAs, but they have also indicated the outlines of a biochemical mechanism for their functions in gene regulation. However, with relatively few exceptions, we know little about the precise roles of the vast majority of miRNAs in regulating gene expression. Furthermore, the precise mechanisms by which miRNA- and siRNA-mediated repression might differ remain to be explained. For example, miRNAs and siRNAs mature in a similar way and join structurally related, if not identical, effector complexes. However, subtle differences between these classes of small RNA are beginning[4], and will no doubt continue, to emerge as we understand more of the precise relationships between miRNAs, siRNAs and the protein components of the RNAi machinery.

1. del Solar, G. & Espinosa, M. Plasmid copy number control: an ever-growing story. *Mol. Microbiol.* **37**, 492–500 (2000).
2. Mlynarczyk, S. K. & Panning, B. X inactivation: Tsix and Xist as yin and yang. *Curr. Biol.* **10**, R899–R903 (2000).
3. Ambros, V. MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. *Cell* **113**, 673–676 (2003).
4. Bartel, D .P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
5. Lai, E. C. microRNAs: runts of the genome assert themselves. *Curr. Biol.* **13**, R925–R936 (2003).
6. Pasquinelli, A. E. & Ruvkun, G. Control of developmental timing by micrornas and their targets. *Annu. Rev. Cell Dev. Biol.* **18**, 495–513 (2002).
7. McManus, M. T. MicroRNAs and cancer. *Semin. Cancer Biol.* **13**, 253–258 (2003).
8. Carrington, J. C. & Ambros, V. Role of microRNAs in plant and animal development. *Science* **301**, 336–338 (2003).
9. Johnston, R. J. & Hobert, O. A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. *Nature* **426**, 845–849 (2003).
10. Chalfie, M., Horvitz, H. R. & Sulston, J. E. Mutations that lead to reiterations in the cell lineages of *C. elegans*. *Cell* **24**, 59–69 (1981).
11. Ambros, V. A hierarchy of regulatory genes controls a larva-to-adult developmental switch in *C. elegans*. *Cell* **57**, 49–57 (1989).
12. Ambros, V. & Horvitz, H. R. Heterochronic mutants of the nematode *Caenorhabditis elegans*. *Science* **226**, 409–416 (1984).
13. Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843–854 (1993). **Described the identification of the first microRNA, *lin-4*, and reported the sequence complementarity between *lin-4* and the 3′ UTR of the *lin-14* mRNA.**
14. Wightman, B., Burglin, T. R., Gatto, J., Arasu, P. & Ruvkun, G. Negative regulatory sequences in the *lin-14* 3′-untranslated region are necessary to generate a temporal switch during *Caenorhabditis elegans* development. *Genes Dev.* **5**, 1813–1824 (1991).
15. Ruvkun, G. & Giusto, J. The *Caenorhabditis elegans* heterochronic gene *lin-14* encodes a nuclear protein that forms a temporal developmental switch. *Nature* **338**, 313–319 (1989).
16. Olsen, P. H. & Ambros, V. The *lin-4* regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev. Biol.* **216**, 671–680 (1999).
17. Ha, I., Wightman, B. & Ruvkun, G. A bulged *lin-4/lin-14* RNA duplex is sufficient for *Caenorhabditis elegans lin-14* temporal gradient formation. *Genes Dev.* **10**, 3041–3050 (1996).
18. Wightman, B., Ha, I. & Ruvkun, G. Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* **75**, 855–862 (1993). **Described the translational repression of LIN-14 by *lin-4* during temporal regulation of larval development. This was the first functional characterization of a microRNA.**
19. Moss, E. G., Lee, R. C. & Ambros, V. The cold shock domain protein LIN-28 controls developmental timing in *C. elegans* and is regulated by the *lin-4* RNA. *Cell* **88**, 637–646 (1997).
20. Reinhart, B.J. *et al.* The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**, 901–906 (2000).
21. Lin, S. Y. *et al.* The *C. elegans hunchback* homolog, *hbl-1*, controls temporal patterning and is a probable microRNA target. *Dev. Cell* **4**, 639–650 (2003).
22. Abrahante, J. E. *et al.* The *Caenorhabditis elegans hunchback*-like gene *lin-57/hbl-1* controls developmental time and is regulated by microRNAs. *Dev. Cell* **4**, 625–637 (2003).
23. Slack, F. J. *et al.* The *lin-41* RBCC gene acts in the *C. elegans* heterochronic pathway between the *let-7* regulatory RNA and the LIN-29 transcription factor. *Mol. Cell* **5**, 659–669 (2000).
24. Vella, M. C., Choi, E. Y., Lin, S. Y., Reinert, K. & Slack, F. J. The *C. elegans* microRNA *let-7* binds to imperfect *let-7* complementary sites from the *lin-41* 3′ UTR. *Genes Dev.* **18**, 132–137 (2004).
25. Lagos-Quintana, M. *et al.* Identification of tissue-specific microRNAs from mouse. *Curr. Biol.* **12**, 735–739 (2002).
26. Sempere, L. F. *et al.* Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. *Genome Biol.* **5**, R13 (2004).
27. Pasquinelli, A. E. *et al.* Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA. *Nature* **408**, 86–89 (2000).
28. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419 (2003). **Described the identification of Drosha and characterizes its function in processing pri-miRNA into pre-miRNA.**
29. Lee, Y., Jeon, K., Lee, J. T., Kim, S. & Kim, V. N. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.* **21**, 4663–4670 (2002).
30. Hannon, G. J. RNA interference. *Nature* **418**, 244–251 (2002).
31. Elbashir, S. M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498 (2001).
32. Elbashir, S. M., Lendeckel, W. & Tuschl, T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15**, 188–200 (2001).
33. Zamore, P. D., Tuschl, T., Sharp, P. A. & Bartel, D. P. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25–33 (2000).

34. Baulcombe, D. Viruses and gene silencing in plants. *Arch. Virol.* **15** (Suppl.), 189–201 (1999).
35. Aufsatz, W., Mette, M. F., van der Winden, J., Matzke, A. J. & Matzke, M. RNA-directed DNA methylation in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **99** (Suppl 4), 16499–16506 (2002).
36. Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A. & Matzke, A. J. Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J.* **19**, 5194–5201 (2000).
37. Grewal, S. I. & Moazed, D. Heterochromatin and epigenetic control of gene expression. *Science* **301**, 798–802 (2003).
38. Volpe, T. A. *et al.* Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**, 1833–1837 (2002).
39. Ketting, R. F., Haverkamp, T. H., van Luenen, H. G. & Plasterk, R. H. *Mut-7* of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell* **99**, 133–141 (1999).
40. Tabara, H. *et al.* The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* **99**, 123–132 (1999).
41. Chen, X. A MicroRNA as a translational repressor of *APETALA2* in *Arabidopsis* flower development. *Science* **303**, 2022–2025 (2004).
42. Llave, C., Xie, Z., Kasschau, K. D. & Carrington, J. C. Cleavage of *Scarecrow-like* mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* **297**, 2053–2056 (2002).
43. Rhoades, M. W. *et al.* Prediction of plant microRNA targets. *Cell* **110**, 513–520 (2002).
   The first bioinfomatic effort to predict microRNA targets on the basis of sequence complementarity between plant miRNAs and their putative targets. It has guided functional studies of several miRNAs.
44. Yekta, S., Shih, I. H. & Bartel, D. P. MicroRNA-directed cleavage of *HOXB8* mRNA. *Science* **304**, 594–596 (2004).
45. Doench, J. G., Petersen, C. P. & Sharp, P. A. siRNAs can function as miRNAs. *Genes Dev.* **17**, 438–442 (2003).
46. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366 (2001).
   Described the identification of Dicer and characterized its function in processing long dsRNAs into small interfering RNAs.
47. Hutvagner, G. *et al.* A cellular function for the RNA-interference enzyme Dicer in the maturation of the *let-7* small temporal RNA. *Science* **293**, 834–838 (2001).
48. Hammond, S. M., Boettcher, S., Caudy, A. A., Kobayashi, R. & Hannon, G. J. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**, 1146–1150 (2001).
   Described the purification of the RISC, and the identification of Argonaute 2 as a key component.
49. Caudy, A. A., Myers, M., Hannon, G. J. & Hammond, S. M. Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev.* **16**, 2491–2496 (2002).
50. Mourelatos, Z. *et al.* miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev.* **16**, 720–728 (2002).
51. Dostie, J., Mourelatos, Z., Yang, M., Sharma, A. & Dreyfuss, G. Numerous microRNPs in neuronal cells containing novel microRNAs. *RNA* **9**, 180–186 (2003).
52. Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A. & Tuschl, T. New microRNAs from mouse and human. *RNA* **9**, 175–179 (2003).
53. Zeng, Y. & Cullen, B. R. Sequence requirements for micro RNA processing and function in human cells. *RNA* **9**, 112–123 (2003).
54. Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. Nuclear export of microRNA precursors. *Science* **303**, 95–98 (2004).
55. Grishok, A. *et al.* Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**, 23–34 (2001).
56. Ketting, R. F. *et al.* Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* **15**, 2654–2659 (2001).
57. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Structure and nucleic-acid binding of the *Drosophila* Argonaute 2 PAZ domain. *Nature* **426**, 465–469 (2003).
58. Song, J. J. *et al.* The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nature Struct. Biol.* **10**, 1026–1032 (2003).
59. Yan, K. S. *et al.* Structure and conserved RNA binding of the PAZ domain. *Nature* **426**, 468–474 (2003).
60. Carmell, M. A. & Hannon, G. J. RNase III enzymes and the initiation of gene silencing. *Nature Struct. Mol. Biol.* **11**, 214–218 (2004).

61. Blaszczyk, J. *et al.* Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage. *Structure* (*Camb*) **9**, 1225–1236 (2001).
62. Papp, I. *et al.* Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors. *Plant Physiol.* **132**, 1382–1390 (2003).
63. Park, W., Li, J., Song, R., Messing, J. & Chen, X. CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr. Biol.* **12**, 1484–1495 (2002).
64. Timmons, L. The long and short of siRNAs. *Mol. Cell* **10**, 435–437 (2002).
65. Schauer, S. E., Jacobsen, S. E., Meinke, D. W. & Ray, A. *DICER-LIKE1*: blind men and elephants in *Arabidopsis* development. *Trends Plant Sci.* **7**, 487–491 (2002).
66. Pham, J. W., Pellino, J. L., Lee, Y. S., Carthew, R. W. & Sontheimer, E. J. A Dicer-2-dependent 80S complex cleaves targeted mRNAs during RNAi in *Drosophila*. *Cell* **117**, 83–94 (2004).
67. Lee, Y. S. *et al.* Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117**, 69–81 (2004).
68. Jin, P. *et al.* Biochemical and genetic interaction between the fragile X mental retardation protein and the microRNA pathway. *Nature Neurosci.* **7**, 113–117 (2004).
69. Liu, Q. *et al.* R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway. *Science* **301**, 1921–1925 (2003).
70. Pellino, J. L. & Sontheimer, E. J. R2D2 leads the silencing trigger to mRNA's death star. *Cell* **115**, 132–133 (2003).
71. Schwarz, D. S. *et al.* Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199–208 (2003).
72. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209–216 (2003).
   This paper, together with reference 71, characterized the regulatory mechanism of the asymmetric assembly of siRNA/miRNA into the RISC complex.
73. Brennecke, J., Hipfner, D. R., Stark, A., Russell, R. B. & Cohen, S. M. *Bantam* encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila*. *Cell* **113**, 25–36 (2003).
74. Hake, S. MicroRNAs: a role in plant development. *Curr. Biol.* **13**, R851–R852 (2003).
75. Carmell, M. A., Xuan, Z., Zhang, M. Q. & Hannon, G. J. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev.* **16**, 2733–2742 (2002).
76. Caudy, A. A. *et al.* A micrococcal nuclease homologue in RNAi effector complexes. *Nature* **425**, 411–414 (2003).
77. Lau, N. C., Lim, L. P., Weinstein, E. G. & Bartel, D. P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**, 858–862 (2001).
78. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853–858 (2001).
79. Lee, R. C. & Ambros, V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**, 862–864 (2001).
   This paper, together with references 77 and 78, was among the first cloning efforts to identify large numbers of miRNAs from worm, fly and mammals.
80. Kim, J. *et al.* Identification of many microRNAs that copurify with polyribosomes in mammalian neurons. *Proc. Natl Acad. Sci. USA* **101**, 360–365 (2004).
81. Griffiths-Jones, S. The microRNA Registry. *Nucleic Acids Res.* **32** (Database issue), D109–D111 (2004).
82. Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B. & Bartel, D. P. MicroRNAs in plants. *Genes Dev.* **16**, 1616–1626 (2002).
83. Lim, L. P., Glasner, M. E., Yekta, S., Burge, C. B. & Bartel, D. P. Vertebrate microRNA genes. *Science* **299**, 1540 (2003).
84. Lim, L. P. *et al.* The microRNAs of *Caenorhabditis elegans*. *Genes Dev.* **17**, 991–1008 (2003).
85. Sempere, L. F., Sokol, N. S., Dubrovsky, E. B., Berger, E. M. & Ambros, V. Temporal regulation of microRNA expression in *Drosophila melanogaster* mediated by hormonal signals and broad-complex gene activity. *Dev. Biol.* **259**, 9–18 (2003).
86. Houbaviy, H. B., Murray, M. F. & Sharp, P. A. Embryonic stem cell-specific microRNAs. *Dev. Cell* **5**, 351–358 (2003).
87. Aravin, A. A. *et al.* The small RNA profile during *Drosophila melanogaster* development. *Dev. Cell* **5**, 337–350 (2003).
88. Metzler, M., Wilda, M., Busch, K., Viehmann, S. & Borkhardt, A. High expression of precursor microRNA-155/*BIC* RNA in children with Burkitt lymphoma. *Genes Chromosomes Cancer* **39**, 167–169 (2004).
89. Calin, G. A. *et al.* Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc. Natl Acad. Sci. USA* **101**, 2999–3004 (2004).

90. Krichevsky, A. M., King, K. S., Donahue, C. P., Khrapko, K. & Kosik, K. S. A microRNA array reveals extensive regulation of microRNAs during brain development. *RNA* **9**, 1274–1281 (2003).
91. Knight, S. W. & Bass, B. L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* **293**, 2269–2271 (2001).
92. Wienholds, E., Koudijs, M. J., van Eeden, F. J., Cuppen, E. & Plasterk, R. H. The microRNA-producing enzyme Dicer1 is essential for zebrafish development. *Nature Genet.* **35**, 217–218 (2003).
93. Bernstein, E. *et al.* Dicer is essential for mouse development. *Nature Genet.* **35**, 215–217 (2003).
94. Moussian, B., Schoof, H., Haecker, A., Jurgens, G. & Laux, T. Role of the *ZWILLE* gene in the regulation of central shoot meristem cell fate during *Arabidopsis* embryogenesis. *EMBO J.* **17**, 1799–1809 (1998).
95. Cox, D. N. *et al.* A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev.* **12**, 3715–3727 (1998).
96. Hipfner, D. R., Weigmann, K. & Cohen, S. M. The *Bantam* gene regulates *Drosophila* growth. *Genetics* **161**, 1527–1537 (2002).
97. Xu, P., Vernooy, S. Y., Guo, M. & Hay, B. A. The *Drosophila* microRNA mir-14 suppresses cell death and is required for normal fat metabolism. *Curr. Biol.* **13**, 790–795 (2003).
98. Palatnik, J. F. *et al.* Control of leaf morphogenesis by microRNAs. *Nature* **425**, 257–263 (2003).
99. Chen, C. Z., Li, L., Lodish, H. F. & Bartel, D. P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* **303**, 83–86 (2004).
100. Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P. & Burge, C. B. Prediction of mammalian microRNA targets. *Cell* **115**, 787–798 (2003).
101. Stark, A., Brennecke, J., Russell, R. B. & Cohen, S. M. Identification of *Drosophila* microRNA targets. *PLoS Biol.* **1**, E60 (2003).
102. Calin, G. A. *et al.* 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 (2002).
103. Michael, M. Z., O'Connor, S. M., van Holst Pellekaan, N. G., Young, G. P. & James, R. J. Reduced accumulation of specific microRNAs in colorectal neoplasia. *Mol. Cancer Res.* **1**, 882–891 (2003).
104. Tang, G., Reinhart, B. J., Bartel, D. P. & Zamore, P. D. A biochemical framework for RNA silencing in plants. *Genes Dev.* **17**, 49–63 (2003).
105. Emery, J. F. *et al.* Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes. *Curr. Biol.* **13**, 1768–1774 (2003).
106. Juarez, M. T., Kui, J. S., Thomas, J., Heller, B. A. & Timmermans, M. C. microRNA-mediated repression of *rolled leaf1* specifies maize leaf polarity. *Nature* **428**, 84–88 (2004).

🌐 Online links

DATABASES
The following terms in this article are linked online to:
Entrez: http://www.ncbi.nih.gov/Entrez
AP2 | *cog-1* | *hbl-1* (*lin-57*) | *Hoxb8* | *let-7* | *lin-4* | *lin-14* | *lin-28* | *lin-41* | *lys-6* | *miR-9* | *miR-99a* | *miR-155* | *miR-181* | *miR-196*
TAIR: http://www.arabidopsis.org
DCL1 | *DCL2* | *DCL3*

FURTHER INFORMATION
miRNA Registry: http://www.sanger.ac.uk/Software/Rfam/mirna/
PHYLIP programs:
http://evolution.genetics.washington.edu/phylip.html
Access to this links box is available online.

CORRECTION

MicroRNAs: SMALL RNAs WITH A BIG ROLE IN GENE REGULATION

Lin He and Gregory J. Hannon

In figure 2, the orientation of some RNA structures was incorrect. The corrected version is shown below. This correction has been made to the online enhanced text and PDF version of this review.