University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Faculty Publications from the Center for Plant Science Innovation

Plant Science Innovation, Center for

1-12-2003

RNA interference: traveling in the cell and gaining functions?

Heriberto D. Cerutti
University of Nebraska - Lincoln, hcerutti1@unl.edu

Follow this and additional works at: http://digitalcommons.unl.edu/plantscifacpub

Part of the Plant Sciences Commons

Cerutti, Heriberto D., "RNA interference: traveling in the cell and gaining functions?" (2003). *Faculty Publications from the Center for Plant Science Innovation.* Paper 21.
http://digitalcommons.unl.edu/plantscifacpub/21

This Article is brought to you for free and open access by the Plant Science Innovation, Center for at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications from the Center for Plant Science Innovation by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in *Trends in Genetics* **19**:1 (January 2003), pp. 39-46; doi:10.1016/S0168-9525(02)00010-0
Copyright © 2002 Elsevier Science Ltd. Used by permission. http://tigs.trends.com

Published online November 23, 2002.

RNA interference: traveling in the cell and gaining functions?

Heriberto Cerutti

School of Biological Sciences and Plant Science Initiative, University of Nebraska–Lincoln, E211 Beadle Center, PO Box 880666, Lincoln, NE 68588-0666, USA; hcerutti1@unl.edu

Abstract: Double-stranded RNA can induce the degradation of homologous RNAs in organisms as diverse as protozoa, animals, plants and fungi, resulting in post-transcriptional gene silencing. But in some species, RNA-mediated processes can also lead to translational repression, DNA methylation, heterochromatin formation or DNA elimination. In some situations, amplification of the "trigger" double-stranded RNA seems to be required for efficient silencing. These findings imply that RNA-mediated mechanisms can control gene expression at both the transcriptional and the post-transcriptional level, and that they can operate in the nuclear and the cytoplasmic compartments.

Double-stranded RNA (dsRNA) can induce many different epigenetic gene-silencing processes in eukaryotes, including the degradation of homologous mRNAs—a process that is known as RNA interference (RNAi) in animals and post-transcriptional gene silencing (PTGS) in plants. Components of the RNAi and PTGS machinery are also involved in the processing and function of microRNAs, a class of small RNAs that were originally identified by their role in translational repression in some animals.

Recent findings have also suggested that dsRNA that has been processed into small RNAs has a role in several chromatin and/or genomic DNA modifications. These observations indicate that dsRNA-dependent mechanisms can affect either post-transcriptional or transcriptional phenomena, raising issues about the subcellular location of the RNAi and PTGS machinery.

1. Basic mechanism of RNA interference

Our mechanistic understanding of RNAi is derived mainly from biochemical work in cell extracts from *Drosophila melanogaster* and *Homo sapiens*, complemented by genetic studies in *Arabidopsis thaliana*, *Caenorhabditis elegans* and *Neurospora crassa* [1–5]. The dsRNA-induced degradation of homologous RNAs can be divided simplistically into initiation and effector steps.

In the initiation step, a long dsRNA is processed into small interfering RNAs (siRNAs) of about 21–23 nt [6]. This cleavage requires ATP, and it is mediated by an RNase-III-like dsRNA-specific ribonuclease, named Dicer in *Drosophila* [6,7] (Figure 1, gray box). In the effector step, the double-stranded siRNAs are incorporated into a multiprotein complex, known as the RNA-induced silencing complex (RISC) [8] (Figure 1, blue box). RISC is then proposed to undergo an ATP-dependent activation step that results in the unwinding of the double-stranded siRNAs [9]. Activated RISC uses a single-stranded siRNA as a guide to identify complementary RNAs [9,10], and an endoribonuclease that has yet to be identified then cleaves the target RNA across from the center of the guide siRNA [8,10,11] (Figure 1, blue box). Finally, the cleaved RNA is probably degraded by exoribonucleases [8].

This basic RNAi pathway is probably present in many, if not all, eukaryotes that show RNAi phenomena. For example, siRNAs have been identified in plants, animals, fungi, protozoa and algae [1,2, 3,4,5], and the apparent products of the *in vivo* endonucleolytic cleavage of mRNAs can be detected in human cells [12] and tomato [13]. Homologs of some of the key components in the pathway, such as Dicer and the *Drosophila* RISC protein AGO2 [14], which is a member of the Argonaute gene family [15], have been implicated in RNA-mediated silencing in several eukaryotes [2,3,5,10,16,17]. Argonaute proteins are characterized by the presence of two structural domains: a PAZ (for "Piwi/Argonaute/Zwille") domain and a C-terminal Piwi domain [18]. But the molecular function of these proteins is unknown, and different homologs might have distinct roles in RNAi. In *C. elegans*, for example, RDE-1 (an Argonaute family member) and RDE-4 (a dsRNA-binding protein) are part of a complex with the Dicer ribonuclease and DHR-1 and DHR-2 (putative RNA helicases) [19]. Genetic analyses indicate that both RDE-1 and RDE-4 are required to initiate silencing, presumably because they aid the conversion of dsRNA into siRNAs [19,20]. But they are not required for the effector step, which is probably mediated by a *C. elegans* analog of RISC that might contain (in germline cells) the putative exonuclease MUT-7 [20,21]. Thus, Argonaute proteins might function at different steps in the production of siRNAs (or microRNAs) and/or in the formation and function of effector complexes.

An unresolved issue is the subcellular locations of siRNA production and target RNA degradation, although most evidence is consistent with a cytosolic, basic RNAi pathway (Figure 1, gray and blue boxes). In mouse oocytes and *Drosophila*, RNAi can target cytoplasmic maternal mRNAs [22,23]. RNA viruses, which replicate exclusively in the cytosol, are inhibited by RNA-mediated silencing mechanisms in both plant and mammalian cells [2,24,25]. In *Drosophila*, in addition to AGO2, two cytoplasmic Argonaute family members, AGO1 and Aubergine (Aub), have been implicated in RNAi

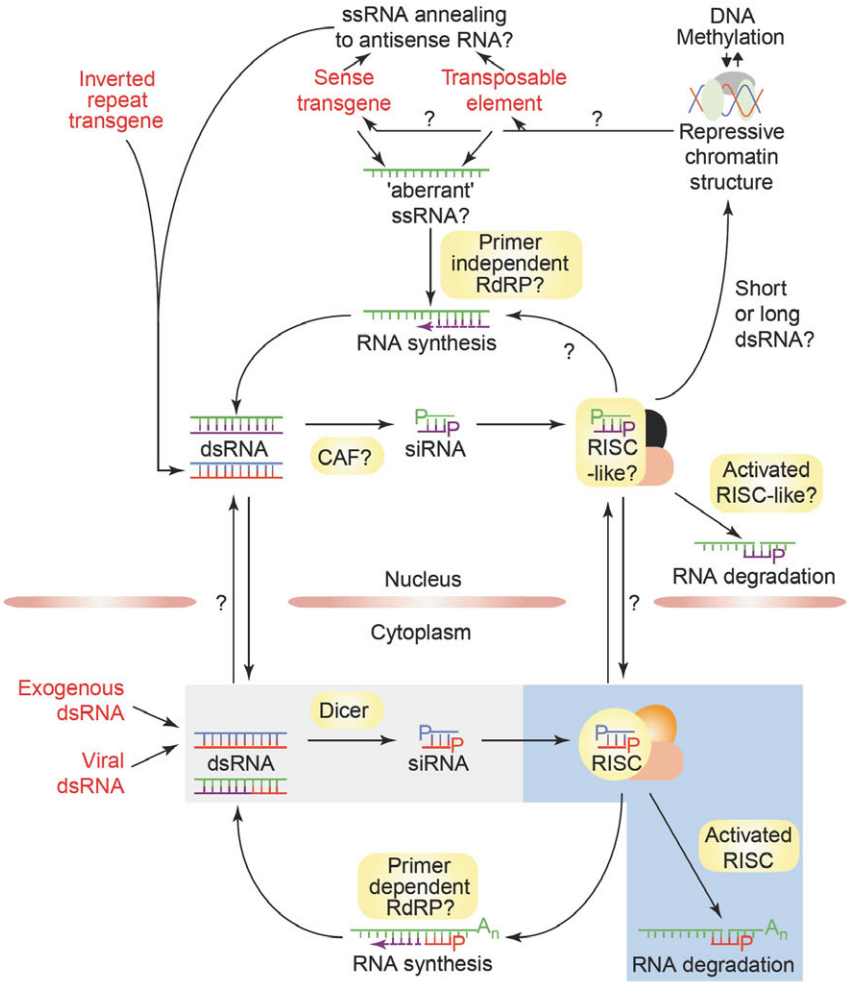

Figure 1. Models of molecular pathways involved in double-stranded RNA (dsRNA)-mediated silencing. The basic mechanism, which is probably present in most eukaryotes undergoing dsRNA-mediated silencing, is indicated by the gray and blue boxes. The remaining steps seem to occur in at least some organisms, but their generality is currently unknown and their subcellular location is mainly hypothetical. The "triggers" of silencing, either direct sources of dsRNA or transcription units producing single-stranded RNAs that can be presumably converted to dsRNA, are colored red. Green RNA, endogenously transcribed single-stranded RNA; purple RNA, RNA synthesized by a putative RNA-directed RNA polymerase; blue and red RNA, double-stranded RNA introduced exogenously or resulting from viral replication, annealing of complementary ssRNAs and/or hairpin RNA. Proteins or protein complexes are indicated by yellow boxes: CAF, an *Arabidopsis* homolog of Dicer; Dicer, an RNase-III-like dsRNA-specific ribonuclease; RdRP, an RNA-directed RNA polymerase; and RISC, RNA-induced silencing complex. Although dsRNA is depicted in single nuclear and cytoplasmic pools, depending on the source these molecules might be delivered differently to the processing Dicer enzymes. Similarly, the RISC and RISC-like complexes might have different components and associated effector proteins depending on their functions. Although a role for dsRNA in directing methylation of homologous DNA sequences has been demonstrated in plants, the molecular machinery involved in this process and the actual nature of the "guide" RNA have not been resolved. Recent evidence suggests that the RISC complex is equivalent to the miRNP complex (Figure 2) in human cells [17].

[26–28] (Table 1). In human cells, RNAi seems to occur predominantly in the cytosol [10,17], because a nucleoplasm confined transcript is completely resistant to degradation induced by a homologous siRNA [29]. In addition, the only known human homolog of Dicer is located in the cytoplasm of HeLa cells [30] (Table 1).

Notably, the rat homolog of eIF2C2 (GERp95), an Argonaute protein and a component of human RISC [10,17], is a peripheral membrane protein that is associated primarily with the Golgi complex or the endoplasmic reticulum [31].

Exogenous (injected or fed) dsRNA and viral dsRNA probably enter the cytoplasmic RNAi pathway directly (Figure 1). In several organisms, however, RNAi can also be induced by inverted repeat transgenes that are transcribed into hairpin dsRNA in the nucleus [1,2,16]. This dsRNA probably needs to be exported to the cytosol to target homologous mRNAs effectively (Figure 1). In both plants [32] and *Drosophila* [33], for example, inverted repeat transgenes silence more efficiently when the hairpin dsRNA contains an intron and polyadenylation signals that presumably facilitate entry of the dsRNA into the mRNA export pathway [34]. By contrast, a transgene designed to produce non-polyadenylated, intronless hairpin dsRNA—which is presumably retained in nuclei—targets the degradation of a homologous mRNA very poorly [1,35].

2. Potential amplification processes in RNAi and PTGS

Genetic analyses suggest that an amplification step might be required for efficient RNA-mediated silencing in several systems.

Table 1. Genetic and/or biochemical components of the RNAi/PTGS machinery discussed in the text,[a, b]

Gene function	*A. thaliana*	*C. elegans*	*C. reinhardtii*	*D. discoideum*	*D. melanogaster*	*H. sapiens*	*N. crassa*	*S. pombe*
RNase III + RNA helicase		CAF	DCR-1			Dicer	Dicer [C]	
DCR1								
Unknown (Argonaute)	AGO1	RDE-1, ALG-1, ALG-2			AGO1 [C], AGO2, AUB [C], PIWI [N]	eIF2C1, eIF2C2	QDE2	AGO1
Putative RdRP	SDE1/SGS2	RRF-1, RRF-3		RRPA			QDE1	RDP1 [N]
RNA helicase		DRH-1, DRH2	MUT6 [N]			Gemin3 [N + C]		
dsRNA-binding protein		RDE-4						
RNase D like		MUT-7						
RecQ DNA helicase							QDE3	
Unknown						Gemin4 [N + C]		

[a] This list is not comprehensive and only some proteins are indicated for each of the following organisms: *Arabidopsis thaliana*, *Caenorhabditis* elegans, *Chlamydomonas reinhardtii*, *Dictyostelium discoideum*, *Drosophila melanogaster*, *Homo sapiens*, *Neurospora crassa*, *Schizosaccharomyces pombe*.
[b] Where known, the subcellular localization of the proteins is shown in brackets: C = cytoplasmic; N = nuclear.

Recent studies in *C. elegans* have led to a model in which primary siRNAs (derived from the trigger dsRNA) might prime the synthesis of additional dsRNA, using the target mRNA as a template, in a reaction catalyzed by a putative RNA-directed RNA polymerase (RdRP) [4,36] (Figure 1, cytoplasm). The newly synthesized dsRNA would then be cleaved by Dicer to generate secondary siRNAs at a sufficient concentration to achieve efficient target mRNA degradation by RISC [4,36]. In support of this model, the injection of short antisense RNA oligomers into *C. elegans* can trigger silencing of endogenous genes, and this effect is dependent on a functional Dicer (DCR-1) [37].

Putative homologs of a tomato RdRP are required for PTGS triggered by sense transgenes in *Arabidopsis*, for quelling in *Neurospora*, and for RNAi in *C. elegans* and *Dictyostelium discoideum* [1–4,36,38–40]; however, biochemical activity remains to be demonstrated definitively for any of the RdRP homologs implicated in RNAi and PTGS [5]. In addition, in both *Drosophila* and humans, the results of many experiments argue against an obligatory role for an RdRP in dsRNA-induced RNAi [5,9,2], although an RdRP-like activity has been reported in extracts from *Drosophila* embryos [41].

Because siRNA production and target mRNA cleavage most probably occur in the cytosol, it is tempting to speculate that the RdRP-dependent amplification step described in *C. elegans* also occurs in the cytoplasm (Figure 1). Notably, RDE-1 and RDE-4 are required to initiate RNAi induced by injected or transgene-produced long dsRNAs [20], but not for silencing triggered by short antisense RNAs [37]. Thus, the RdRP-generated dsRNA might be delivered to Dicer differently to the way in which exogenous or transgenic dsRNA is delivered [19] (Figure 1, cytoplasm). In *Dictyostelium*, RRPA, a putative RdRP required for RNAi, contains an N-terminal RNA helicase domain that has not been found in other RdRPs [40]. Intriguingly, the closest homolog of this helicase domain is that of *C. elegans* Dicer, suggesting that domain swapping might have occurred between *Dictyostelium* Dicer and RRPA [40]. If this is true, then perhaps Dicer and an RdRP interact in a complex such that the dsRNA generated by RdRP activity is immediately accessible and rapidly processed by Dicer.

In *C. elegans*, RDE-1, RDE-4 and RRF-1 (a putative RdRP) are required for dsRNA-induced silencing in the soma, even when the trigger dsRNA is expressed from a transgene in the nuclei of target tissues [16,19,36]. SDE1/SGS2 (a putative RdRP) and AGO1 (the first described member of the Argonaute family [15]) are also required for PTGS induced by sense transgenes in *Arabidopsis* [38,39,42]. They are, however, completely dispensable for PTGS induced by the expression of inverted repeat transgenes that produce hairpin dsRNA [42]. This suggests that the putative amplification step catalyzed by *Arabidopsis* SDE1/SGS2 might be different to that mediated by *C. elegans* RRF-1 [42].

In plants, there is also an SDE1/SGS2-dependent spreading of silencing from the region homologous to the trigger dsRNA into the adjacent non-homologous 5′ and 3′ regions of a target transgene [43]. This is not consistent with the simple notion that primary siRNAs prime 5′→3′ dsRNA synthesis on sense mRNA. A possible explanation is that trigger dsRNAs or siRNAs interact with the transgene DNA and induce changes in chromatin structure [38,43] (Figure 1, nucleus, repressive chromatin structure). This would lead to the production of "aberrant", but nearly full-length transgenic RNAs that could be used by SDE1/SGS2 for primer-independent amplification [38,43] (Figure 1, nucleus, "aberrant ssRNA?").

In several plant species, PTGS of transgenes usually correlates with DNA methylation in transcribed regions [1–3,44]. Chromatin modifications associated with (or preceding) DNA methylation might lead to the production of truncated aberrant RNAs, as has been reported in filamentous fungi [45]. A role for DNA methylation and/or chromatin modification in PTGS of sense transgenes is also supported by the observation that mutations in MET1/DDM2 (a DNA methyltransferase) and in DDM1 (a member of the SWI2/SNF2 family of chromatin-remodeling proteins) can affect both the degree and persistence of post-transcriptional silencing in *Arabidopsis* [44], although indirect effects, such as saturation of the PTGS machinery by the unregulated expression of endogenous RNAs, cannot be ruled out.

The nature of the postulated aberrant RNA remains unknown [1,2,5,38], but PTGS can be induced in soybean and

tobacco by transgenes expressing ribozyme-truncated transcripts that are not polyadenylated and are preferentially retained in the nucleus [46]. In *Drosophila*, the post-transcriptional silencing of alcohol dehydrogenase transgenes—a phenomenon that is similar to PTGS induced by sense transgenes in plants—is dependent on the Argonaute family member Piwi [47]. Notably, Piwi is localized predominantly in the nucleus of interphase cells [48] (Table 1). In addition, in *Schizosaccharomyces pombe*, RDP1 (a putative RdRP homolog) is associated with centromeric DNA [49] (Table 1).

It is therefore tempting to speculate that in *Arabidopsis* the mechanism dependent on SDE1/SGS2 and AGO1 might function to produce dsRNA from non-polyadenylated, prematurely terminated or misprocessed RNAs in the nuclear compartment (Figure 1, nucleus, "primer independent RdRP?"). Perhaps the RdRP itself operates as a sensor for the accumulation of aberrant RNAs above a certain threshold [5]. The resulting dsRNA could be either processed to siRNAs by a nuclear Dicer (see below) or exported to the cytoplasm as part of a ribonucleoprotein complex that delivers the dsRNA to a cytosolic Dicer. But only the latter might be able to induce PTGS because, as discussed above, dsRNA from inverted repeat transgenes apparently needs to be exported for efficient silencing [32,33].

A similar model could explain co-suppression triggered by single-copy dispersed transgenes as well as certain transposon and repetitive DNA-silencing phenomena [3,4,13,19,38]. But it should be noted that some aberrant RNAs could exit the nucleus and the SDE1/SGS2-dependent step could occur in the cytosol. Furthermore, this RdRP might have dual roles in both the nuclear and the cytoplasmic compartments.

The apparent differences in the requirement or the characteristics of an amplification step, when comparing *Arabidopsis*, *C. elegans*, *Drosophila* and human, might reflect true mechanistic differences among organisms. Alternatively, at least in species in which RdRPs homologs are encoded by multigene families, such as *Arabidopsis*, *Dictyostelium*, *Neurospora* and *C. elegans* [3,4,36,38–40], different family members might participate in different steps of the pathway. Thus, the discrepancies might ultimately reflect our current lack of understanding of the role of several components. In addition, as mentioned above, the enzymatic activity of the putative RdRPs involved in RNAi and PTGS remains to be verified. This is particularly relevant because loss of function of RRF-3 (a possible RdRP) in *C. elegans* enhances the sensitivity to RNAi in several tissues [50]. Thus, RRF-3 behaves as a negative modulator of the RNAi response, which is difficult to reconcile with its postulated activity (but see [50]).

3. Nuclear phenomena associated with RNAi and PTGS

In *C. elegans*, RNAi mediated by injected long dsRNA virtually abolished the accumulation of a homologous cytoplasmic transcript, but it also produced a substantial reduction of nascent RNA in the nucleus [51]. RNAi is also triggered by dsRNA homologous to some introns, suggesting that at least certain precursor mRNAs (pre-mRNAs) can be targeted for degradation [52].

In the green alga *Chlamydomonas reinhardtii*, MUT6 (a DEAH-box RNA helicase required for PTGS [53]) is localized in the nucleus (K. van Dijk, B. Jeong and H. Cerutti, unpublished; Table 1). This helicase associates with a single-stranded RNA-binding protein and with a ribonuclease containing a staphylococcal nuclease-like domain. Because MUT6 is required for the degradation of non-polyadenylated transposon transcripts [53], which are presumably retained in nuclei, these observations suggest the existence of a nuclear RISC-like complex (Figure 1, nucleus). In organisms with a single, cytoplasmic Dicer, RISC or a similar ribonucleoprotein (RNP) complex might be imported into the nucleus. But evidence from *C. elegans* suggests that most pre-mRNAs do not seem to be targeted by RNAi [51,52], possibly because they are protected by RNP complexes and/or rapidly processed to mature mRNAs and exported to the cytosol [34]. Thus, a nuclear RISC-like complex might be capable of degrading only certain transcripts, such as non-polyadenylated transposon RNAs or accessible pre-mRNAs.

In human cells, RNAi seems to be limited mostly to the cytoplasm [10,17,29]. A nucleoplasm-sequestered transcript was found to be resistant to degradation induced by a homologous siRNA. But the same RNA, in the biochemically defined nuclear fraction, was partially degraded when undergoing export from the nucleus [29]. A possible explanation is that transcripts become accessible to the cytoplasmic RNAi machinery before complete transit of the mRNAs through the nuclear pores [29]. Alternatively, nuclear RISC-like complexes might exist in human cells but only target accessible transcripts, such as those that undergo remodeling of associated proteins concomitant with nuclear export [34]. As suggested previously [54], a fraction of the RISC complexes might be located at the nuclear pores and act as gatekeepers, scanning the RNAs as they are being exported.

The human RISC complex is proposed to contain eIF2C1 and/or eIF2C2, Gemin3 (a putative DEAD-box RNA helicase) and Gemin4 [10,17]. Gemin3 and Gemin4 are also components of another multiprotein complex, containing the SMN (survival of motor neurons) protein, that functions in the assembly and restructuring of diverse RNP particles involved in transcription, mRNA splicing and rRNA processing [55,56]. Gemin3 and Gemin4 are localized both in the nucleus and in the cytosol [56], whereas the rat homolog of eIF2C2 is associated mainly with Golgi or endoplasmic reticulum membranes [31] (Table 1). The subcellular location of the whole RISC complex remains to be determined.

The *Arabidopsis* genome encodes four putative Dicer homologs [57,58]. One of these, Carpel Factory (CAF), has been implicated recently in the generation of microRNAs [59], and it is predicted to be a nuclear protein [60]. Thus, at least in *Arabidopsis*, dsRNA might also be processed to siRNAs in the nucleus (Figure 1). Indeed, non-polyadenylated transgenic hairpin dsRNA, which is designed to inhibit the transcription of homologous promoters and is presumably retained in the nucleus, is processed to small RNAs [1,35]. Putative siRNAs corresponding to a viroid, which is replicated in the nucleus by the host RNA polymerase II, have also been detected in to-

mato [61]. In addition, HC-Pro, a cytoplasmic viral suppressor of PTGS, inhibits the production of siRNAs associated with the post-transcriptional silencing of transgenes [2,62] but does not reduce the small RNAs corresponding to promoter-directed dsRNA in tobacco [1,35].

Recent work involving viral suppressors of RNA-mediated silencing in *Nicotiana benthamiana* and silencing-defective *Arabidopsis* mutants has demonstrated the existence of two size classes of siRNAs [58]. Long (24–26 nt) siRNAs are correlated with systemic silencing in *N. benthamiana* and methylation of retroelement sequences in *Arabidopsis* [58]. Retrotransposon siRNAs in both plant species correspond exclusively to this long class. By contrast, short (21–22 nt) siRNAs are correlated with sequence-specific degradation of a transgenic mRNA [58]. All of these observations support the existence of at least two populations of small RNAs in plants (in addition to microRNAs). Because subtle alterations in Dicer structure have been postulated to alter the spacing between the catalytic centers [3], long and short siRNAs might be processed by different Dicer homologs, which are presumably located in the nuclear (for long siRNAs?) or in the cytoplasmic (for short siRNAs?) compartments.

In several plant species, dsRNA can direct methylation of homologous DNA sequences [1,2,35]. Methylation of genomic DNA occurs even when silencing is induced by RNA viruses, with sequences homologous to nuclear DNA, that replicate exclusively in the cytoplasm [1,2,54]. This again suggests that there is communication between the cytoplasm and the nucleus. When the dsRNA has homology to a promoter, it induces transcriptional silencing in association with DNA methylation [1,35]; however, it is not known whether long dsRNA or processed small RNAs are involved in this process (Figure 1, nucleus).

Connections between the RNAi and the PTGS machinery and chromatin and/or genomic DNA modifications are also starting to emerge in other organisms. In *C. elegans*, mutations in the putative RNA exonuclease MUT-7 reactivate transgenic arrays that are silenced by a polycomb-dependent, presumably transcriptional, mechanism [16]. Some polycomb group homologs, which are normally involved in chromatin repression, are also required for RNAi under certain experimental conditions [63]. Likewise, QDE3, a putative RecQ DNA helicase, is necessary for quelling in *Neurospora* [64]. In *Drosophila*, mutations in Piwi block PTGS and one aspect of transcriptional silencing [47].

Several recent reports have directly implicated the RNAi and PTGS machinery in heterochromatin formation and genome rearrangements [49,65–68] (Figure 1, nucleus). In many eukaryotes, heterochromatin is characterized by a high density of histone H3 methylated at lysine 9 (H3-Lys9) [69]. This modification results in the binding to histone H3 of heterochromatin protein 1 (HP1), and presumably other factors, and formation of a transcriptionally repressive chromatin structure [69]. H3-Lys9 methylation also leads to DNA methylation in *Neurospora* and *Arabidopsis* [69]. But DNA methylation might be a secondary modification that contributes, in certain organisms, to the stability and inheritability of the silent chromatin state.

In *S. pombe*, deletion of DCR1 (a Dicer homolog), RDP1 (a putative RdRP homolog) or AGO1 (an Argonaute homolog) results in the transcriptional derepression of transgenes integrated in centromeric regions, the loss of H3-Lys9 methylation and the dissociation of SWI6 (the *S. pombe* homolog of HP1) from this chromatin domain [49]. In addition, small RNAs complementary to both strands of the centromeric repeats are detected in wild-type cells [65]. These observations have led to a model in which dsRNA originating from the pericentric repeats by bi-directional transcription is processed to siRNAs, which in turn induce H3-Lys9 methylation and heterochromatin formation [49]. Analysis of repetitive DNA integrated into a euchromatic region has shown that the RNAi machinery is required for initiating but not maintaining the heterochromatic state [66].

Studies in the ciliated protozoan *Tetrahymena thermophila* support the model of heterochromatin formation postulated for *S. pombe* [67,68]. *Tetrahymena* contains a transcriptionally inactive micronucleus (with an intact genome), which gives rise to a transcriptionally active macronucleus with extensive deletions corresponding to about 15% of the micronuclear genome. This programmed DNA rearrangement is impaired in a mutant with defects in a homolog of Piwi [67]. Wild-type but not mutant cells contain small RNAs that hybridize preferentially to micronucleus-specific sequences [67]. In addition, H3-Lys9 methylation is required for the targeted DNA elimination in *Tetrahymena* [68]. It is thought that the small RNAs, which are processed from bi-directionally transcribed micronucleus-specific sequences, scan the macronuclear genome, thereby directing H3-Lys9 methylation and subsequent genomic deletions [67,68]. Although small RNAs presumably target chromatin modification through a pairing mechanism, the recognition step and the components that link small RNAs to histone modification are currently unclear. Nevertheless, these observations extend the range of dsRNA-mediated processes and encourage the examination of transcriptional regulation by dsRNA.

4. MicroRNAs

MicroRNAs constitute a class of noncoding small RNAs that are phylogenetically widespread in invertebrates, vertebrates and plants [55,57,59,70–72]. The small temporal RNAs (stRNAs) *lin-4* and *let-7*, which belong to a subclass of miRNAs, were initially identified in *C. elegans* as essential regulators of the timing of development [70]. Mature stRNAs are about 21 nt and can form imperfect duplexes with sequences in the 3′ untranslated regions of target mRNAs, leading to translational repression [70]. Although the biological role of most miRNAs is unknown, they were thought originally to function in translational control or other undefined mechanisms of genetic regulation that are clearly distinct from the RNA cleavage directed by siRNAs. But several new findings are blurring the differences between miRNAs and siRNAs [17,71,72].

MicroRNAs are usually recovered as single-stranded RNAs of 20–25 nt that have been processed from longer stem-loop precursors. In animals, precursor miRNAs (pre-miRNAs) are

apparently encoded as imperfectly pairing inverted repeats of 60–70 nt [55,70]. In plants, however, some predicted pre-miRNAs are more than three times longer than those in animals and involve more extensive or complex stem-loop structures [57,59]. Most pre-miRNAs are encoded in intergenic regions and are probably transcribed from autonomous promoters [57,59,70] (Figure 2); however, several pre-miRNAs are clustered in genomic regions and are apparently synthesized as a single polycistronic RNA [70,73]. In addition, a few miRNAs might also be processed from introns, as byproducts of pre-mRNA splicing, or even from putative protein-coding RNAs [57,59,70].

In mammalian cells, maturation of miRNAs involves at least two steps [73]. Both single and clustered miRNAs seem to be expressed as longer transcripts that are processed into pre-miRNAs of 60–70 nt in the nucleus [73] (Figure 2, nucleus, "pre-processing"). The pre-miRNAs are then exported to the cytoplasm and processed by Dicer, and possibly by other factors, into mature miRNAs of about 21 nt [73] (Figure 2, cytoplasm). Recently, numerous human miRNAs have been found to be associated in a 15S RNP complex that includes Gemin3, Gemin4 and eIF2C2 [55]. The human homolog of *let-7* is a component of this microRNP (miRNP) and

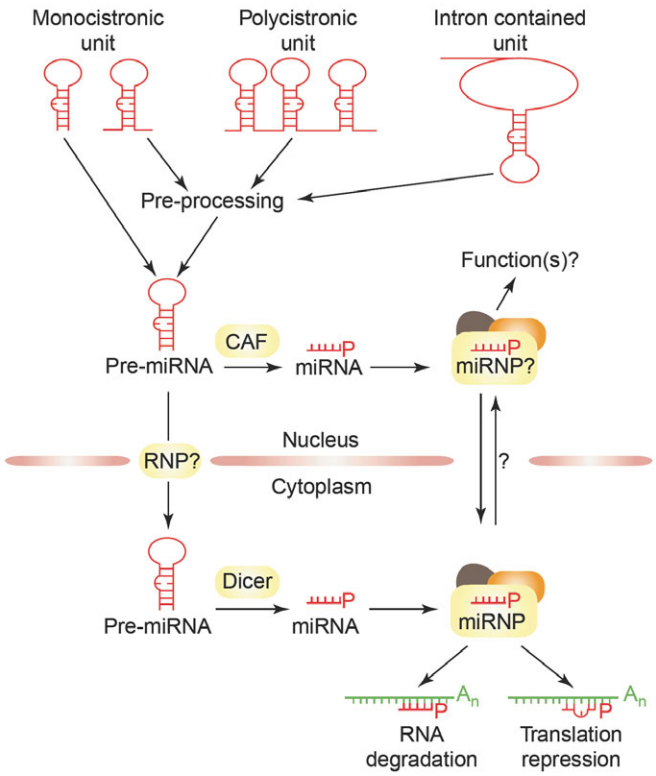

Figure 2. Model of the origin and potential functions of microRNAs. The indicated pathways are mainly supported by experimental evidence in *Arabidopsis*, *Caenorhabditis elegans*, *Drosophila* and human. Green RNA, endogenously transcribed single-stranded RNA; red RNA, microRNAs and microRNA precursors. Protein or protein complexes are indicated by yellow boxes: miRNP, ribonucleoprotein complex containing human Gemin3, Gemin4 and eIF2C2. Several steps related to miRNA biogenesis and function are hypothetical (see text for details). Recent evidence suggests that the miRNP complex is equivalent to the RISC complex (Figure 1) in human cells [17].

can direct the cleavage of a perfectly complementary target RNA [17]. These observations have led to the proposal that the miRNP is the human RISC and can carry out both target cleavage in the RNAi pathway and translational control in the miRNA pathway [17].

The biogenesis of at least some plant miRNAs seems to be different to that in animals. Although CAF, a homolog of Dicer, has been implicated in the generation of miRNAs in *Arabidopsis* [59], pre-miRNAs are detected poorly or not at all in *Arabidopsis*, whereas mature miRNAs are observed readily [57,59]. In addition, whereas in metazoans pre-miRNAs increase when Dicer activity is reduced [74], mutants of CAF show significantly lower levels of miRNAs without any concomitant accumulation of precursor [59]. This suggests that because CAF is predicted to be a nuclear protein [60], at least some plant miRNAs might be processed either co-transcriptionally or shortly after transcription from transient primary transcripts (Figure 2, nucleus). Mature miRNAs might then be transported to the cytosol as part of a RNP complex and/or they might have a role in the nuclear compartment (Figure 2).

Unlike animal miRNAs, whose targets are difficult to identify by sequence complementarity, some *Arabidopsis* and rice miRNAs have been found to show perfect or near-perfect complementarity to potential mRNA targets [71,72]. This suggested that they might function as siRNAs [71,72]. Indeed, an *Arabidopsis* miRNA (miR171/miRNA39) was found to direct specific cleavage of complementary transcripts corresponding to a family of transcriptional regulators [72]. Thus, despite the presence of miRNAs in both plants and animals, current studies imply differences in the structure of predicted precursor RNAs, in the timing and (presumed) subcellular localization of processing and, possibly, in the actual function of mature miRNAs.

5. Biological functions of RNA-mediated silencing

Mutational inactivation of components of the RNAi and PTGS machinery affects at least three distinct eukaryotic processes: the defense response against viruses, transposon mobility and the development of multicellular organisms [1–5,54]. The RNAi and PTGS processes were originally proposed to have evolved to counteract genomic parasites [1–4, 54,75], but it is becoming apparent that dsRNA-mediated mechanisms are also involved in the normal regulation of endogenous genes. *Drosophila* males use an RNAi mechanism to degrade *Stellate* transcripts [75]; one miRNA has been implicated in the specific cleavage of target mRNAs corresponding to the SCARECROW-like transcription factors in *Arabidopsis* [72]; and several miRNAs that are complementary to protein-coding sequences have been identified in *Arabidopsis* and rice [57,59,71].

Intriguingly, a high proportion of the predicted miRNA targets function as developmental regulators in plants, suggesting that miRNAs might have a role in coordinating growth and development [71]. Consistent with this interpretation, *Arabidopsis* CAF mutants, which are defective in miRNA processing, show pronounced developmental alterations [59,60]. In *C. elegans*, developmental defects resulting from reduced function of Dicer and the Argonaute-like proteins ALG-1 and ALG-2 have also been attributed to the improper processing

of miRNA precursors and a reduction in mature stRNA expression [3,74]. It is therefore tempting to propose the existence of ancient, miRNA-mediated mechanisms that regulate endogenous genes in eukaryotes. But most endogenous small RNAs cloned from animals, and several from plants, do not match protein-coding or structural RNAs and their mechanistic roles, with the exception of *C. elegans* stRNAs, remain unknown [55,70,74].

Dicer processes precursor dsRNAs to make both siRNAs and miRNAs. In organisms encoding only one Dicer, a single pathway might handle small RNAs. In other words, the miRNA and siRNA pathways might be interchangeable from biogenesis of the small RNA to interaction with its target. The final outcome, such as mRNA degradation or, for example, translational repression, would depend on the degree of complementarity to the target RNA and, presumably, on associated effector proteins [3,17]. This model is consistent with the findings that short hairpin RNAs, resembling miRNA precursors, can induce RNAi on perfectly homologous target mRNAs [3,17], and that the human RISC seems to be equivalent to the 15S miRNP that is associated with many miRNAs [17,55]. Alternatively, two distinct pathways, intersecting at the Dicer catalyzed step, might be involved in the generation and function of at least some miRNAs and siRNAs [3]. This is supported by the requirement of different Argonaute proteins for the production of functional stRNAs or siRNAs in *C. elegans* [16,19,74]. In organisms where Dicer is encoded by a multigene family, such as *Arabidopsis* [57,58], cytoplasmic and nuclear processing pathways might operate.

Recent findings have also implicated small RNAs in chromatin and/or DNA modifications and genome rearrangements, such as heterochromatin formation in *S. pombe* and DNA elimination in *Tetrahymena* [1,35,49,65–68]. This suggests that dsRNA-mediated processes might have a role in genome organization and transcriptional control. It is clear that despite much progress resulting from a combination of genetics and biochemistry, we are only just beginning to understand the mechanistic complexity of RNA-mediated silencing, its biological implications, and the differences and similarities among different eukaryotes.

Acknowledgements

I thank S. Jacobsen, T. Clemente and members of my laboratory for helpful comments. My apologies to colleagues whose work could not be cited due to space limitations. Supported by NIH grant GM62915.

References

1. M. Matzke *et al.*, RNA: guiding gene silencing. *Science* **293** (2001), pp. 1080–1083.
2. V. Vance and H. Vaucheret, RNA silencing in plants – defense and counter-defense. *Science* **292** (2001), pp. 2277–2280.
3. G. J. Hannon, RNA interference. *Nature* **418** (2002), pp. 244–251.
4. R. H. A. Plasterk, RNA silencing: the genome's immune system. *Science* **296** (2002), pp. 1263–1265.
5. P. D. Zamore, Ancient pathways programmed by small RNAs. *Science* **296** (2002), pp. 1265–1269.
6. P. D. Zamore *et al.*, RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101** (2000), pp. 25–33.
7. E. Bernstein *et al.*, Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409** (2001), pp. 363–366.
8. S. M. Hammond *et al.*, An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404** (2000), pp. 293–296.
9. A. Nykänen *et al.*, ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **107** (2001), pp. 309–321.
10. J. Martinez *et al.*, Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110** (2002), pp. 563–574.
11. S. M. Elbashir *et al.*, RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15** (2001), pp. 188–200.
12. T. Holen *et al.*, Positional effects of short interfering RNAs targeting the human coagulation trigger tissue factor. *Nucleic Acids Res.* **30** (2002), pp. 1757–1766.
13. Y. Han and D. Grierson, Relationship between small antisense RNAs and aberrant RNAs associated with sense transgene mediated gene silencing in tomato. *Plant J.* **29** (2002), pp. 509–519.
14. S. M. Hammond *et al.*, Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293** (2001), pp. 1146–1150.
15. K. Bohmert *et al.*, *AGO1* defines a novel locus of *Arabidopsis* controlling leaf development. *EMBO J.* **17** (1998), pp. 170–180.
16. H. Tabara *et al.*, The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* **99** (1999), pp. 123–132.
17. G. Hutvágner and P. D. Zamore, A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297** (2002), pp. 2056–2060.
18. L. Cerutti *et al.*, Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain. *Trends Biochem. Sci.* **25** (2000), pp. 481–482.
19. H. Tabara *et al.*, The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in *C. elegans*. *Cell* **109** (2002), pp. 861–871.
20. A. Grishok *et al.*, Genetic requirements for inheritance of RNAi in *C. elegans*. *Science* **287** (2000), pp. 2494–2497.
21. R. F. Ketting *et al.*, Mut-7 of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell* **99** (1999), pp. 133–141.
22. J. R. Kennerdell and R. W. Carthew, Use of dsRNA-mediated genetic interference to demonstrate that *frizzled* and *frizzled 2* act in the *wingless* pathway. *Cell* **95** (1998), pp. 1017–1026.
23. P. Svoboda *et al.*, Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development* **127** (2000), pp. 4147–4156.
24. F. Ratcliff *et al.*, A similarity between viral defense and gene silencing in plants. *Science* **276** (1997), pp. 1558–1560.
25. V. Bitko and S. Barik, Phenotypic silencing of cytoplasmic genes using sequence-specific double-stranded short interfering RNA and its application in the reverse genetics of wild type negative-strand RNA viruses. *BMC Microbiol.* **1** (2001), pp. 34–45.
26. R. W. Williams and G. M. Rubin, ARGONAUTE1 is required for efficient RNA interference in *Drosophila* embryos. *Proc. Natl Acad. Sci. USA* **99** (2002), pp. 6889–6894.
27. Y. Kataoka *et al.*, Developmental roles and molecular characterization of a *Drosophila* homologue of *Arabidopsis Argonaute1*, the founder of a novel gene superfamily. *Genes Cells* **6** (2001), pp. 313–325.
28. A. N. Harris and P. M. Macdonald, *aubergine* encodes a *Drosophila* polar granule component required for pole cell formation and related to eIF2C. *Development* **128** (2001), pp. 2823–2832.

29. Y. Zeng and B. R. Cullen, RNA interference in human cells is restricted to the cytoplasm. *RNA* **8** (2002), pp. 855–860.

30. E. Billy *et al.*, Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. *Proc. Natl Acad. Sci. USA* **98** (2001), pp. 14428–14433.

31. D. E. Cikaluk *et al.*, GERp95, a membrane-associated protein that belongs to a family of proteins involved in stem cell differentiation. *Mol. Biol. Cell* **10** (1999), pp. 3357–3372.

32. S. V. Wesley *et al.*, Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant J.* **27** (2001), pp. 581–590.

33. S. Kalidas and D. P. Smith, Novel genomic cDNA hybrids produce effective RNA interference in adult *Drosophila*. *Neuron* **33** (2002), pp. 177–184.

34. R. Reed and E. Hurt, A conserved mRNA export machinery coupled to pre-mRNA splicing. *Cell* **108** (2002), pp. 523–531.

35. M. F. Mette *et al.*, Resistance of RNA-mediated TGS to HC-Pro, a viral suppressor of PTGS, suggests alternative pathways for dsRNA processing. *Curr. Biol.* **11** (2001), pp. 1119–1123.

36. T. Sijen *et al.*, On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107** (2001), pp. 465–476.

37. M. Tijsterman *et al.*, RNA helicase MUT-14-dependent gene silencing triggered in *C. elegans* by short antisense RNAs. *Science* **295** (2002), pp. 694–697.

38. T. Dalmay *et al.*, 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** (2000), pp. 543–553.

39. P. Mourrain *et al.*, *Arabidopsis* SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell* **101** (2000), pp. 533–542.

40. H. Martens *et al.*, RNAi in *Dictyostelium*: the role of RNA-directed RNA polymerases and double-stranded RNase. *Mol. Biol. Cell* **13** (2002), pp. 445–453.

41. C. Lipardi *et al.*, RNAi as random degradative PCR: siRNA primers convert mRNA into dsRNAs that are degraded to generate new siRNAs. *Cell* **107** (2001), pp. 297–307.

42. C. Béclin *et al.*, A branched pathway for transgene-induced RNA silencing in plants. *Curr. Biol.* **12** (2002), pp. 684–688.

43. F. E. Vaistij *et al.*, 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** (2002), pp. 857–867.

44. J.-B. Morel *et al.*, DNA methylation and chromatin structure affect transcriptional and post-transcriptional transgene silencing in *Arabidopsis*. *Curr. Biol.* **10** (2000), pp. 1591–1594.

45. M. R. Rountree and E. U. Selker, DNA methylation inhibits elongation but not initiation of transcription in *Neurospora crassa*. *Genes Dev.* **11** (1997), pp. 2383–2395.

46. T. Buhr *et al.*, Ribozyme termination of RNA transcripts down-regulate seed fatty acid genes in transgenic soybean. *Plant J.* **30** (2002), pp. 155–163.

47. M. Pal-Bhadra *et al.*, RNAi related mechanisms affect both transcriptional and posttranscriptional transgene silencing in *Drosophila*. *Mol. Cell* **9** (2002), pp. 315–327.

48. D. N. Cox *et al.*, *piwi* encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. *Development* **127** (2000), pp. 503–514.

49. T. A. Volpe *et al.*, Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297** (2002), pp. 1833–1837.

50. F. Simmer *et al.*, Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi. *Curr. Biol.* **12** (2002), pp. 1317–1319.

51. M. K. Montgomery *et al.*, RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **95** (1998), pp. 15502–15507.

52. J. M. Bosher *et al.*, RNA interference can target pre-mRNA: consequences for gene expression in a *Caenorhabditis elegans* operon. *Genetics* **153** (1999), pp. 1245–1256.

53. D. Wu-Scharf *et al.*, Transgene and transposon silencing in *Chlamydomonas reinhardtii* by a DEAH-box RNA helicase. *Science* **290** (2000), pp. 1159–1162.

54. P. M. Waterhouse *et al.*, Gene silencing as an adaptive defence against viruses. *Nature* **411** (2001), pp. 834–842.

55. Z. Mourelatos *et al.*, miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev.* **16** (2002), pp. 720–728.

56. B. Charroux *et al.*, Gemin4: a novel component of the SMN complex that is found in both gems and nucleoli. *J. Cell Biol.* **148** (2000), pp. 1177–1186.

57. C. Llave *et al.*, Endogenous and silencing-associated small RNAs in plants. *Plant Cell* **14** (2002), pp. 1605–1619.

58. A. Hamilton *et al.*, Two classes of short interfering RNA in RNA silencing. *EMBO J.* **21** (2002), pp. 4671–4679.

59. B. J. Reinhart *et al.*, MicroRNAs in plants. *Genes Dev.* **16** (2002), pp. 1616–1626.

60. S. E. Jacobsen *et al.*, Disruption of an RNA helicase/RNAse III gene in *Arabidopsis* causes unregulated cell division in floral meristems. *Development* **126** (1999), pp. 5231–5243.

61. I. Papaefthimiou *et al.*, Replicating potato spindle tuber viroid RNA is accompanied by short RNA fragments that are characteristic of post-transcriptional gene silencing. *Nucleic Acids Res.* **29** (2001), pp. 2395–2400.

62. A. C. Mallory *et al.*, HC-Pro suppression of transgene silencing eliminates the small RNAs but not transgene methylation or the mobile signal. *Plant Cell* **13** (2001), pp. 571–583.

63. N. R. Dudley *et al.*, Using RNA interference to identify genes required for RNA interference. *Proc. Natl Acad. Sci. USA* **99** (2002), pp. 4191–4196.

64. C. Cogoni and G. Macino, Posttranscriptional gene silencing in *Neurospora* by a RecQ DNA helicase. *Science* **286** (1999), pp. 2342–2345.

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

66. I. M. Hall *et al.*, Establishment and maintenance of a heterochromatin domain. *Science* **297** (2002), pp. 2232–2237.

67. K. Mochizuki *et al.*, Analysis of a *piwi*-related gene implicates small RNAs in genome rearrangement in *Tetrahymena*. *Cell* **110** (2002), pp. 689–699.

68. S. D. Taverna *et al.*, Methylation of histone H3 at lysine 9 targets programmed DNA elimination in *Tetrahymena*. *Cell* **110** (2002), pp. 701–711.

69. M. Lachner and T. Jenuwein, The many faces of histone lysine methylation. *Curr. Opin. Cell Biol.* **14** (2002), pp. 286–298.

70. A. E. Pasquinelli, MicroRNAs: deviants no longer. *Trends Genet.* **18** (2002), pp. 171–173.

71. M. W. Rhoades *et al.*, Prediction of plant microRNA targets. *Cell* **110** (2002), pp. 513–520.

72. C. Llave *et al.*, Cleavage of *Scarecrow-like* mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* **297** (2002), pp. 2053–2056.

73. Y. Lee *et al.*, MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.* **21** (2002), pp. 4663–4670.

74. A. Grishok *et al.*, Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106** (2001), pp. 23–34.

75. A. A. Aravin *et al.*, Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr. Biol.* **11** (2001), pp. 1017–1027.