Specialization and evolution of endogenous small RNA pathways

Elisabeth J. Chapman and James C. Carrington‡*

Abstract | The specificity of RNA silencing is conferred by small RNA guides that are processed from structured RNA or dsRNA. The core components for small RNA biogenesis and effector functions have proliferated and specialized in eukaryotic lineages, resulting in diversified pathways that control expression of endogenous and exogenous genes, invasive elements and viruses, and repeated sequences. Deployment of small RNA pathways for spatiotemporal regulation of the transcriptome has shaped the evolution of eukaryotic genomes and contributed to the complexity of multicellular organisms.

**Department of Biology, Myers Hall, 915 East Third Street, Indiana University, Bloomington, Indiana 47405, USA.*
‡Center for Genome Research and Biocomputing, 3021 ALS, Oregon State University, Corvallis, Oregon 97331, USA.
Correspondence to J.C.C.
e-mail: carrington@cgrb.oregonstate.edu
doi:10.1038/nrg2179

Genome expansion in higher plants, animals and other multicellular eukaryotes is partially attributed to an increased capacity for the spatiotemporal control of gene expression that is required for development[1,2]. Throughout evolution, protein components of transcriptional, post-transcriptional and post-translational regulatory pathways have proliferated, expanding the potential for regulatory specificity. Proteins that are involved in transcription, for example, represent as much as 10% of the coding sequence in the genomes of metazoans[1]. By contrast, specificity for RNA-mediated silencing-based regulation is conferred by small RNA guides (for example, microRNAs (miRNAs), small interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs)) that are generated through distinct biogenesis pathways and function with specialized effector proteins (members of the Argonaute (AGO)–Piwi family). Therefore, the evolution of the RNA-mediated silencing mechanism as a regulatory device in multicellular organisms has involved a proliferation of small RNA biogenesis and effector proteins, and the *de novo* formation, proliferation and refinement of small RNA guides. How have the core biogenesis and effector modules proliferated and specialized to generate diverse regulatory pathways? How have new specificities emerged, and how has RNA silencing been integrated into regulatory networks?

Here, we review the diverse functions of the basic RNA silencing modules, focusing on three areas in which regulatory proliferation and functional specialization has occurred: diversification and specialization of small-RNA-processing factors, proliferation and specialization of effectors, and *de novo* evolution of small RNA regulators. Although much of the evidence discussed here refers to plant systems, in which the diversity and genome-wide deployment of silencing mechanisms is high, it also includes a broader perspective from other systems. We focus primarily on endogenous regulatory pathways, with only brief discussions of antiviral silencing mechanisms in cases in which they are of particular relevance to the central topic of this Review.

Diversity of small RNA pathways

Early studies of transgene-induced RNA-silencing phenomena, particularly in plants, revealed puzzling differences in the molecular and epigenetic properties that are associated with silencing in distinct individuals[3]. In some cases, silencing was associated with DNA methylation of promoter regions and transcriptional repression whereas, in other cases, it was associated with high levels of transcription of targeted loci but low stability of the resulting transcripts. In retrospect, it is obvious that many of these differences resulted from the entry of trigger loci or transcripts into distinct silencing pathways that evolved through proliferation and specialization of a basic set of modules[4,5]. The core mechanism, as revealed through pioneering studies in worms, flies, plants and other eukaryotic models, is triggered by perfect or near-perfect dsRNA that is processed into small RNA duplexes by RNaseIII-like enzymes[6–10]. The diversity of RNA silencing modules is therefore reflected in distinct types of trigger loci, and in biogenesis and effector factors that form and use distinct classes of small RNA (TABLE 1).

Diversity of small-RNA-generating loci

There are several types of loci that generate functional small RNAs. Some form precursor transcripts that adopt secondary structures with Dicer substrate activity,

Table 1 | **Classes of small RNA identified in eukaryotes.**

Class	Description	Biogenesis and genomic origin	Function
miRNA	MicroRNA	Processing of foldback miRNA gene transcripts by members of the Dicer and RNaseIII-like families	Post-transcriptional regulation of transcripts from a wide range of genes
Primary siRNA	Small interfering RNA	Processing of dsRNA or foldback RNA by members of the Dicer family	Binding to complementary target RNA; guide for initiation of RdRP-dependent secondary siRNA synthesis
Secondary siRNA	Small interfering RNA	RdRP activity at silenced loci (*Caenorhabditis elegans*); processing of RdRP-derived long dsRNA or long foldback RNA by members of the Dicer family (*Arabidopsis thaliana*)	Post-transcriptional regulation of transcripts; formation and maintenance of heterochromatin
tasiRNA	*Trans*-acting siRNA	miRNA-dependent cleavage and RdRP-dependent conversion of *TAS* gene transcripts to dsRNA, followed by Dicer processing	Post-transcriptional regulation of transcripts
natsiRNA	Natural antisense transcript-derived siRNA	Dicer processing of dsRNA arising from sense- and antisense-transcript pairs	Post-transcriptional regulation of genes involved in pathogen defense and stress responses in plants
piRNA	Piwi-interacting RNA	A biogenesis mechanism is emerging, which is Argonaute-dependent but Dicer-independent	Suppression of transposons and retroelements in the germ lines of flies and mammals

RdRP, RNA-dependent RNA polymerase.

and do not require RNA-dependent RNA polymerase (RdRP) or amplification mechanisms. Other loci form small RNAs after the primary transcripts are processed through dsRNA-forming mechanisms that involve RdRP activity. Yet other loci yield small RNAs through non-RdRP-dependent amplification mechanisms.

miRNA loci. miRNAs, which act as post-transcriptional regulators of gene expression, are encoded in the genomes of multicellular eukaryotes and unicellular plants. miRNAs arise from primary transcripts with self-complementarity that undergo a multistep maturation process involving specialized RNaseIII-type enzymes, including Dicer[11,12]. miRNA genes are most often transcribed by RNA polymerase II (PolII), can form in mono- or polycistronic configurations, and occur in both intronic and exonic sequences[13,14]. In plants and animals, miRNAs of similar sequence are grouped into families, the members of which are sometimes regulated differentially and expressed in tissue-specific patterns[15].

siRNAs from inverted and direct repeat sequences. Endogenous siRNAs can arise from loci that contain direct or inverted repeats, and from dispersed repetitive elements[16]. Inverted repeats that yield transcripts with near-perfect self-complementarity are common in plant genomes, and often yield heterogeneous siRNA populations that do not require an RdRP[17,18] (FIG. 1a). Tandem direct repeats, such as those at 5S rRNA loci, and all classes of transposons and retroelements, also spawn siRNA populations, but these are frequently dependent on RdRP activity[19]. The proportion of total small RNA that derives from repeat sequences in *Arabidopsis thaliana* is relatively high, although this overrepresentation is reduced when the proportion is adjusted to account for numbers of repeats[18–20]. The mechanisms by which small RNAs are produced from direct repeat sequences are not entirely clear, but they include specific Dicers, RdRPs and specialized Argonaute proteins such as *A. thaliana* ARGONAUTE4 (AGO4) and *Drosophila melanogaster* PIWI[21–25].

LINE1 (L1) elements
A class of self-replicating retrotransposons that are highly abundant in the human genome.

Small RNA from bidirectional transcripts. Transcription from opposing promoters can yield dsRNA (FIG. 1b). For example, the 5′ untranslated region (UTR) of the human *LINE1* (L1) element contains promoters that support internal bidirectional transcription, and L1-specific siRNAs that derive from the putative double-stranded region accumulate in cultured cells. Production of L1-specific siRNAs is associated with degradation of L1 transcripts and repression of L1 transposition[26].

In plants, natural antisense transcript-derived siRNAs (natsiRNAs) arise from overlapping transcripts that are induced by abiotic or biotic stress. In two examples reported to date, stress-induced transcription results in production of dsRNA from which a specific natsiRNA arises by the activity of one or more Dicer-like (DCL) proteins[27,28]. The natsiRNA then guides repression of the complementary transcript. In some cases, the activity of natsiRNA is accompanied by production of other small RNAs, triggered from the natsiRNA-guided cleavage of the target transcript[27].

Computational analysis of natural antisense transcripts (NATs) and small RNA accumulation patterns in *A. thaliana* indicates that NAT pairs are not a major contributor to small RNA populations in the absence of stress[29]. However, natsiRNA accumulation is strongly induced following relevant stimuli, and NAT pairs are found at more than 7% of the annotated transcriptional units in the *A. thaliana* genome[27–29]. This suggests that the potential for small RNA production from NAT pairs in response to general or specific stimuli might be underestimated.

TAS loci. *Trans*-acting siRNAs (tasiRNAs) arise in plants from specific *TAS* loci (FIG. 1c). *TAS* transcripts are PolII-dependent, and function as highly specialized precursors that feed into an RdRP-dependent siRNA-biogenesis pathway. They are targets for cleavage by miRNA-guided mechanisms, and yield siRNAs that are in a 21-nucleotide register with the cleavage site. Specific tasiRNAs from at least four families function post-transcriptionally, as do miRNAs[20,30–34].

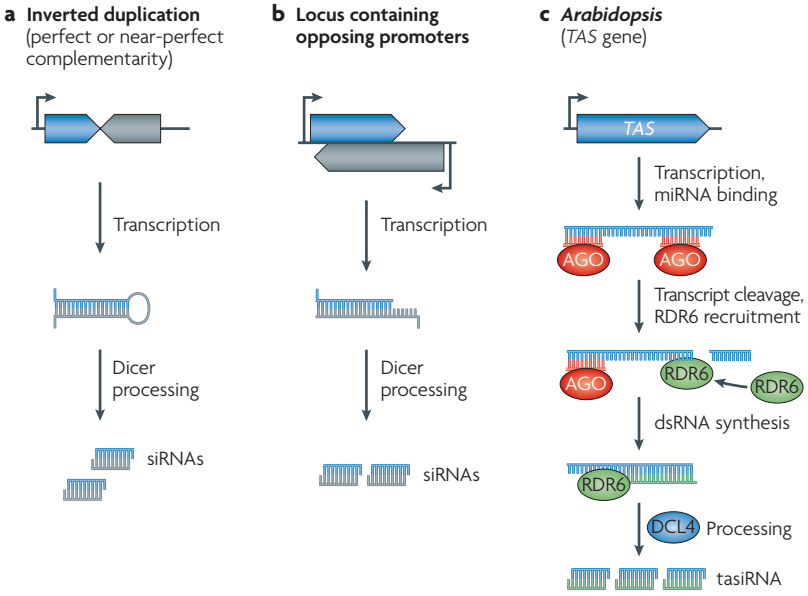


Figure 1 | **Endogenous siRNA-generating loci. a** | Small interfering RNAs (siRNAs) arise from some loci that contain inverted repeats. Transcripts from these loci form foldback structures with near-perfect complementarity. siRNA production probably results from Dicer processing of foldback transcripts from these loci. Accumulation of siRNAs from direct or tandem repeats (not shown) is frequently dependent on RNA-dependent RNA polymerase activity. **b** | Loci that generate bidirectional transcript pairs can spawn siRNAs, such as natural antisense transcript-derived siRNAs (natsiRNAs) in *Arabidopsis thaliana* and small RNAs that are associated with genome rearrangement in *Tetrahymena thermophila*. Specific mechanisms for small RNA biogenesis from these loci differ among lineages, but each is likely to involve processing of dsRNA into siRNA by members of the Dicer family. **c** | *Trans*-acting siRNA (tasiRNA) are endoded at *TAS* loci in plants and result from DICER-LIKE 4 (DCL4) processing of RNA-DEPENDENT RNA POLYMERASE 6 (RDR6)-dependent dsRNA. dsRNA synthesis is triggered by microRNA (miRNA)-guided cleavage of the primary transcript, and is processed in a phased, 21-nucleotide register from the miRNA-guided cleavage site. AGO, Argonaute protein.

piRNA-generating loci. Piwi-interacting RNAs (piRNAs, also referred to as repeat-associated siRNAs) are found in the germline cells of mammals and insects. In *D. melanogaster* piRNAs arise from transposons and repeat regions such as *Supressor of Stellate* (*Su(Ste)*), whereas in mammals piRNAs originate from developmentally regulated gene clusters that frequently contain transposons and retroelements[35–37]. In both systems, piRNAs associate with germline-specific Argonaute proteins of the Piwi subclass, which are required for their biogenesis. piRNAs are associated with control of transposition in the germ line[23,38–41].

siRNAs from exogenous agents. Exogenous triggers, such as experimentally introduced dsRNA, transgenes and viruses, are well-known initiators of RNA silencing. In plants and some animals, such as *Caenorhabditis elegans*, distinct siRNA populations form during primary and secondary phases. Primary siRNAs arise directly from trigger RNA molecules. Secondary siRNAs arise by distinct mechanisms in plants and *C. elegans*, but in both cases an RdRP engages with a target transcript that interacts with a primary small RNA–AGO complex[42–44]. RdRP-dependent siRNAs can function at cell-autonomous and non-cell-autonomous levels, leading to the spread of silencing signals throughout a tissue or organism[45].

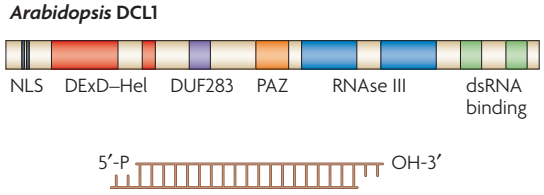
Proliferation of biogenesis factors

Most types of small RNA are processed from precursor RNA by members of the Dicer family (BOX 1). Several mechanisms have evolved that regulate the biogenesis of specific classes of small RNA using variations on the same core processing machinery. Most animals encode a single Dicer protein that catalyses the formation of multiple classes of small RNA with diverse functions, and specificity is achieved through the interaction of Dicer with other proteins. In other cases, specificity results from the existence of multiple members of Dicer-family proteins in a single species.

Specificity through Dicer-interacting proteins. In *C. elegans*, specification of Dicer function requires interactions between Dicer 1 (DCR-1) and other proteins to produce siRNAs from exogenous dsRNA, process miRNAs from hairpin precursors, or produce endogenous siRNAs[46–49]. The DCR-1-interacting proteins Argonaute-like 1 and 2 (ALG-1 and ALG-2) are required for miRNA processing. RNAi defective 4 (RDE-4), a dsRNA-binding protein that interacts with DCR-1, is required for the accumulation of small RNAs from a locus on the X chromosome (the X cluster)[46]. Another interesting group of DCR-1 interactors consists of proteins that were previously identified in *C. elegans* mutants with an enhanced RNAi response (*eri* mutants)[50,51]. Loss of pathway-specific ERI proteins seems to free DCR-1 to enter other RNA silencing pathways, thus boosting activity in those pathways[46]. This suggests that DCR-1 is limiting, and competition among RNA silencing pathways for DCR-1 might prevent aberrant dicing by titrating out DCR-1 to suitable templates[46,49]. A similar competition model was proposed to explain small RNA accumulation patterns in *A. thaliana*, in which specific DCL proteins faithfully generate size-specific small RNA classes from endogenous and exogenous precursors[52]. Dicer-interacting dsRNA-binding proteins are also implicated in the production of specific small RNA classes in *D. melanogaster*[53–55].

siRNA-specific biogenesis factors. Two members of the Dicer family are encoded in the *D. melanogaster* genome, Dicer 1 (DCR1) and Dicer 2 (DCR2). DCR2 acts in concert with the dsRNA-binding protein R2D2 to process siRNA precursors (dsRNA)[53,54]. The fungi *Neurospora crassa* and *Magnaporthe oryzae* also encode two Dicers. In *M. oryzae* the Dicer protein MDL-2 is required for siRNA accumulation and RNA silencing, whereas in *N. crassa* the two Dicers are redundant[56,57]. The ciliate protozoan *Tetrahymena thermophila* encodes three Dicers, of which *DCL1* is required for the accumulation of ~27–30-nucleotide RNAs that are associated with genome rearrangement in conjugating cells, but is dispensable for the accumulation of 23–24-nucleotide RNAs[58,59].

Box 1 | **Dicer proteins and the anatomy of small RNA duplexes**



Dicers are large, ~200 kDa proteins that generally contain DExD and –C ATPase/helicase domains (DExD–Hel), a domain of unknown function (DUF283), a PAZ domain, two catalytic RNase III domains and a dsRNA-binding (dsRB) domain. The genomes of *Schizosaccharomyces pombe*, *Chlamydomonas reinhardtii*, *Caenorhabditis elegans* and vertebrate animals encode one Dicer protein, whereas *Drosophila melanogaster*, *Tetrahymena thermpohila* and *Neurospora crassa* each encode two Dicers. The single *C. reinhardtii* Dicer DCR1 lacks the PAZ and dsRB domains, and Dicers of the ciliate *T. thermophila* lack these as well as the DUF283 domain[178]. In other organisms, all seven characteristic elements of Dicer are represented, although only plants encode Dicer proteins that contain all seven elements in a single protein. Plants encode four or more Dicer-like (DCL) proteins, and of these, DCL1, DCL3 and DCL4, but not DCL2, contain duplicated dsRNA-binding domains[60].

Dicer processing of dsRNA results in production of a small RNA duplex, with each strand being 21–24 nucleotides in length and containing a 5′ monophosphate (5′-P) and a 3′ hydroxyl (3′-OH). The duplex is paired such that each 3′ end contains a 2-nucleotide overhang. NLS, nuclear localization signal.

The *A. thaliana* genome encodes three siRNA-generating DCL enzymes, which cleave both endogenous and exogenous precursors[60]. Although DCL3 is the primary enzyme responsible for generating 24-nucleotide siRNAs on a genome-wide scale, in *dcl3* mutants DCL2 and DCL4 have access to DCL3 substrates and produce 22-nucleotide and 21-nucleotide siRNAs from RDR2-dependent precursors[18,52]. DCL3 does not contribute significantly to miRNA accumulation, and is dispensable for the production of most RNA-virus-derived siRNAs, and thus might be specialized towards processing of RDR2-dependent dsRNA products[19]. DCL2 contributes to siRNA production from exogenous elements, and is also involved in processing of natsiRNA, although the mechanisms by which it does this are unclear[27]. DCL4, along with dsRNA-binding protein 4 (DRB4), is required for processing of 21-nucleotide tasiRNAs[31,52,61–65]. DCL4 also acts in a hierarchical manner with DCL2 to produce secondary siRNAs from viruses and transgenes[66–68].

The notable proliferation of DCL-family members in plants might in part be a consequence of the deployment of RNA silencing as an antiviral defense mechanism, as suggested by Deleris *et al.*[66]. Virus-derived siRNAs programme antiviral RNA-induced silencing complex (RISC) activity and propagate an antiviral signal to distal tissues[66,67,69,70]. Plant viruses encode RNA silencing suppressor proteins that inhibit one or more points in the silencing pathway[71]. The coat protein of *Turnip crinkle virus* functions as a suppressor of RNA-mediated silencing by inhibiting DCL4 activity[66,72–74], which leads to hyper-accumulation of DCL2-dependent 22-nucleotide siRNAs. Similarly, loss of the *Cauliflower mosaic virus*-derived 24-nucleotide siRNA in *dcl3* mutants is accompanied by an increased accumulation of virus-derived 21-nucleotide siRNAs[75]. These findings indicate that multiple members of the DCL family collectively contribute to the plant's defence against a diverse range of viruses, and that loss or suppression of activity of one DCL can be compensated for by the activities of other DCL-family members. Thus, a defensive advantage that is conferred by the existence of multiple antiviral DCL activities might have contributed to proliferation of the DCL family in plants.

Role of RdRPs in siRNA biogenesis. In *C. elegans,* the RdRP RRF-1 is required for RNAi in somatic cells, and is specifically required for accumulation of secondary siRNAs at the target loci of primary siRNAs[42,76]. Accumulation of secondary siRNAs, which are primarily antisense to the target transcript, is required for potent silencing of the target gene. Recent, surprising results indicate that secondary siRNAs are probably manufactured exclusively during the RdRP reaction, without the need for subsequent dicing of dsRNA[42,77]. Synthesis of secondary siRNAs can initiate in a primer-independent manner at positions that lie proximal to the primary-siRNA–mRNA target-duplex site, and generates pools of siRNAs that have a unit length of 22 nucleotides[42,77] (FIG. 2a). Strong antisense-strand bias, and the presence of 5′-triphosphate residues, suggest that secondary siRNAs are the products of staggered RdRP-initiation and RdRP-termination events, although the mechanisms by which the unit length of siRNAs is 'measured' during a synthetic reaction are unclear.

RdRPs also have key roles in several silencing pathways in plants, although their functions are distinct from those of *C. elegans*. Accumulation of secondary siRNAs during post-transcriptional silencing of exogenous nucleic acids, including many viruses and transgene transcripts, is dependent on RDR6 in *A. thaliana*[68,70,78,79]. RDR6 also functions to amplify siRNAs from dozens of endogenous loci, many of which form transcripts that are targeted by one or more miRNAs[34,44,80]. RDR6-dependent secondary siRNAs accumulate in a non-strand-biased pattern, often initiate near the primary-siRNA–miRNA target site and form primarily through the activity of DCL4 (FIG. 2b). Complementarity between the 3′ end of the siRNA–miRNA initiators and the RDR6 template strand is required, which is consistent with a priming model. However, primary-siRNA–miRNA-guided cleavage of the transcript at a single site is not sufficient for the accumulation of secondary siRNAs[34,68].

A. thaliana RDR6 is also required for the biogenesis of tasiRNAs through a refined RNAi-derived mechanism[31–33,81] (FIG. 1b). RDR6 converts processed precursor transcripts into dsRNAs that are diced by DCL4. Although phased tasiRNA accumulation patterns are reminiscent of secondary-siRNA accumulation patterns in *C. elegans,* tasiRNAs accumulate as a result of sequential, phased (21-nucleotide) processing by DCL4 from one end of the dsRNA. The evolutionary refinement of this process involves the integration of an miRNA as a site-specific RNA-processing factor to form one end of the precursor. This facilitates the formation of highly specific tasiRNA guides, which have precise ends and are complementary to discrete sequences in their target mRNA.

RISC
(RNA-induced silencing complex). An Argonaute protein–small RNA complex that inhibits translation of target RNAs through degradative or non-degradative mechanisms.

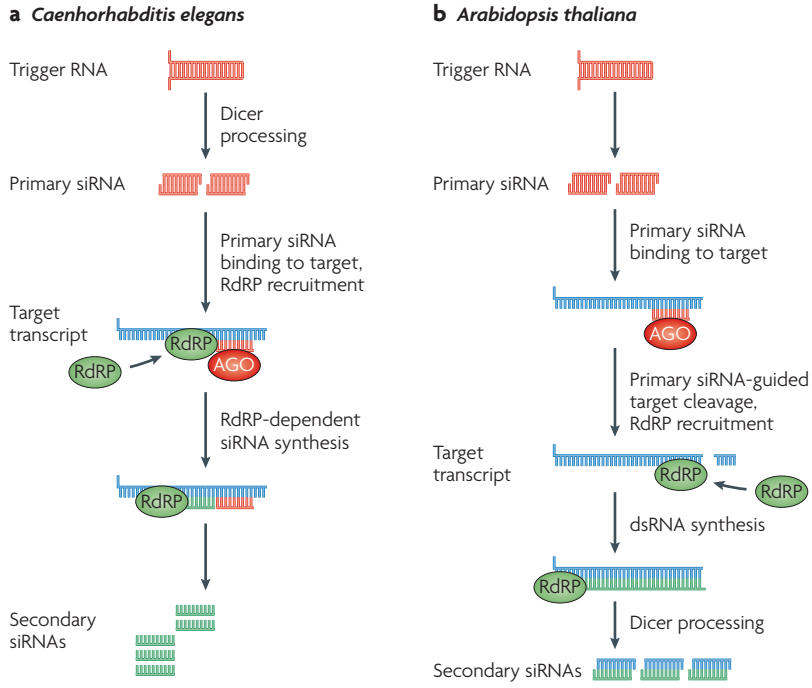


Figure 2 | **Models for amplification of silencing signals in *Caenorhabditis elegans* and *Arabidopsis thaliana*. a** | Processing of trigger dsRNA in *C. elegans* by Dicer 1 (DCR-1) releases primary small interfering RNAs (siRNAs). The primary siRNA associates with an Argonaute (AGO) protein, such as RDE-1, and the complexes bind to complementary target RNA. The bound complexes might then recruit RNA-dependent RNA polymerase (RdRP), which uses the target transcript as template for synthesis of the secondary siRNA. Abundant secondary siRNAs are formed by independent initiation events (rather than by Dicer-mediated processing), are complementary to the target RNA and accumulate in phased pools. **b** | Processing of trigger dsRNA in *A. thaliana* by one or more Dicer-like enzymes releases primary siRNAs, which associate with an Argonaute protein, such as AGO1, and guide cleavage of the target RNA. This event is proposed to recruit an RdRP, such as RDR6, which uses the target transcript as template for synthesis of a long dsRNA. This dsRNA precursor is processed by DCL enzymes to release abundant secondary siRNAs.

How are RdRPs specifically recruited to transcripts that generate small RNAs? In plants, production of siRNAs from miRNA-regulated transcripts is surprisingly rare on a genome-wide scale, suggesting that small-RNA-guided cleavage or interaction *per se* is insufficient. However, RDR6–DCL4-dependent siRNAs are frequently associated with transcripts that are targeted by more than one small RNA[20,34,44,80]. Notable in this group are transcripts from *TAS3* loci, which contain two target sites for miR390 and require a specific Argonaute protein (AGO7) for tasiRNA accumulation[33,82,83], and transcripts that encode a set of pentatricopeptide repeat (PPR) proteins[34,80]. In the *C. elegans* system, accumulation of secondary siRNAs also requires a specific Argonaute, RDE-1, which associates with the primary siRNA[42]. However, aside from transcript association with one or more silencing complexes, the mechanisms by which RdRP is recruited to target transcripts are unclear.

In addition to RDR6-dependent siRNAs, *A. thaliana* accumulates endogenous 24-nucleotide siRNAs from transposons, retroelements and other features through an RDR2–DCL3-dependent mechanism[17–19,84]. The RDR2 homologue MOP1 is required for production of 24-nucleotide siRNAs in maize[85]. Both RDR2 and MOP1 are associated with RNA-directed DNA methylation and chromatin modification at multiple loci[19,84,86,87].

miRNA-specific biogenesis factors. Most miRNAs arise from PolII transcripts, and in many lineages miRNA production is dependent on specialized biogenesis factors (FIG. 3). The basis for the existence of biogenesis requirements that are distinct from those of siRNA production centres on miRNA foldback precursors, which adopt imperfectly paired stem–loop structures. A unique adaptation of miRNA foldbacks is recognition by an enzyme complex that catalyses the initial cleavage events that form the precursor miRNA (pre-miRNA)[55,88]. In animals, this initial step involves the nuclear 'microprocessor' complex (which contains the Drosha and DiGeorge syndrome critical region gene 8 protein (DCGR8))[89]. Initial processing in flies is done by Drosha in concert with Pasha, a DGCR8 homologue[55]. Further processing to yield a miRNA–miRNA* duplex (where miRNA* is largely complementary to the miRNA, and is later discarded) is catalysed by Dicer and *trans*-activation responsive RNA-binding protein (TRBP) in mammals[90,91], or by DCR1 and Loquacious (a TRBP homologue) in flies[92–94]. Mutational analysis of primary miRNA transcripts indicates that the length of the stem and other structural features, such as the terminal loop, determine accurate processing[95,96]. Intriguingly, some intron-derived miRNAs bypass microprocessor steps, and instead use debranched introns that mimic structural features of pre-miRNA Dicer substrates[97,98].

Specialization of factors involved in miRNA biogenesis is also apparent in plants. miRNA accumulation in *A. thaliana* requires DICER-LIKE 1 (DCL1, which primarily yields 21-nucleotide products) and the dsRNA-binding protein HYPONASTIC LEAVES 1 (HYL1), which contains regions of similarity to *D. melanogster* R2D2 (REFS 63,99–104). Weak *dcl1* mutant alleles trigger severe developmental and reproductive defects, whereas strong alleles are embryonic lethal[105]. Null mutations in *HYL1* are mild in comparison, and many miRNAs still accumulate to detectable levels in *hyl1* mutants, suggesting that this gene is partially redundant. Plants encode several other dsRNA-binding proteins in the HYL1–DRB family, and it is possible that one or more of these proteins is able to substitute for HYL1 function[73,106]. DCL1 is not required for accumulation of heterochromatin-associated 24-nucleotide siRNAs, and does not contribute significantly to accumulation of virus-derived siRNAs from several RNA viruses[19,73]. However, DCL1 is necessary for accumulation of virus-derived siRNAs that arise from a predicted imperfect foldback RNA structure from *Cauliflower mosaic virus*, and contributes to siRNA accumulation from a foldback trigger of RNA silencing[43,75]. These findings reinforce the idea that DCL1 has specialized to produce 21-nucleotide RNAs from imperfect foldbacks, the most recognizable of which are transcripts from miRNA genes.

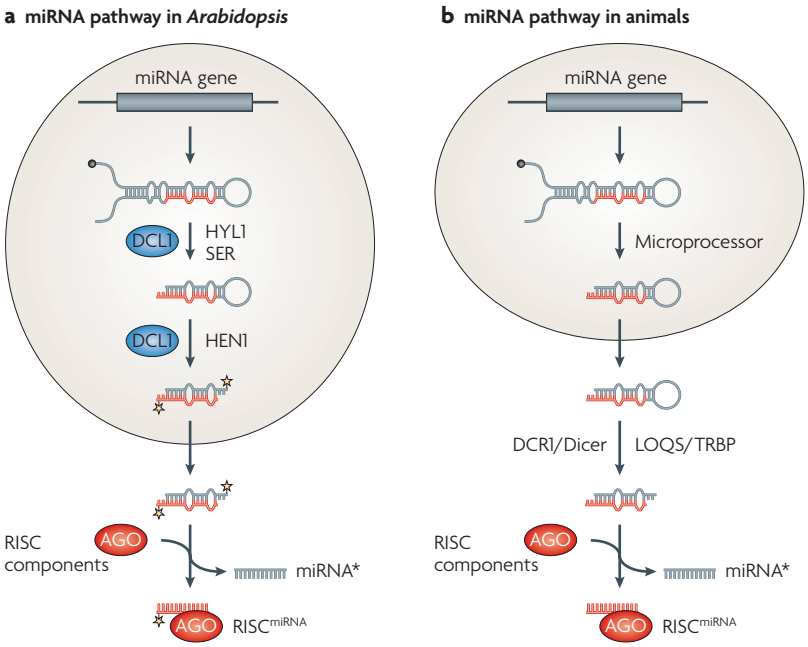


Figure 3 | **RNA biogenesis in plants and animals. a** | Transcripts from *Arabidopsis thaliana* microRNA (miRNA) genes adopt an imperfect foldback structure, and are processed by DICER-LIKE 1 (DCL1) in concert with HYPONASTIC LEAVES 1 (HYL1) and SERRATE (SE) to release a miRNA–miRNA* (miRNA* represents the antisense strand) duplex. This duplex is methylated by the methyltransferase HUA ENHANCER 1 (HEN1) and exported to the cytoplasm where the mature miRNA strand associates with an Argonaute (AGO) protein, most commonly AGO1, to form the RNA-induced silencing complex (RISC). **b** | Transcripts from animal miRNA genes adopt an imperfect foldback structure and undergo sequential processing by the microprocessor complex (Drosha and Pasha in flies; Drosha and DiGeorge syndrome region gene 8 protein (DCGR8) in mammals) to form precursor miRNA (pre-miRNA) intermediates. Pre-miRNAs are exported to the cytoplasm and processed by Dicer in concert with a specialized RNA-binding protein, such as Loquacious (LOQS) in flies or *trans*-activation responsive RNA-binding protein (TRBP) in humans.

Arabidopsis thaliana DCL2, DCL3 and DCL4 collectively have a limited capacity to process imperfect foldbacks from miRNA loci[19,52,107]. A notable exception to this generalization is the dependence of non-conserved *A. thaliana* miR822 and miR839 on DCL4 processing[20]. These loci are likely to be evolutionarily 'young', and encode foldback arms that are relatively long and more perfectly paired. These loci, as described below, might have yet to 'adapt' to the canonical DCL1-dependent pathway that operates on imperfect miRNA foldbacks. Thus, although plant Dicers have diversified in a functional sense, together they provide a set of small-RNA-processing activities that contribute to the evolution of novel regulatory circuits.

piRNA biogenesis factors. piRNAs do not form through Dicer-dependent mechanisms. Rather, piRNA guide strands form by a process that yields overlapping (offset by 10 nucleotides) complementary strands (FIG. 4). In flies, guide strands and complementary strands are frequently divided between Aubergine (AUB)-interacting or Piwi-interacting pools and Argonaute 3 (AGO3)-interacting pools. The model proposed for piRNA biogenesis involves cleavage of transposon-derived transcripts by a guide piRNA–Piwi or piRNA–AUB complex, generating a new RNA 5′ end at a position opposite nucleotides 10 and 11 of the piRNA[22,23]. The cleavage product is then incorporated into an AGO3 complex and undergoes further 3′-end processing to release a new piRNA. Similar models have been proposed for the roles of Piwi-like proteins in piRNA biogenesis in mammals[37]. These models suggest that hierarchical, specialized Argonaute proteins working in concert have acquired new functions in the biogenesis of a new small RNA class[22,23].

Proliferation and specialization of effectors

The effector steps of RNA silencing are mediated by the RISC, the activity of which can be reconstituted with a small RNA guide and an Argonaute protein[108–111]. Of the factors involved in RNA silencing, it is the members of the Argonaute family that have proliferated to the greatest extent in a number of organisms. Studies of this family indicate that significant diversification and specialization of function supported the evolution of new small RNA classes and regulatory cascades in several lineages.

The *A. thaliana* genome encodes ten members of the Argonaute-like subfamily of Argonaute proteins and no Piwi-like members. Each of these ten proteins contains the Arg–Arg–Arg–His peptide motif that is associated with slicer function[112], and slicer activity has been demonstrated for AGO4 (REF. 113) and AGO1 (REFS 111,114). These two Argonautes interact preferentially with different small RNA classes and thus have specialized roles: AGO4 is required for functionality of 24-nucleotide siRNAs at heterochromatic sites[21,113,115], and AGO1 interacts with miRNAs and tasiRNAs, but not with 24-nucleotide siRNAs[111,113], and is also required for miRNA activity[111,116,117]. Enriched AGO1 complexes that contain endogenous miR165 were shown to catalyse cleavage of target mRNA *in vitro*, and this slicer activity was dependent on conserved residues in the catalytic Piwi domain[111]. These results, coupled with the phenotype of *ago1* mutants, suggest that AGO1 is the predominant effector of miRNA-mediated regulation in *A. thaliana*. AGO1 also interacts with transgene- and virus-derived siRNAs, which is consistent with the impaired silencing and antiviral defence that are observed in *ago1* mutants[43,116–118].

Is there a molecular basis for the selective interaction of AGO1 with 21-nucleotide RNAs, specifically miRNAs, and of AGO4 with 24-nucleotide RNAs? It has been proposed that small RNA biogenesis and effector programming are linked such that specific Dicer products are funnelled to the corresponding Argonaute through Dicer–Argonaute interactions[111,119]. Indeed, DCL1 products (miRNA) represent more than 90% of AGO1-associated small RNAs, and DCL3 products (24-nucleotide siRNAs) represent the majority of AGO4-associated RNA[113]. In *D. melanogaster*, the structure of small RNA duplexes allows small RNA 'sorting', in which miRNAs and siRNAs programme AGO1 and AGO2 complexes, respectively[119,120]. Sorting of siRNAs into AGO2 complexes is facilitated by DCR2–R2D2,

and a parallel process is proposed to facilitate AGO2 programming[119,120]. Further characterization of putative Dicer–Argonaute interactions is needed to understand the mechanisms of Argonaute programming in *A. thaliana* and other organisms.

The Argonaute family in *C. elegans* has expanded to contain 27 predicted members, all of which have been analysed using forward or reverse genetics or biochemical approaches[121]. Several members of this family seem to have specialized functions. The highly similar ALG-1 and ALG-2 are dispensable for RNAi but are required for growth, development and germline maintenance, as well as for miRNA processing and function[122]. CSR-1 is required for chromosome segregation and germline RNAi, and PRG-1 is thought to function in germline maintenance[121,123].

Other Argonaute proteins in *C. elegans* serve specialized functions in the RNAi pathway, which involves production of primary siRNAs from an RNAi trigger, followed by interaction of primary siRNAs with the target mRNA[42,77] (FIG. 4). This event recruits an RdRP to the target, which is used as a template for the synthesis of secondary siRNAs. The Argonaute protein RDE-1 interacts with low-abundance primary siRNAs derived from trigger dsRNAs, but does not significantly interact with the abundant secondary siRNAs that arise from RdRP-dependent amplification of silencing at the target locus[42,77,121,124]. Secondary siRNAs arise from regions within and upstream of the target sequence and interact with SAGO-1 and SAGO-2, two Argonaute proteins that seem to be limiting factors in the RNAi response and that lack the catalytic residues associated with slicer function[121,125]. This Argonaute protein hierarchy is proposed to protect cells against off-target effects that might accompany transitivity[121]. According to the model, if interactions were to occur between RDE-1 and secondary siRNAs, RdRP-dependent amplification of siRNA production could spread into non-targeted, essential genes. Limiting primary-siRNA–Argonaute interactions to non-slicer Argonautes would prevent the biogenesis of secondary siRNAs from regions distal to the initially targeted element. Thus, specialization among the *C. elegans* Argonaute proteins through interaction with specific small RNA classes allows the animal to mount a potent, amplified RNAi response against target loci while protecting neighbouring genes.

A most intriguing example of Argonaute protein specificity lies in the proposed role of Piwi-like Argonaute proteins in piRNA biogenesis and repression of transposon and retroelement loci in the germ lines of flies and mammals. In *D. melanogaster*, protection of the germ line from mobilization of transposons and other mobile genetic elements is essential to prevent transposon and retroelement reactivation and eventual hybrid dysgenesis[126]. piRNAs arise in the germ line and silence retrotransposons and repetitive elements such as *gypsy* transposons and *Stellate* (*Ste*)[23,40]. Piwi-family members AUB, Piwi and AGO3 are expressed in the germ line[23,123], and AUB is required for silencing of *Ste* and repetitive loci, such as transposons that contain long terminal repeats (LTRs)[40]. These expression patterns and activities are distinct from those of AGO1, which is broadly expressed and interacts with miRNAs, and of AGO2, which interacts with exogenous siRNAs but is not required for silencing at the *Ste* locus[127,128]. The intriguing finding that Piwi-type proteins are involved in piRNA biogenesis and activity in both flies and mammals suggests that the proliferation of Argonaute proteins might have enabled the evolution of a new small RNA pathway for the control of mobile elements in the germ line[22,23,41].

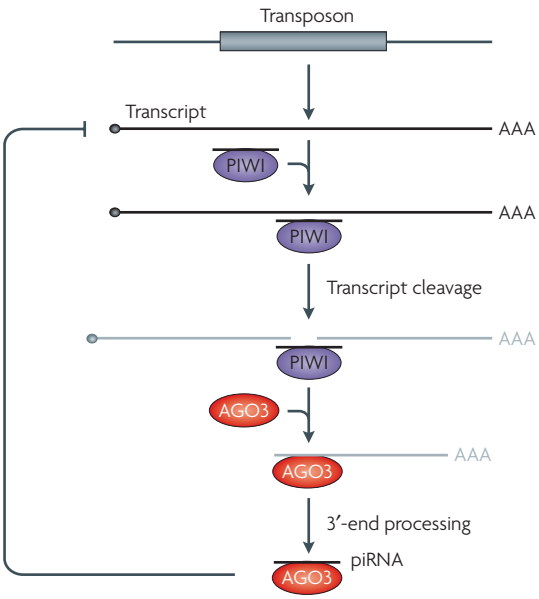


Figure 4 | **Model for Piwi-interacting RNA (piRNA) biogenesis.** piRNAs are found in association with members of the Piwi subfamily of Argonaute proteins. The proposed piRNA-biogenesis model involves initial targeting of transcripts from transposons and retroelements by a Piwi-like protein that is programmed with a small RNA. Cleavage of the transcript generates the 5′ end of a new piRNA. Further 3′-end processing might require a distinct Piwi-like protein, such as *Drosophila melanogaster* AGO3, generating a new piRNA with a 3′ end that is offset by 10 nucleotides from the initial small RNA.

Functional consequences of diversification

A consequence of the diversification of small RNA biogenesis and effector factors is an expansion of functionality of silencing at different regulatory levels. These include regulation at the post-transcriptional, DNA-rearrangement and epigenetic levels. Regarding the first of these levels, the connections between small RNA functions and translational repression, slicer-dependent and slicer-independent degradation of transcripts, and transcript localization were reviewed recently in REF. 129, and will not be elaborated here.

Regulation of DNA rearrangement. Ciliates such as *T. thermophila* are characterized by nuclear dualism, in which two distinct nuclei have roles in germline maintenance (micronucleus) and gene expression (macronucleus).

Transitivity
The spreading of silencing to regions that flank the original target sequence.

Hybrid dysgenesis
Descibes phenotypes that result from a high rate of mutation in germline cells of *Drosophila melanogaster*, triggered by P-element transposition.

The transcriptionally active macronucleus differs from the silent micronucleus in DNA content. Formation of the macronucleus involves site-specific loss of DNA, including transposon sequences, through a pathway that depends on RNA silencing factors[130]. Bidirectional transcripts that arise from deletion elements in conjugating cells accumulate before DNA deletion, and probably provide dsRNA precursors for *DCL1*-dependent production of 27–30-nucleotide sRNAs. TWI1, an Argonaute protein, is proposed to complex with 27–30-nucleotide sRNAs and mark complementary DNA sequences for methylation and subsequent elimination[131,132].

Epigenetic regulation. In some cases, endogenous siRNAs and piRNAs trigger epigenetic effects at target loci and are associated with RNA-directed DNA methylation and chromatin remodelling[19,21,84,133]. For example, transcription of the maize Mu killer locus (*Muk*), an inverted duplication of a transposon sequence from the *Mutator* family, results in accumulation of *Muk*-derived siRNAs and heritable silencing of the *Mutator* transposon family[134,135]. *Mutator* and other transposable elements are implicated in the regulation of multiple distinct epialleles[136,137]. siRNAs that arise from direct duplications (tandem repeats) in the promoter of the *FWA* locus in *A. thaliana* are implicated in the initiation of DNA methylation and epigenetic control[84,138]. In mice, loss of the piRNA-associated Argonaute proteins MILI (also known as PIWIL2) or MIWI2 (also known as PIWIL4) is associated with reduced DNA methylation of *LINE1* elements and increased *LINE1* expression[37,139]. In *D. melanogaster,* deficienct accumulation of piRNAs from various transposable elements in *spn-E* mutants is associated with derepression of chromatin at retrotransposon loci[41]. Thus, the influence of small RNA pathways on DNA methylation, chromatin structure and epigenetic control in multicellular plants and animals is now well established. Furthermore, in fission yeast (*Schizosaccharomyces pombe*), small RNA pathways are associated with repressive chromatin[140] and centromere function[141].

In *A. thaliana,* small-RNA-dependent DNA methylation and chromatin modification are mediated through RDR2, DCL3 and AGO4 (REF. 84) (FIG. 5). Two isoforms of PolIV, a DNA-dependent RNA-polymerase-like protein that is specific to plants, are also required at some loci[142–146] (FIG. 4). One model states that primary transcripts that are stalled by methylation are recognized as aberrant and attract one isoform, PolIVa, to the locus. However, it must be noted that there are no positive data indicating that either PolIV isoform has polymerase activity. Regardless of how they are made, transcripts from target loci are converted by RDR2 into a dsRNA substrate for DCL3, which catalyses 24-nucleotide-siRNA formation, in nucleolar Cajal bodies[145,146]. These siRNAs then programme AGO4, which, in association with a subunit of Pol IVb[145,146], may form a complex that guides DNA methylation and chromatin modification (such as histone 3 lysine 9 (H3K9) methylation) at the locus. Genetic analysis also implicates chromatin-remodelling factors, several

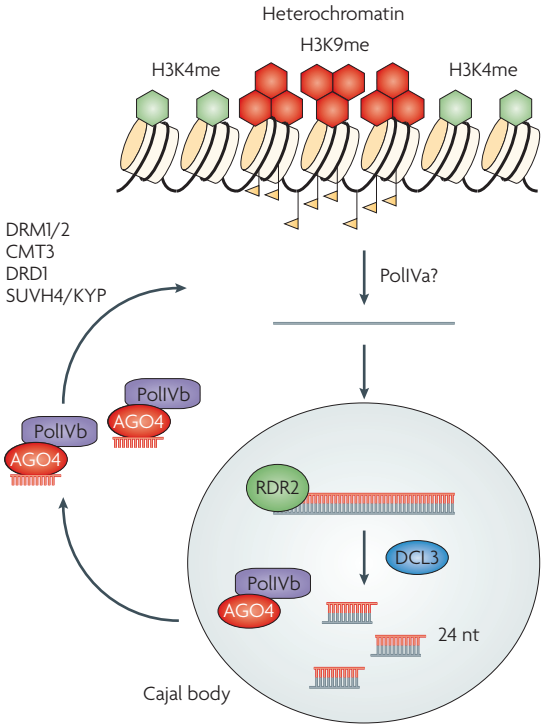


Figure 5 | **The chromatin-associated small interfering RNA (siRNA) pathway in *Arabidopsis thaliana*.** RNA-directed DNA methylation (for example, at histone H3 lysines 4 and 9 (H3K4me and H3K9me, respectively) and chromatin remodelling in *A. thaliana* involves 24-nucleotide siRNAs formed through an RNA-DEPENDENT RNA POLYMERASE 2 (RDR2)–DICER-LIKE 3 (DCL3)–POLYMERASE IVA (PolIVa)-dependent pathway. Effector complexes containing siRNAs, ARGONAUTE 4 (AGO4) and PolIVb direct DNA and chromatin modifications through the activities of many factors, including DOMAINS REARRANGED METHYLASE 1 and 2 (DRM1 and DRM2), CHROMOMETHYLASE 3 (CMT3), DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1) and SU(VAR)3-9 HOMOLOGUE 4 (SUVH4) (also known as KYP).

DNA methyltransferases, histone methyltransferases and other chromatin proteins in RNA-directed DNA and chromatin modifications[147].

In *S. pombe*, an RNAi-induced transcriptional silencing complex (RITS) contains Argonaute 1 (Ago1), a chromodomain protein (Chp1) and other proteins[148]. This complex recruits an RdRP-containing complex[149,150], and its distribution across the genome correlates with patterns of H3K9 methylation[151]. RITS and the RdRP-containing complex are proposed to act in a self-reinforcing loop in which DNA-interacting proteins and the siRNAs in RITS guide H3K9 methylation and heterochromatin formation, as well as RdRP-mediated conversion of transcripts into siRNA precursors[149,152].

De novo evolution of small RNAs

At present, there are no generally accepted examples of canonical miRNAs that are conserved between plants and animals. Yet, both plants and animals contain

Epiallele
An allele for which variable methylation or chromatin states confer heritable variable expression among individuals.

Cajal bodies
Nuclear bodies that are associated with the assembly of the gene expression machinery.

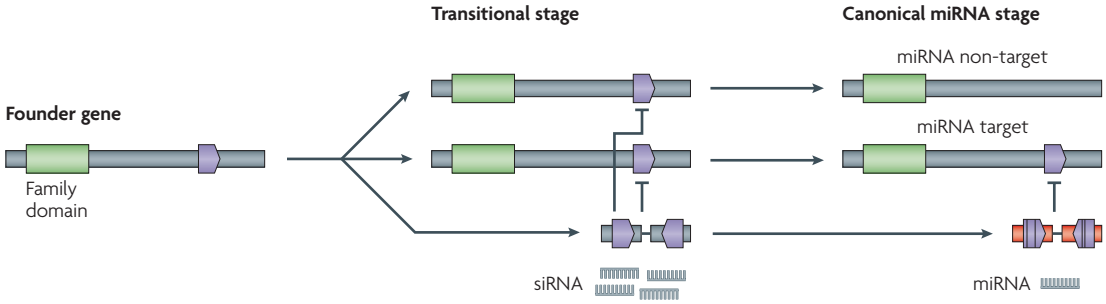


Figure 6 | **Models for genesis and evolution of microRNA (miRNA) loci in plants.** Plant miRNA genes are proposed to originate and evolve from inverted duplications of protein-coding genes. During a transitional stage, transcripts from inverted duplications fold into long hairpin-like structures that are processed to heterogeneous small interfering RNAs (siRNAs) by Dicer-like (DCL)-family members. The potential for conformation to the canonical miRNA-biogenesis pathway, which involves site-specific processing by DCL1, occurs by mutations that result in an imperfect foldback structure.

well-characterized miRNAs that function as negative regulators at the post-transcriptional level. How do miRNA loci arise, and how does miRNA-target specificity evolve?

Origin of miRNA genes. Several non-conserved plant miRNA genes have recently been shown to contain extensive similarity to protein-coding genes[153,154]. This led to development of a model for the genesis of miRNA loci in plants in which products of inverted duplications produce small RNAs with regulatory potential[153]. Inverted duplications that contain sequences from protein-coding genes may form a locus with regulatory potential to control the gene or gene family of origin. If small RNAs that arise from the novel locus are coexpressed with and regulate transcripts from the progenitor gene, and if this regulation were advantageous, the novel locus and its regulatory function would be maintained. Over time, mutation and sequence changes that occur by genetic drift would lead to imperfect pairing in the foldback structure and adaptation to the specialized miRNA-biogenesis machinery (FIG. 6).

This evolutionary model makes several predictions: primary transcripts from relatively new, non-conserved miRNA loci will contain extensive similarity to their progenitor genes; new miRNA genes will encode near-perfect foldback precursors that are longer than those encoded by conserved, highly evolved miRNA genes; and these precursors might not be strictly dependent on miRNA-biogenesis factors, but rather might be processed by siRNA-biogenesis factors to form heterogeneous populations. Recent findings are in accordance with these predictions. Several miRNA-generating loci in *A. thaliana* encode hairpin-forming RNA transcripts that have extensive similarity to protein-coding genes[20,155]. Some of these loci encode long foldback precursors with near-perfect base-pair complementarity, including a few that give rise to heterogeneous, DCL4-dependent miRNAs[20]. These small RNAs might represent evolutionary intermediates that have yet to conform to the canonical miRNA-biogenesis pathway[20]. Interestingly, evolutionary intermediates of miRNAs might also be represented in the unicellular alga *Chlamydomonas reinhardtii*. Several miRNAs in *C. reinhardtii* arise from long, nearly perfectly paired hairpins that spawn heterogeneous small RNA populations[156,157].

Recent analysis of miRNAs in *A. thaliana* suggests that new foldback-yielding loci follow several different fates (FIG. 7). In the first fate, a small RNA with the capacity to regulate one or more members of the originating gene family is maintained. In the second fate, targeting specificity for an unrelated gene or gene family is acquired, most likely through genetic drift. Obviously, these first two fates require the presence of a selective advantage. The third, and perhaps most common, fate involves elimination of the small-RNA-generating locus through mutations in the promoter, foldback sequence and/or targeting elements. Thus, mechanisms for small RNA production in plants provide opportunities for the evolution of regulators with new specificities. As small RNA populations from different species throughout the plant kingdom are characterized, it will be interesting to learn how this putative evolutionary mechanism has been deployed in regulatory diversification among plant lineages.

Genetic screens and sequencing of small RNA populations have revealed the existence of hundreds of distinct miRNAs in primates and other animals[158–161]. Although some primate miRNAs arise from processed pseudogenes, and thus might have a similar origin to that proposed in the plant gene-duplication model above, this is probably uncommon[162]. A growing body of evidence suggests that several animal miRNAs originate from genomic repeats (FIG. 5). Between 5% and 20% of human miRNA precursor sequences, including those for several highly conserved miRNAs, contain sequences derived from repeat elements and transposable elements[163,164]. Several miRNA precursors in this group contain *LINE2*-like sequences, and hairpins within these miRNA transcripts arise from two adjacent, inverted *LINE2* elements[163,165]. Other miRNAs in this group arise from the MADE1 family of miniature inverted-repeat transposable elements (MITEs)[166]. Finally, miRNAs in a cluster on human chromosome 19 (C19MC) are interspersed among *Alu* repeats and rely on promoters in these *Alu* elements for transcription[167]. These findings suggest a functional relationship between transposable elements and animal miRNAs, and point to the possibility that repeat elements contribute in part to the genesis of miRNA genes.

Genetic drift
Fluctuations in gene frequencies in a population due to chance.

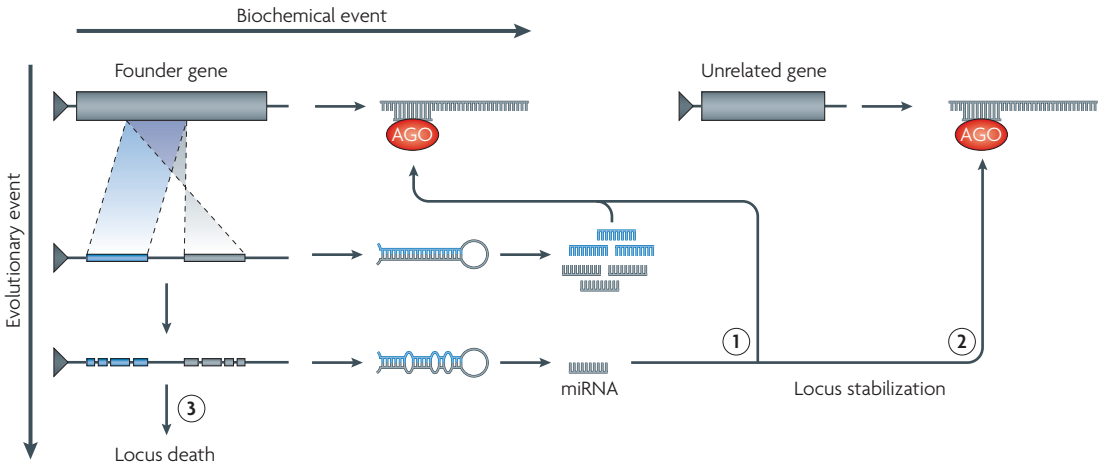


Figure 7 | **The three-fate model for the evolution of new microRNA (miRNA) loci in plants.** In the first fate, a new miRNA is selected with the capacity to regulate the progenitor gene or a related member of the progenitor family. In the second fate, a miRNA is selected with targeting specificity for an unrelated gene sequence. The first two fates are relatively rare, and are proposed to require selection for stabilization. The third fate involves loss of the locus by sequence drift or deletion in the absense of positive selection. AGO, Argonaute protein.

If such relationships exist between transposable elements and the origin of miRNA genes, how does specificity for their targets arise? An intriguing possibility is that these miRNAs target repeat elements that are integrated into transcribed genes (FIG. 6). Bioinformatic analyses show that numerous human miRNAs have complementarity to a conserved *Alu* element found in the 3′ UTRs of human mRNAs[163,164,168]. Similarly, *LINE2*-derived miRNAs might target a large number of mRNAs that carry *LINE2* elements in 3′ UTRs[163]. Together, these findings suggest a model for the origin of some animal miRNA regulatory circuits in which transposable elements gives rise to small RNAs that regulate related elements that are integrated into transcriptional units at other loci.

How has the deployment of miRNA-based regulation affected evolution of the transcriptome in eukaryotes? In animals, in which 'seeds' of miRNA targets are located predominantly in their 3′ UTRs and occupy only 7–8 nucleotides, the impact is expected to be particularly strong, as potential target sites may accumulate by chance[169]. Computational analysis of *D. melanogaster* and human genes indicates that genes that are coexpressed with miRNAs avoid evolving target seeds for those miRNAs[169,170]. Depletion of target seeds is also observed in orthologues from related species that are not miRNA targets, suggesting that these genes are under pressure to avoid target sites. This has been proposed to explain why several ubiquitously expressed 'housekeeping' genes have notably short 3′ UTRs that are perhaps less likely to accumulate target sites.

Evolution of miRNA families. miRNAs are grouped into families on the basis of sequence similarity of the mature miRNA. It is clear that animal miRNA families, such as the human mir17 cluster and the miRNAs in the Hox gene cluster, have expanded by the duplication of individual genes and gene clusters[165,171,172]. Duplications also yielded plant miRNA families, such as the seven-member miR166 family in *A. thaliana*[173].

What are the consequences of the expansion of miRNA families? miRNAs in a family might be redundant, or might function additively to exact a stronger regulatory effect on target mRNAs. miRNAs in a family might have distinct targeting specificities due to sequence differences. Also, family members might have distinct expression patterns to regulate specific targets across a spatiotemporal range. Experimental evidence supports each of these ideas and suggests that they are not mutually exclusive.

The related *D. melanogaster* miRNAs miR-310, miR-311, miR-312 and miR-313 have overlapping expression patterns, and antisense depletion of each miRNA triggers a similar phenotype, suggesting that they have redundant functions[174]. Expression domains of the three *A. thaliana MIR164* precursors partially overlap in developing shoots, and these three genes function redundantly to regulate phyllotaxis[175]. The expression domains of *A. thaliana MIR159A*, *MIR159B* and *MIR159C* overlap significantly, and plants that ectopically express members of the miR159 family have similar phenotypes[176], implying that there is some level of redundancy.

In at least some families, however, miRNAs have specific roles. Although *A. thaliana MIR164*-family members are partially redundant, the expression pattern of *MIR164C* is distinct from those of the other *MIR164* members. *MIR164C* functions as the predominant source of miR164 to negatively regulate *CUP-SHAPED COTYLEDON 2* (*CUC2*) transcripts in flowers[175]. This indicates that target specificity is conferred in part through specific expression patterns among miR164 family members. Similarly, expression patterns and the limited abundance of miR319 both prevent regulation of predicted targets in the *MYB* gene family, although *MYB* family members are effectively regulated by the similar miR159. Conversely, target sites in *TCP*-gene

transcripts are specifically recognized by miR319, but not miR159, owing to sequence differences[177]. From these data, a picture is emerging in which duplication of functional miRNA genes, followed by specialization through fine-tuning of base-pair complementarity and expression patterns, generates a suite of related miRNAs that confer regulatory specificity among sets of targets.

Conclusion

The evolution of RNA silencing as a regulatory mechanism in eukaryotes has involved proliferation and specialization of all of the core components of RNA silencing pathways. Gene duplications and DNA rearrangements have generated new small RNA regulators. Dicer proteins and processing cofactors, as well as RdRPs in some lineages, have specialized to produce unique small RNA classes for various downstream processes. Argonaute proteins have proliferated and evolved a range of functions for amplification and biogenesis of secondary small RNAs, enzymatic cleavage of RNA targets, translational repression and recruitment of chromatin-modification factors. Numerous specialized miRNA genes have emerged and evolved as post-transcriptional regulators of gene expression. The products of these specialization events now constitute an array of diverse small-RNA-based regulatory pathways that control gene expression.

Importantly, we have only begun to explore the function of small RNAs in controlling transcriptional networks. Overcoming challenges in identifying functional roles of small-RNA-based regulation in specific tissues and at the whole-organism level will be crucial to a more integrative view of how RNA silencing functions in a cellular and developmental context.

1. Levine, M. & Tjian, R. Transcription regulation and animal diversity. *Nature* **424**, 147–151 (2003).
2. Moore, M. J. From birth to death: the complex lives of eukaryotic mRNAs. *Science* **309**, 1514–1518 (2005).
3. Jorgensen, R. in *RNA Interference* (ed. Hannon, G.) 5–21 (Cold Spring Harbor Press, Cold Spring Harbor, 2003).
4. Mello, C. C. & Conte, D. Jr. Revealing the world of RNA interference. *Nature* **431**, 338–342 (2004).
5. Baulcombe, D. RNA silencing in plants. *Nature* **431**, 356–363 (2004).
6. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
This seminal paper was the first to define dsRNA as the trigger for RNAi.
7. Hamilton, A. J. & Baulcombe, D. C. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286**, 950–952 (1999).
The first report to link small RNAs to post-transcriptional gene silencing in plants.
8. Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**, 293–296 (2000).
9. 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).
10. 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).
11. Kim, V. N. MicroRNA biogenesis: coordinated cropping and dicing. *Nature Rev. Mol. Cell Biol.* **6**, 376–385 (2005).
12. Vazquez, F. *Arabidopsis* endogenous small RNAs: highways and byways. *Trends Plant Sci.* **11**, 460–468 (2006).
13. Kim, V. N. & Nam, J. W. Genomics of microRNA. *Trends Genet.* **22**, 165–173 (2006).
14. Du, T. & Zamore, P. D. microPrimer: the biogenesis and function of microRNA. *Development* **132**, 4645–4652 (2005).
15. Bartel, D. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
16. Slotkin, R. K. & Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nature Rev. Genet.* **8**, 272–285 (2007).
This paper provides a thorough review of the mechanisms and consequences of small-RNA-based regulation of transposable elements.
17. Lu, C. *et al.* MicroRNAs and other small RNAs enriched in the *Arabidopsis* RNA-dependent RNA polymerase-2 mutant. *Genome Res.* **16**, 1276–1288 (2006).
18. Kasschau, K. D. *et al.* Genome-wide profiling and analysis of *Arabidopsis* siRNAs. *PLoS Biol.* **5**, e57 (2007).
19. Xie, Z. *et al.* Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.* **2**, e104 (2004).
20. Rajagopalan, R., Vaucheret, H., Trejo, J. & Bartel, D. P. A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev.* **20**, 3407–3425 (2006).
See reference 153.
21. Zilberman, D., Cao, X. & Jacobsen, S. E. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* **299**, 716–719 (2003).
22. Gunawardane, L. S. *et al.* A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. *Science* **315**, 1587–1590 (2007).
23. Brennecke, J. *et al.* Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128**, 1089–1103 (2007).
With reference 22, this paper proposes a novel biogenesis mechanism for the production of piRNAs.
24. Houwing, S. *et al.* A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in zebrafish. *Cell* **129**, 69–82 (2007).
25. Martienssen, R. A. Maintenance of heterochromatin by RNA interference of tandem repeats. *Nature Genet.* **35**, 213–214 (2003).
26. Yang, N. & Kazazian, H. H. Jr. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nature Struct. Mol. Biol.* **13**, 763–771 (2006).
27. Borsani, O., Zhu, J., Verslues, P. E., Sunkar, R. & Zhu, J. K. Endogenous siRNAs derived from a pair of natural *cis*-antisense transcripts regulate salt tolerance in *Arabidopsis*. *Cell* **123**, 1279–1291 (2005).
28. Katiyar-Agarwal, S. *et al.* A pathogen-inducible endogenous siRNA in plant immunity. *Proc. Natl Acad. Sci. USA* **103**, 18002–18007 (2006).
29. Henz, S. R. *et al.* Distinct expression patterns of natural antisense transcripts in *Arabidopsis*. *Plant Physiol.* **144**, 1247–1255 (2007).
30. Williams, L., Carles, C. C., Osmont, K. S. & Fletcher, J. C. A database analysis method identifies an endogenous *trans*-acting short-interfering RNA that targets the *Arabidopsis ARF2*, *ARF3*, and *ARF4* genes. *Proc. Natl Acad. Sci. USA* **102**, 9703–9708 (2005).
31. Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H. L. & Poethig, R. S. SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of *trans*-acting siRNAs in *Arabidopsis*. *Genes Dev.* **18**, 2368–2379 (2004).
32. Vazquez, F. *et al.* Endogenous *trans*-acting siRNAs regulate the accumulation of *Arabidopsis* mRNAs. *Mol. Cell* **16**, 69–79 (2004).
33. Allen, E., Xie, Z., Gustafson, A. M. & Carrington, J. C. microRNA-directed phasing during *trans*-acting siRNA biogenesis in plants. Cell **121**, 207–221 (2005).
34. Howell, M. D. *et al.* Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in *Arabidopsis* reveals dependency on miRNA- and tasiRNA-directed targeting. *Plant Cell* **19**, 926–942 (2007).
35. Girard, A., Sachidanandam, R., Hannon, G. J. & Carmell, M. A. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442**, 199–202 (2006).
36. Aravin, A. *et al.* A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442**, 203–207 (2006).
37. Aravin, A. A., Sachidanandam, R., Girard, A., Fejes-Toth, K. & Hannon, G. J. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* **316**, 744–747 (2007).
38. Kalmykova, A. I., Klenov, M. S. & Gvozdev, V. A. Argonaute protein PIWI controls mobilization of retrotransposons in the *Drosophila* male germline. *Nucleic Acids Res.* **33**, 2052–2059 (2005).
39. Saito, K. *et al.* Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev.* **20**, 2214–2222 (2006).
40. Vagin, V. V. *et al.* A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313**, 320–324 (2006).
41. Klenov, M. S. *et al.* Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the *Drosophila melanogaster* germline. *Nucleic Acids Res.* **35**, 5430–5438 (2007).
42. Sijen, T., Steiner, F. A., Thijssen, K. L. & Plasterk, R. H. Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. *Science* **315**, 244–247 (2007).
See reference 77.
43. Dunoyer, P., Himber, C., Ruiz-Ferrer, V., Alioua, A. & Voinnet, O. Intra- and intercellular RNA interference in *Arabidopsis thaliana* requires components of the microRNA and heterochromatic silencing pathways. *Nature Genet.* **39**, 848–856 (2007).
44. Axtell, M. J., Jan, C., Rajagopalan, R. & Bartel, D. P. A two-hit trigger for siRNA biogenesis in plants. *Cell* **127**, 565–577 (2006).
This paper presents the common features between *trans*-acting siRNA loci in plants, providing a model for the guidance of transcripts into tasiRNA pathways.
45. Voinnet, O. Non-cell autonomous RNA silencing. *FEBS Lett.* **579**, 5858–5871 (2005).
46. Duchaine, T. F. *et al.* Functional proteomics reveals the biochemical niche of *C. elegans* DCR-1 in multiple small-RNA-mediated pathways. *Cell* **124**, 343–354 (2006).
47. 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).
48. Ambros, V., Lee, R. C., Lavanway, A., Williams, P. T. & Jewell, D. MicroRNAs and other tiny endogenous RNAs in *C. elegans*. *Curr. Biol.* **13**, 807–818 (2003).
49. Lee, R. C., Hammell, C. M. & Ambros, V. Interacting endogenous and exogenous RNAi pathways in *Caenorhabditis elegans*. *RNA* **12**, 589–597 (2006).
50. Simmer, F. *et al.* Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi. *Curr. Biol.* **12**, 1317–1319 (2002).

51. Kennedy, S., Wang, D. & Ruvkun, G. A conserved siRNA-degrading RNase negatively regulates RNA interference in *C. elegans*. *Nature* **427**, 645–649 (2004).
52. Gasciolli, V., Mallory, A. C., Bartel, D. P. & Vaucheret, H. Partially redundant functions of *Arabidopsis* DICER-like enzymes and a role for DCL4 in producing *trans*-acting siRNAs. *Curr. Biol.* **15**, 1494–1500 (2005).
53. Tomari, Y., Matranga, C., Haley, B., Martinez, N. & Zamore, P. D. A protein sensor for siRNA asymmetry. *Science* **306**, 1377–1380 (2004).
54. Liu, Q. *et al.* R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway. *Science* **301**, 1921–1925 (2003).
55. Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**, 231–235 (2004).
56. Kadotani, N., Nakayashiki, H., Tosa, Y. & Mayama, S. One of the two Dicer-like proteins in the filamentous fungi *Magnaporthe oryzae* genome is responsible for hairpin RNA-triggered RNA silencing and related small interfering RNA accumulation. *J. Biol. Chem.* **279**, 44467–44474 (2004).
57. Catalanotto, C. *et al.* Redundancy of the two dicer genes in transgene-induced posttranscriptional gene silencing in *Neurospora crassa*. *Mol. Cell. Biol.* **24**, 2536–2545 (2004).
58. Mochizuki, K. & Gorovsky, M. A. A Dicer-like protein in *Tetrahymena* has distinct functions in genome rearrangement, chromosome segregation, and meiotic prophase. *Genes Dev.* **19**, 77–89 (2005).
59. Lee, S. R. & Collins, K. Two classes of endogenous small RNAs in *Tetrahymena thermophila*. *Genes Dev.* **20**, 28–33 (2006).
60. Margis, R. *et al.* The evolution and diversification of Dicers in plants. *FEBS Lett.* **580**, 2442–2450 (2006).
61. Xie, Z., Allen, E., Wilken, A. & Carrington, J. C. DICER-LIKE 4 functions in *trans*-acting small interfering RNA biogenesis and vegetative phase change in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **102**, 12984–12989 (2005).
62. Vazquez, F., Gasciolli, V., Crete, P. & Vaucheret, H. The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. *Curr. Biol.* **14**, 346–351 (2004).
63. Hiraguri, A. *et al.* Specific interactions between Dicer-like proteins and HYL1/DRB-family dsRNA-binding proteins in *Arabidopsis thaliana*. *Plant Mol. Biol.* **57**, 173–188 (2005).
64. Han, M. H., Goud, S., Song, L. & Fedoroff, N. The *Arabidopsis* double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proc. Natl Acad. Sci. USA* **101**, 1093–1098 (2004).
65. Adenot, X. *et al.* DRB4-dependent TAS3 *trans*-acting siRNAs control leaf morphology through AGO7. *Curr. Biol.* **16**, 927–932 (2006).
66. Deleris, A. *et al.* Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science* **313**, 68–71 (2006).
67. Dunoyer, P., Himber, C. & Voinnet, O. DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. *Nature Genet.* **37**, 1356–1360 (2005).
68. Moissiard, G., Parizotto, E. A., Himber, C. & Voinnet, O. Transitivity in *Arabidopsis* can be primed, requires the redundant action of the antiviral Dicer-like 4 and Dicer-like 2, and is compromised by viral-encoded suppressor proteins. *RNA* **13**, 1268–1278 (2007).
69. Pantaleo, V., Szittya, G. & Burgyan, J. Molecular bases of viral RNA Targeting by viral small interfering RNA-programmed RISC. *J. Virol.* **81**, 3797–3806 (2007).
70. Himber, C., Dunoyer, P., Moissiard, G., Ritzenthaler, C. & Voinnet, O. Transitivity-dependent and -independent cell-to-cell movement of RNA silencing. *EMBO J.* **22**, 4523–4533 (2003).
71. Li, F. & Ding, S. W. Virus counterdefense: diverse strategies for evading the RNA-silencing immunity. *Annu. Rev. Microbiol.* **60**, 503–531 (2006).
72. Thomas, C. L., Leh, V., Lederer, C. & Maule, A. J. Turnip crinkle virus coat protein mediates suppression of RNA silencing in *Nicotiana benthamiana*. *Virology* **306**, 33–41 (2003).
73. Bouche, N., Lauressergues, D., Gasciolli, V. & Vaucheret, H. An antagonistic function for *Arabidopsis* DCL2 in development and a new function for DCL4 in generating viral siRNAs. *EMBO J.* **25**, 3347–3356 (2006).
74. Qu, F., Ren, T. & Morris, T. J. The coat protein of *Turnip crinkle virus* suppresses posttranscriptional gene silencing at an early step. *J. Virol.* **77**, 511–522 (2003).
75. Moissiard, G. & Voinnet, O. RNA silencing of host transcripts by *cauliflower mosaic virus* requires coordinated action of the four *Arabidopsis* Dicer-like proteins. *Proc. Natl Acad. Sci. USA* **103**, 19593–19598 (2006).
76. Grishok, A., Sinskey, J. L. & Sharp, P. A. Transcriptional silencing of a transgene by RNAi in the soma of *C. elegans*. *Genes Dev.* **19**, 683–696 (2005).
77. Pak, J. & Fire, A. Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science* **315**, 241–244 (2007).
With reference 42, this paper proposes a new model for RdRP-dependent biogenesis of secondary siRNAs for amplification of silencing.
78. Vaistij, F. E., Jones, L. & Baulcombe, D. C. Spreading of RNA targeting and DNA methylation in RNA silencing requires transcription of the target gene and a putative RNA-dependent RNA polymerase. *Plant Cell* **14**, 857–867 (2002).
79. Dalmay, T., Hamilton, A., Rudd, S., Angell, S. & Baulcombe, D. C. An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* **101**, 543–553 (2000).
80. Chen, H. M., Li, Y. H. & Wu, S. H. Bioinformatic prediction and experimental validation of a microRNA-directed tandem *trans*-acting siRNA cascade in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **104**, 3318–3323 (2007).
81. Yoshikawa, M., Peragine, A., Park, M. Y. & Poethig, R. S. A pathway for the biogenesis of *trans*-acting siRNAs in *Arabidopsis*. *Genes Dev.* **19**, 2164–2175 (2005).
82. Fahlgren, N. *et al.* Regulation of *AUXIN RESPONSE FACTOR3* by *TAS3* ta-siRNA affects developmental timing and patterning in *Arabidopsis*. *Curr. Biol.* **16**, 939–944 (2006).
83. Hunter, C. *et al.* *Trans*-acting siRNA-mediated repression of *ETTIN* and *ARF4* regulates heteroblasty in *Arabidopsis*. *Development* **133**, 2973–2981 (2006).
84. Chan, S. W. *et al.* RNA silencing genes control *de novo* DNA methylation. *Science* **303**, 1336 (2004).
85. Woodhouse, M. R., Freeling, M. & Lisch, D. Initiation, establishment, and maintenance of heritable MuDR transposon silencing in maize are mediated by distinct factors. *PLoS Biol.* **4**, e339 (2006).
86. Alleman, M. *et al.* An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature* **442**, 295–298 (2006).
87. Lisch, D., Carey, C. C., Dorweiler, J. E. & Chandler, V. L. A mutation that prevents paramutation in maize also reverses *Mutator* transposon methylation and silencing. *Proc. Natl Acad. Sci. USA* **99**, 6130–6135 (2002).
88. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419 (2003).
89. Han, J. *et al.* The Drosha–DGCR8 complex in primary microRNA processing. *Genes Dev.* **18**, 3016–3027 (2004).
90. Chendrimada, T. P. *et al.* TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**, 740–744 (2005).
91. Haase, A. D. *et al.* TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep.* **6**, 961–967 (2005).
92. Saito, K., Ishizuka, A., Siomi, H. & Siomi, M. C. Processing of pre-microRNAs by the Dicer-1–Loquacious complex in *Drosophila* cells. *PLoS Biol.* **3**, e235 (2005).
93. Forstemann, K. *et al.* Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biol.* **3**, e236 (2005).
94. 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).
95. Zeng, Y., Yi, R. & Cullen, B. R. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *EMBO J.* **24**, 138–148 (2005).
96. Han, J. *et al.* Molecular basis for the recognition of primary microRNAs by the Drosha–DGCR8 complex. *Cell* **125**, 887–901 (2006).
97. Ruby, J. G., Jan, C. H. & Bartel, D. P. Intronic microRNA precursors that bypass Drosha processing. *Nature* **448**, 83–86 (2007).
98. Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M. & Lai, E. C. The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* **130**, 89–100 (2007).
99. Papp, I. *et al.* Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors. *Plant Physiol.* **132**, 1382–1390 (2003).
100. 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).
101. Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B. & Bartel, D. P. MicroRNAs in plants. *Genes Dev.* **16**, 1616–1626 (2002).
102. Kurihara, Y. & Watanabe, Y. *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. *Proc. Natl Acad. Sci. USA* **101**, 12753–12758 (2004).
103. Kurihara, Y., Takashi, Y. & Watanabe, Y. The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. *RNA* **12**, 206–212 (2006).
104. Song, L., Han, M. H., Lesicka, J. & Fedoroff, N. *Arabidopsis* primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. *Proc. Natl Acad. Sci. USA* **104**, 5437–5442 (2007).
105. 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).
106. Nakazawa, Y., Hiraguri, A., Moriyama, H. & Fukuhara, T. The dsRNA-binding protein DRB4 interacts with the Dicer-like protein DCL4 *in vivo* and functions in the *trans*-acting siRNA pathway. *Plant Mol. Biol.* **63**, 777–785 (2007).
107. Henderson, I. R. *et al.* Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nature Genet.* **38**, 721–725 (2006).
108. Meister, G. *et al.* Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* **15**, 185–197 (2004).
109. Liu, J. *et al.* Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437–1441 (2004).
110. Rivas, F. V. *et al.* Purified Argonaute2 and an siRNA form recombinant human RISC. *Nature Struct. Mol. Biol.* **12**, 340–349 (2005).
111. Baumberger, N. & Baulcombe, D. C. *Arabidopsis* ARGONAUTE1 is an RNA slicer that selectively recruits microRNAs and short interfering RNAs. *Proc. Natl Acad. Sci. USA* **102**, 11928–11933 (2005).
112. Tolia, N. H. & Joshua-Tor, L. Slicer and the Argonautes. *Nature Chem. Biol.* **3**, 36–43 (2007).
113. Qi, Y. *et al.* Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature* **443**, 1008–1012 (2006).
114. Qi, Y., Denli, A. M. & Hannon, G. J. Biochemical specialization within *Arabidopsis* RNA silencing pathways. *Mol. Cell* **19**, 421–428 (2005).
115. Zilberman, D. *et al.* Role of *Arabidopsis* ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats. *Curr. Biol.* **14**, 1214–1220 (2004).
116. Vaucheret, H., Vazquez, F., Crete, P. & Bartel, D. P. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev.* **18**, 1187–1197 (2004).
117. Morel, J. B. *et al.* Fertile hypomorphic ARGONAUTE (*ago1*) mutants impaired in post-transcriptional gene silencing and virus resistance. *Plant Cell* **14**, 629–639 (2002).
118. Zhang, X. *et al.* *Cucumber mosaic virus*-encoded 2b suppressor inhibits *Arabidopsis* ARGONAUTE1 cleavage activity to counter plant defense. *Genes Dev.* **20**, 3255–3268 (2006).
119. Tomari, Y., Du, T. & Zamore, P. D. Sorting of *Drosophila* small silencing RNAs. *Cell* **130**, 299–308 (2007).
120. Forstemann, K., Horwich, M. D., Wee, L., Tomari, Y. & Zamore, P. D. *Drosophila* microRNAs are sorted into functionally distinct argonaute complexes after production by Dicer-1. *Cell* **130**, 287–297 (2007).
121. Yigit, E. *et al.* Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell* **127**, 747–757 (2006).

122. 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).
123. 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).
124. Sijen, T. *et al.* On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107**, 1–20 (2001).
125. Steiner, F. A. & Plasterk, R. H. Knocking out the Argonautes. *Cell* **127**, 667–668 (2006).
126. Lozovskaya, E. R., Hartl, D. L. & Petrov, D. A. Genomic regulation of transposable elements in *Drosophila*. *Curr. Opin. Genet. Dev.* **5**, 768–773 (1995).
127. Okamura, K., Ishizuka, A., Siomi, H. & Siomi, M. C. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev.* **18**, 1655–1666 (2004).
128. Rehwinkel, J. *et al.* Genome-wide analysis of mRNAs regulated by Drosha and Argonaute proteins in *Drosophila melanogaster*. *Mol. Cell. Biol.* **26**, 2965–2975 (2006).
129. Valencia-Sanchez, M. A., Liu, J., Hannon, G. J. & Parker, R. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev.* **20**, 515–524 (2006).
130. Yao, M. C. & Chao, J. L. RNA-guided DNA deletion in *Tetrahymena*: an RNAi-based mechanism for programmed genome rearrangements. *Annu. Rev. Genet.* **39**, 537–559 (2005).
131. Mochizuki, K., Fine, N. A., Fujisawa, T. & Gorovsky, M. A. Analysis of a *piwi*-related gene implicates small RNAs in genome rearrangement in *Tetrahymena*. *Cell* **110**, 689–699 (2002).
132. Liu, Y. *et al.* RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in *Tetrahymena*. *Genes Dev.* **21**, 1530–1545 (2007).
133. Hamilton, A., Voinnet, O., Chappell, L. & Baulcombe, D. Two classes of short interfering RNA in RNA silencing. *EMBO J.* **21**, 4671–4679 (2002).
134. Slotkin, R. K., Freeling, M. & Lisch, D. Mu killer causes the heritable inactivation of the Mutator family of transposable elements in *Zea mays*. *Genetics* **165**, 781–797 (2003).
135. Slotkin, R. K., Freeling, M. & Lisch, D. Heritable transposon silencing initiated by a naturally occurring transposon inverted duplication. *Nature Genet.* **37**, 641–644 (2005).
136. Martienssen, R. & Baron, A. Coordinate suppression of mutations caused by Robertson's mutator transposons in maize. *Genetics* **136**, 1157–1170 (1994).
137. Walker, E. L. Paramutation of the *r1* locus of maize is associated with increased cytosine methylation. *Genetics* **148**, 1973–1981 (1998).
138. Lippman, Z. & Martienssen, R. The role of RNA interference in heterochromatic silencing. *Nature* **431**, 364–370 (2004).
139. Carmell, M. A. *et al.* MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev. Cell* **12**, 503–514 (2007).
140. Grewal, S. I. & Jia, S. Heterochromatin revisited. *Nature Rev. Genet.* **8**, 35–46 (2007).
141. Heit, R., Underhill, D. A., Chan, G. & Hendzel, M. J. Epigenetic regulation of centromere formation and kinetochore function. *Biochem. Cell Biol.* **84**, 605–618 (2006).
142. Onodera, Y. *et al.* Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**, 613–622 (2005).
143. Herr, A. J., Jensen, M. B., Dalmay, T. & Baulcombe, D. C. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**, 118–120 (2005).
144. Pontier, D. *et al.* Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. *Genes Dev.* **19**, 2030–2040 (2005).
145. Pontes, O. *et al.* The *Arabidopsis* chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. *Cell* **126**, 79–92 (2006).
146. Li, C. F. *et al.* An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. *Cell* **126**, 93–106 (2006).
147. Chan, S. W., Henderson, I. R. & Jacobsen, S. E. Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nature Rev. Genet.* **6**, 351–360 (2005).
148. Verdel, A. *et al.* RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**, 672–676 (2004).
149. Sugiyama, T., Cam, H., Verdel, A., Moazed, D. & Grewal, S. I. RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proc. Natl Acad. Sci. USA* **102**, 152–157 (2005).
150. Motamedi, M. R. *et al.* Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119**, 789–802 (2004).
151. Cam, H. P. *et al.* Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nature Genet.* **37**, 809–819 (2005).
152. Noma, K. *et al.* RITS acts *in cis* to promote RNA interference-mediated transcriptional and post-transcriptional silencing. *Nature Genet.* **36**, 1174–1180 (2004).
153. Allen, E. *et al.* Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. *Nature Genet.* **36**, 1282–1290 (2004).
In this report, a model is proposed for *de novo* evolution of miRNA regulators in plants. This model is further explored in reference 20.
154. Wang, Y., Hindemitt, T. & Mayer, K. F. Significant sequence similarities in promoters and precursors of *Arabidopsis thaliana* non-conserved microRNAs. *Bioinformatics* **22**, 2585–2589 (2006).
155. Fahlgren, N. *et al.* High-throughput sequencing of *Arabidopsis* microRNAs: evidence for frequent birth and death of miRNA genes. *PLoS ONE* **2**, e219 (2007).
156. Molnar, A., Schwach, F., Studholme, D. J., Thuenemann, E. C. & Baulcombe, D. C. miRNAs control gene expression in the single-cell alga *Chlamydomonas reinhardtii*. *Nature* **447**, 1126–1129 (2007).
157. Zhao, T. *et al.* A complex system of small RNAs in the unicellular green alga *Chlamydomonas reinhardtii*. *Genes Dev.* **21**, 1190–1203 (2007).
References 156 and 157 describe endogenous miRNAs in a unicellular organism.
158. 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).
159. 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).
160. Wienholds, E. & Plasterk, R. H. MicroRNA function in animal development. *FEBS Lett.* **579**, 5911–5922 (2005).
161. Kloosterman, W. P. & Plasterk, R. H. The diverse functions of microRNAs in animal development and disease. *Dev. Cell* **11**, 441–450 (2006).
162. Devor, E. J. Primate microRNAs *miR-220* and *miR-492* lie within processed pseudogenes. *J. Hered.* **97**, 186–190 (2006).
163. Smalheiser, N. R. & Torvik, V. I. Mammalian microRNAs derived from genomic repeats. *Trends Genet.* **21**, 322–326 (2005).
164. Piriyapongsa, J., Marino-Ramirez, L. & Jordan, I. K. Origin and evolution of human microRNAs from transposable elements. *Genetics* **176**, 1323–1337 (2007).
In this report, several human miRNA species are proposed to originate from transposable elements.
165. Hertel, J. *et al.* The expansion of the metazoan microRNA repertoire. *BMC Genomics* **7**, 25 (2006).
166. Piriyapongsa, J. & Jordan, I. K. A family of human microRNA genes from miniature inverted-repeat transposable elements. *PLoS ONE* **2**, e203 (2007).
167. Borchert, G. M., Lanier, W. & Davidson, B. L. RNA polymerase III transcribes human microRNAs. *Nature Struct. Mol. Biol.* **13**, 1097–1101 (2006).
168. Smalheiser, N. R. & Torvik, V. I. *Alu* elements within human mRNAs are probable microRNA targets. *Trends Genet.* **22**, 532–536 (2006).
169. Farh, K. K. *et al.* The widespread impact of mammalian microRNAs on mRNA repression and evolution. *Science* **310**, 1817–1821 (2005).
170. Stark, A., Brennecke, J., Bushati, N., Russell, R. B. & Cohen, S. M. Animal microRNAs confer robustness to gene expression and have a significant impact on 3′UTR evolution. Cell **123**, 1133–1146 (2005).
171. Tanzer, A., Amemiya, C. T., Kim, C. B. & Stadler, P. F. Evolution of microRNAs located within Hox gene clusters. *J. Exp. Zoolog. B Mol. Dev. Evol.* **304**, 75–85 (2005).
172. Tanzer, A. & Stadler, P. F. Molecular evolution of a microRNA cluster. *J. Mol. Biol.* **339**, 327–335 (2004).
173. Maher, C., Stein, L. & Ware, D. Evolution of *Arabidopsis* microRNA families through duplication events. *Genome Res.* **16**, 510–519 (2006).
174. Leaman, D. *et al.* Antisense-mediated depletion reveals essential and specific functions of microRNAs in *Drosophila* development. *Cell* **121**, 1097–1108 (2005).
175. Sieber, P., Wellmer, F., Gheyselinck, J., Riechmann, J. L. & Meyerowitz, E. M. Redundancy and specialization among plant microRNAs: role of the *MIR164* family in developmental robustness. *Development* **134**, 1051–1060 (2007).
176. Schwab, R. *et al.* Specific effects of microRNAs on the plant transcriptome. *Dev. Cell* **8**, 517–527 (2005).
177. Palatnik, J. F. *et al.*, Weigel, D. Sequence and expression differences underlie functional specialization of *Arabidopsis* microRNAs *miR159* and *miR319*. *Dev. Cell* **13**, 115–125 (2007).
178. Cerutti, H. & Casas-Mollano, J. A. On the origin and functions of RNA-mediated silencing: from protists to man. *Curr. Genet.* **50**, 81–99 (2006).

Acknowledgements
We thank members of the Carrington laboratory for productive discussions. E.J.C. was supported in part by a P.F. and Nellie Buck Yerex Fellowship. Research in the J.C.C. laboratory was supported by grants from the US National Institutes of Health, the US National Science Foundation and the US Department of Agriculture.

DATABASES

Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
AGO1 | AGO4 | AGO7 | AUB | CSR-1 | *CUC2* | DCL2 | DCL3 | DCL4 | DRB4 | *FWA* | HEN1 | HYL1 | RDR6 | RRF-1 | SAGO-1 | SAGO-2 | SE | TRBP

UniProtKB: http://ca.expasy.org/sprot
Ago1 | ALG-1 | ALG-2 | Chp1 | DCR-1 | Dicer | Drosha | Loquacious | MILI | MIWI2 | Pasha | PIWI | RDE-1 | RDE-4 | R2D2

FURTHER INFORMATION

The Carrington laboratory:
http://jcclab.science.oregonstate.edu/

ALL LINKS ARE ACTIVE IN THE ONLINE PDF