Cell
PRESS

Recognizing the enemy within: licensing RNA-guided genome defense

Phillip A. Dumesic and Hiten D. Madhani

Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94158, USA

How do cells distinguish normal genes from transposons? Although much has been learned about RNAi-related RNA silencing pathways responsible for genome defense, this fundamental question remains. The literature points to several classes of mechanisms. In some cases, double-stranded RNA (dsRNA) structures produced by transposon inverted repeats or antisense integration trigger endogenous small interfering RNA (siRNA) biogenesis. In other instances, DNA features associated with transposons – such as their unusual copy number, chromosomal arrangement, and/or chromatin environment – license RNA silencing. Finally, recent studies have identified improper transcript processing events, such as stalled pre-mRNA splicing, as signals for siRNA production. Thus, the suboptimal gene expression properties of selfish elements can enable their identification by RNA silencing pathways.

RNA silencing pathways suppress transposons to protect genome integrity

Transposons have parasitized nearly all eukaryotic genomes, including the human genome, over half of which is derived from transposon sequences. In so doing, transposons have shaped eukaryotic evolution, in part by contributing regulatory and coding information to nearby host genes [1]. In fact, ~4% of human protein-coding open reading frames contain transposon-derived sequences [2], as do 25% of promoter regions [3]. Remarkably, host organisms have even co-opted the protein activities encoded by transposon genes, as in the case of Recombination activating gene 1 (RAG1), a transposon-derived endonuclease that now serves to catalyze V(D)J recombination in humans [4]. Despite these positive contributions to host biology, transposons are primarily selfish elements that threaten the integrity of host genomes. Transposon mobilization can disrupt host genes and promote deleterious chromosomal rearrangements [5,6], and these perturbations contribute to Mendelian disease and cancer in humans [5,7]. To combat this threat, host organisms have evolved multiple genome defense mechanisms to suppress transposon mobility.

RNAi-related RNA silencing pathways are deeply conserved genome defense mechanisms that act in organisms from protist to human to recognize and suppress transposons [8]. In all of these pathways, 20–30 nucleotide small RNAs act in complex with Argonaute family proteins to silence complementary transcripts. Depending on the pathway and context, silencing proceeds by a variety of mechanisms, which include RNA degradation, translational repression, and the establishment of repressive histone modifications [9,10]. The pathways also differ in the means by which they generate small RNAs. In the canonical RNAi pathway, RNase-III-type Dicer enzymes convert long dsRNA into small interfering RNA (siRNA). In contrast, PIWI-interacting RNA (piRNA) pathways do not require Dicer, and instead generate small RNA from single-stranded precursors. Still other pathways, such as the endogenous siRNA pathways of worm and plant, require RNA-dependent RNA polymerases (RdRPs) for small RNA biogenesis; these enzymes are thought to produce small RNAs directly or to generate long dsRNA substrates for Dicer. The diversity of small RNA biogenesis mechanisms permits a wide variety of RNA sequences to template the production of small RNAs for genome defense.

The adaptability of RNA silencing pathways, whose targeting depends primarily on the sequences of their small RNA guides, makes them well suited to defend against transposons, which differ extensively in sequence and distribution among eukaryotic species. Indeed, orthologous Argonaute proteins silence different transposon families in different host organisms, suggesting that RNA silencing pathways can adapt to recognize novel transposons that a host has not previously encountered [11,12]. This adaptability raises a central question in the study of these pathways: how does RNA silencing specifically distinguish transposons from host genes? Here, we review our current understanding of principles by which transposons can be recognized as non-self by RNA silencing pathways of fungi, plants, and metazoans.

The challenges of transposon recognition

RNA silencing pathways must overcome several properties of transposons in order to target them specifically. First, like host genes, transposons reside in the nuclear genome and utilize host gene expression machinery. Thus, transposons are not broadly distinguished by the use of alternative gene expression mechanisms. Second, transposons themselves differ widely in sequence, with many families exhibiting no homology to each other, which hinders sequence-specific transposon recognition mechanisms (Box 1). Finally, even

---

Box 1. Common types of eukaryotic transposable elements

Transposons are nucleic acid elements that can mobilize to new chromosomal locations. They co-opt the host gene expression machinery to produce their own protein products, which facilitate mobilization. Although common in these respects, transposons are remarkably diverse in sequence and transposition mechanism, which hinders their identification by cellular genome defense pathways.

Transposons can be broadly divided into retrotransposons (Class I) and DNA transposons (Class II) Figure I [85,86]. Retrotransposons mobilize through an RNA intermediate, which is reverse transcribed to allow its integration into the genome. This 'copy and paste' mechanism does not alter the original transposon locus and therefore acts to increase transposon copy number. Some retro-transposons encode long terminal repeats (LTRs), which act as promoters and polyadenylation signals. Like their retrovirus rela-tives, these transposons undergo reverse transcription in virus-like particles in the cytoplasm. Other retrotransposons, such as LINEs, also encode their own promoters and 3′ end formation signals, but lack LTRs. These elements undergo reverse transcription in the nucleus using nicked genomic DNA as a primer. By contrast, short interspersed nuclear elements (SINEs) mobilize in a manner similar to that of LINEs, but do not themselves encode the proteins required for mobilization. They are therefore nonautonomous, and depend on LINE-encoded factors.

DNA transposons do not utilize an RNA intermediate, but instead generally mobilize through a 'cut and paste' mechanism in which the original transposon locus is excised and reinserted in a new location. These transposons generally encode terminal inverted repeat (TIR) sequences, which recruit transposase to the transposon DNA locus in order to initiate its excision.
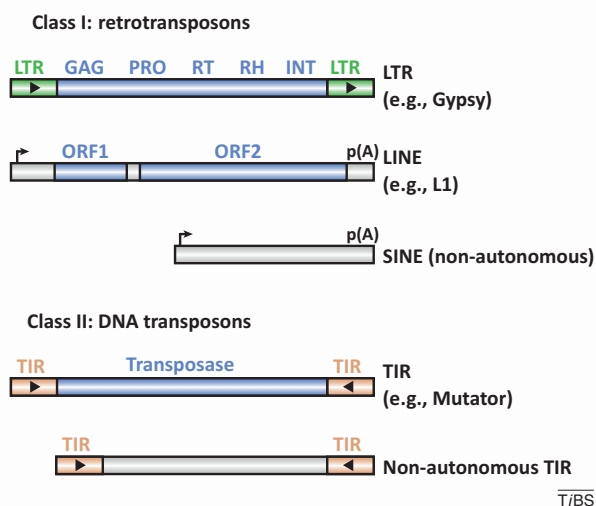

Figure I. Examples of common eukaryotic transposons. Common Class I transposons include LTR retrotransposons, which generally encode a capsid protein (GAG), protease (PRO), reverse transcriptase (RT), RNaseH (RH), and integrase (INT). By contrast, LINEs typically contain two open reading frames: one of unknown function and one that encodes endonuclease and reverse transcriptase activities. SINEs are nonautonomous elements whose sequence features are recognized by transposon-derived proteins acting *in trans*. Most Class II elements are TIR transposons, which can mobilize either autonomously or nonautonomously. Blue rectangles indicate protein-coding regions.

Studies of RNA silencing pathways have begun to reveal strategies by which these pathways recognize transposons while avoiding inappropriate silencing of host sequences. These strategies are not mutually exclusive, and an individual RNA silencing pathway sometimes utilizes multiple strategies. One general strategy recognizes transposons by virtue of their tendency to generate dsRNA. A second strategy distinguishes transposons by their unique ability to mobilize, which makes them more likely to exist in unusual chromosomal arrangements or in high copy number. A third class of mechanism exploits the suboptimal gene expression properties of transposons, which might arise due to their distinct evolutionary histories, to distinguish them from host genes. A fourth class of mechanism licenses small RNA production against transposons based on the prior capture of transposon sequences by specialized chromatin niches. These four transposon recognition strategies are discussed in turn in this review.

Transposons are identified by their production of dsRNA

Early studies of *Caenorhabditis elegans* RNAi have indicated that some mutants defective in gene silencing triggered by exogenous dsRNA are also defective in suppressing transposons, raising the possibility that endogenous dsRNA initiates transposon silencing [14,15]. In such a model, transposon-derived dsRNA is processed by Dicer enzymes to yield siRNA, which then acts to repress homologous transposon sequences throughout the genome. An attractive feature of this model is that transposons exhibit several properties that might increase their likelihood of generating dsRNA, thereby enabling them to be distinguished from host genes [16]. For instance, some transposon families encode repeats and antisense promoters that can produce dsRNA. Furthermore, the mobilization of transposons into existing host transcriptional units may lead to the production of antisense transposon transcripts. Finally, the repetitiveness of transposon sequences in the genome, together with their tendency to undergo rearrangements, may promote the formation of structured loci in which duplicated transposon sequences give rise to transcripts that fold to form dsRNA (Figure 1A).

Several of the above mechanisms of transposon dsRNA production have been validated experimentally. For instance, the inverted repeats of Tc1 DNA transposons in *C. elegans* have been shown to form dsRNA *in vivo*; likely due to intramolecular folding of precursor transcripts that originate from host promoters and read through an entire transposon locus [17]. By contrast, an internal antisense promoter in the human long interspersed element (LINE)-1 is required for the production of small RNA targeting this transposon, suggesting that intermolecular dsRNA triggers LINE-1 siRNA production, although the Dicer dependence of these small RNAs remains to be demonstrated [18]. Finally, unusual tandem arrangements of transposon-derived sequences, such as inverted duplications, have been found in mouse, nematode, plant, and yeast to be hotspots of siRNA production, typically owing to formation of intramolecular dsRNA [12,17,19,20]. Remarkably, a single locus of this type in maize, the Mu killer locus, is sufficient to trigger silencing of Mutator DNA transposons

for a single transposon type, individual loci differ extensively, from fully intact, active transposons to degenerated, inactive loci. This means that transposon sequences are not always distinguished from host genes simply by their ability to mobilize. Furthermore, both active and inactive transposon loci might be deleterious to their hosts, because the expression of these abundant sequences, which constitute >80% of some eukaryotic genomes, confers a metabolic cost and may induce RNA toxicity [13].
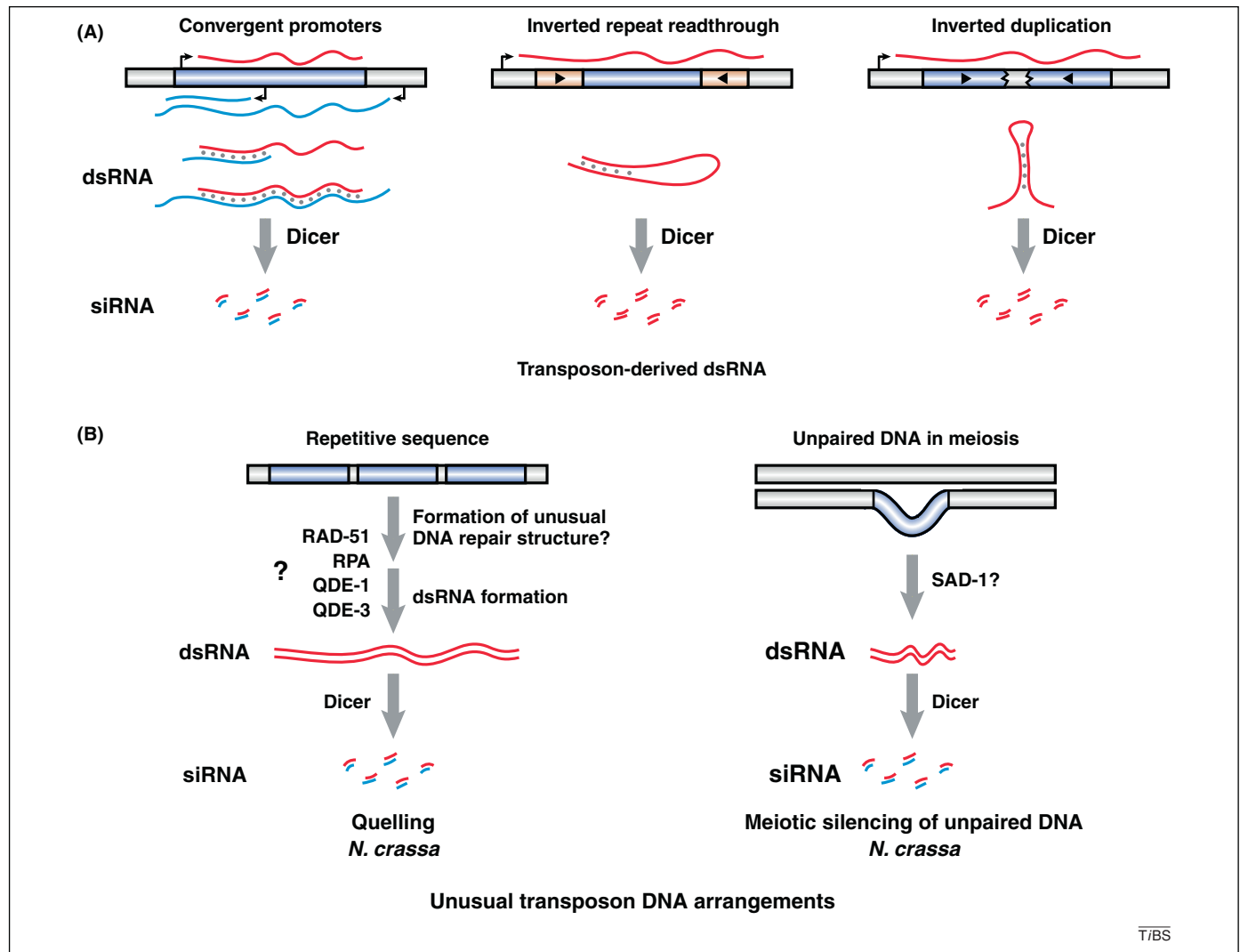
Figure 1. Transposon features recognized by RNA silencing pathways. **(A)** Transposon-derived double-stranded RNA (dsRNA) provides a substrate for Dicer activity and thereby triggers endogenous small interfering RNA (siRNA) pathways. dsRNA can be generated intermolecularly by the action of convergent promoters encoded either by the transposon (blue bar) or by the host. Intramolecular dsRNA is generated from a single transposon when a transcript contains both of its inverted repeat sequences. Inverted duplications of transposons can also cause intramolecular dsRNA formation. **(B)** Unusual chromosomal arrangements of transposons allow their identification by genome defense pathways. In *Neurospora crassa*, the quelling pathway (left) silences repetitive sequences. The mechanism by which these sequences are detected and used to template siRNA production is unclear, but may involve the formation of unusual DNA repair intermediates, because quelling requires proteins, such as RAD-51, that mediate homologous DNA recombination. Subsequent siRNA production requires QDE-1 – a DNA- and RNA-dependent RNA polymerase – and its binding partners replication protein (RP)A and QDE-3. The meiotic silencing of unpaired DNA pathway (right) detects loci that lack a partner during homologous chromosome pairing in meiosis I. siRNA production from these loci requires SAD-1, an RNA-dependent RNA polymerase (RdRP) that is paralogous to QDE-1.

genome-wide [19]. Thus, the capacity of RNA silencing to act *in trans* enables an entire family of transposons to be recognized, even if only a subset of these loci generates dsRNA.

The above examples suggest that the transposon-derived dsRNA can, in at least some cases, trigger the production of small RNAs for the purpose of genome defense. Nevertheless, the general rules that dictate the *in vivo* formation of dsRNA, and its subsequent processing into siRNA, remain unclear. For instance, although the intramolecular dsRNA produced by the Mu killer locus can induce RNA silencing of homologous Mutator transposons, this effect cannot be achieved by expression of sense and antisense Mutator transcripts from distinct loci [21], suggesting that intermolecular dsRNA comprising these transcripts is either inefficiently formed or poorly processed by Dicer. Similarly, the 31 genomic copies of Tc1 in *C. elegans* generate both sense and antisense transcripts, but these transcripts do not appear to form intermolecular dsRNA, as assessed by analysis of *in vivo* dsRNA editing by adenosine deaminase [17]. These findings suggest that the initiation of RNA silencing by dsRNA may require a licensing step, in which only a subset of potential dsRNA substrates gains the capacity to trigger RNA silencing. Such a licensing step may also explain why endogenous non-transposon loci that produce complementary transcripts are often poor triggers of siRNA biogenesis, as in *Drosophila* [22].

Future experiments that systematically address the relations between cellular single-stranded RNA (ssRNA), dsRNA, and siRNA will help elucidate the rules by which dsRNA triggers small RNA production *in vivo*. These rules may involve the ability of given RNAs to form dsRNA, which could be influenced by their expression levels, subcellular localizations, or associated proteins. Alternatively, specific features of the dsRNA itself, such as whether the chemistry of its ends allows efficient Dicer processing,

could determine the efficiency of siRNA production. A more complete understanding of the rules for dsRNA and siRNA production could explain why only some transposon families are targeted by endogenous siRNA, and why different targeted families exhibit siRNAs corresponding to distinct regions of their sequence [23].

Unusual DNA arrangements distinguish transposons from host genes

Transposons are often found in unusual chromosomal arrangements. The repetitive sequences they encode can cause replication fork slippage, leading to the formation of tandem arrays [24,25]. Furthermore, their ability to mobilize can lead to the asymmetric distribution of transposon sequences on homologous chromosomes. These unusual arrangements, which represent targets of the two RNA silencing pathways in the filamentous fungus *Neurospora crassa*, may thus distinguish transposons for detection by host genome defense pathways (Figure 1B). The quelling pathway operates in vegetative cells, where it targets tandem array sequences for silencing. By contrast, the meiotic silencing of unpaired DNA (MSUD) pathway acts in mating cells, where it silences loci that lack a partner during homologous chromosome pairing in meiosis I. Each pathway requires its own RdRP, suggesting that silencing is triggered not by a naturally occurring dsRNA, but rather by the directed action of RdRP activity at particular targets. These targets, however, do not include active transposons, which are not found in the modern-day genome of *N. crassa*. Instead, RNA silencing pathways in this organism are studied primarily in their capacity to silence transgenes. Nevertheless, *N. crassa* produces small RNAs corresponding to ancient, degenerated transposon sequences in its genome, pointing to genome defense as a biological role for these pathways [26].

The quelling pathway targets repetitive DNA arrangements

Quelling, one of the first known RNA silencing pathways, was discovered >20 years ago [27,28]. In this pathway, repetitive transgenes, which are often oriented in tandem arrays, are used as templates for the production of siRNA, which post-transcriptionally silences homologous loci throughout the genome. Like other endogenous siRNA pathways, quelling requires Dicer enzymes, an Argonaute protein (QDE-2), and an RdRP (QDE-1) [27]. Interestingly, QDE-1 can act not only as an RdRP, but also as a DNA-dependent RNA polymerase, and recombinant QDE-1 is sufficient to generate dsRNA from a DNA template [29]. These findings, together with the observation that quelling can be triggered by transgenes lacking Pol II promoters [30], suggest that quelling may be initiated by QDE-1-mediated production of ssRNA and dsRNA from transgene sequences. In such a model, the recruitment of QDE-1 to particular loci would underlie the specificity of quelling for transgenes and other foreign genetic elements.

Recent findings outline a model in which QDE-1 recruitment to transgenic loci is triggered by the repetitiveness of the DNA itself. An initial clue came from the finding that QDE-1 physically interacts with replication protein (RP)A in a manner that requires QDE-3, a quelling factor related

to RecQ helicases [29,31]. Because RPA binds ssDNA to facilitate DNA replication and repair, and RecQ helicases play roles in DNA repair, these findings suggest a functional link between DNA metabolism and quelling. Further studies have indicated that factors involved in homologous DNA recombination bind to transgenic loci and are required for quelling, supporting a model in which homologous recombination generates recombination intermediates that recruit QDE-3, and subsequently, QDE-1 for the purpose of small RNA production [32]. These events may take place more often at repetitive loci because these loci are more likely to undergo homologous recombination, or because their repetitiveness leads to the production of aberrant recombination structures that are resolved by QDE-3. Consistent with such a model, treatment of cells with hydroxyurea, an agent that induces double-stranded DNA breaks, enhances quelling [32]. Although the detailed mechanism of QDE-3 recruitment to and QDE-1 action at repetitive loci remain areas of active investigation, these findings point to homologous recombination as a mechanism to sense the repetitiveness of particular genomic sequences, thereby identifying them as targets for a genome defense pathway.

Meiotic silencing of unpaired DNA

In contrast to quelling, which targets foreign sequences present in multiple genomic copies, the MSUD pathway can recognize even a single-copy sequence as foreign based on its inability to pair with a homologous chromosome during meiosis [33]. This pathway requires a Dicer protein, an Argonaute, and an RdRP (SAD-1), which generate small RNA corresponding specifically to the unpaired region [34]. All of these proteins exhibit perinuclear localization, suggesting a specialized site for small RNA biogenesis [27]. In fact, only a single known MSUD factor localizes to the nucleus: SAD-5, a fungal-specific protein with no predicted domains [35]. This has hindered progress on a central question in the study of meiotic silencing: how precisely are unpaired regions detected in the nucleus and engaged for small RNA production?

Studies to date suggest that the detection of unpaired regions is not restricted to particular DNA sequences or chromosomal regions [33]. Rather, detection efficiency correlates with the size of the unpaired region as well as the degree of nonidentity it shares with its homolog [36]. MSUD can detect even subtle amounts of sequence non-identity (~5% across 2–4 kb), and its sensitivity is further increased if the unpaired region carries DNA methylation marks, suggesting that not only DNA sequence but also chromatin context may contribute to detection of unpaired regions [37]. Detection proceeds even if the unpaired region lacks a Pol II promoter, further suggesting that the initial detection event involves recognition of some aspect of DNA structure [36]. Presumably, noncanonical transcription produces the RNA transcripts that template dsRNA and siRNA production. An attractive, but so far untested, model is that SAD-1 acts successively as a DNA-dependent and RNA-dependent RNA polymerase – like its paralog, QDE-1, in the quelling pathway – in order to synthesize dsRNA from unpaired loci. Future studies will therefore benefit from characterization of the precursor transcripts

that give rise to small RNAs from unpaired regions. Another area of investigation will be to determine whether chromatin states influence the MSUD pathway, either in its detection or small RNA production steps. In this regard, it is interesting to note that a meiotic silencing pathway in *C. elegans* also requires an RdRP (Ego-1), and is coupled to histone H3 lysine 9 (H3K9) methylation of unpaired loci [38,39].

Sensing of transposons via the gene expression machinery

Although the quelling and MSUD pathways recognize properties of the DNA copy number or chromosomal arrangement of a foreign element, other RNA silencing pathways appear instead to target features of transposon-encoded RNAs. What features could distinguish transposon transcripts from host transcripts, because both are produced and undergo processing by the same host cellular machineries? As discussed further below, recent findings in the yeast *Cryptococcus neoformans* suggest that transposon-derived transcripts encode suboptimal splicing signals and tend to stall on spliceosomes, which directs their use as templates for dsRNA and siRNA production [40]. Although transposon gene expression signal strength has not been systematically examined in other organisms, plant genome defense pathways mediated by endogenous siRNA may also detect features of transcript maturation, as evidenced by the finding that 5′ and 3′ end formation features can dictate RNA silencing specificity in these systems. Together, these results raise the possibility that gene expression signal strength generally distinguishes transposons from host genes, perhaps because the evolutionary history of transposons differs from that of their host organisms.

The spliceosome is a transposon sensor in a basidiomycetous yeast

In *C. neoformans*, endo-siRNA mediates post-transcriptional gene silencing and is produced primarily from transposon-derived sequences [40,41]. Strikingly, these small RNAs correspond to both exons and introns of their targets, suggesting that incompletely spliced mRNA precursors are preferred templates for siRNA production. Furthermore, several *C. neoformans* siRNA biogenesis factors, including an Argonaute and an RdRP, form a spliceosome-coupled and nuclear RNAi (SCANR) complex that physically interacts with the spliceosome. These findings point to a kinetic competition model for RNA silencing specificity, in which efficiently spliced host transcripts are not targeted for siRNA production, whereas transposon transcripts, by virtue of their poor splicing kinetics, become susceptible to SCANR-mediated siRNA production (Figure 2). Consistent with this model, the transcripts targeted for siRNA production exhibit intron sequence features predictive of poor splicing and tend to accumulate abnormally on spliceosomes *in vivo* [40].

Experimental manipulations of siRNA target transcripts in *C. neoformans* further indicate that splicing signals dictate RNA silencing specificity. For instance, elimination of the introns of a transcript suppresses its ability to template siRNA biogenesis. Conversely,
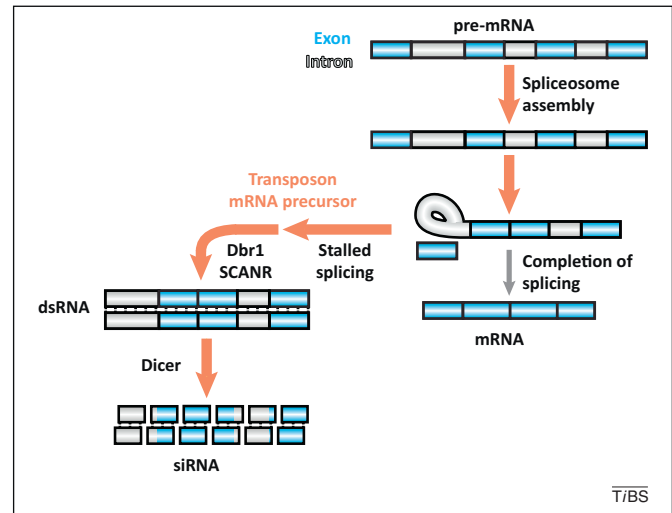

Figure 2. Stalled spliceosomes license RNA silencing in an endogenous small interfering RNA (siRNA) pathway. In the yeast *Cryptococcus neoformans*, transcripts targeted by RNA silencing, which primarily include transposons, exhibit sequence features predictive of poor splicing and tend to stall in spliceosomes. The stalled splicing of transposon mRNA precursors is required for siRNA biogenesis mediated by a spliceosome-coupled and nuclear RNAi (SCANR) complex that contains an RNA-dependent RNA polymerase and physically associates with the spliceosome. Lariat debranching enzyme (Dbr1) is also required for siRNA production, suggesting that transposon mRNA precursors in the lariat intermediate stage are linearized to enable double-stranded RNA (dsRNA) formation by SCANR. In this hypothetical example, splicing of the first intron of a transcript stalls at the lariat intermediate stage, whereas downstream introns remain partially spliced.

introduction of a 3′ splice site mutation, which stalls splicing at the stage of lariat intermediate formation, dramatically increases siRNA production from the mutated transcript. Importantly, the effect of a 3′ splice site mutation is suppressed by introduction of a 5′ splice site mutation, which prevents intron engagement with the spliceosome. The effect of a 3′ splice site mutation is also suppressed in the absence of lariat debranching enzyme, further supporting a model in which the processing of stalled splicing intermediates is required for the production of dsRNA and subsequent siRNA. Together, these findings establish stalled spliceosomes as a necessary signal by which transposon-derived transcripts can be distinguished for the purpose of genome defense [40].

Although the reason why siRNA targets in *C. neoformans*, which are primarily transposon-derived, possess weaker splicing signals than those of host genes is not yet understood, this property may have been shaped by the distinct features of transposon propagation. First, intron-containing retrotransposons must produce both spliced and unspliced transcripts, with the latter used selectively as transposition substrates [42], in order to maintain intron sequences in new genomic copies. Weak splicing signals may thus be favored as a mechanism for producing both transcript types. Second, the capacity of transposons to be horizontally transferred between different organisms may predispose them to possess suboptimal splicing signals. For instance, because gene expression signals, such as splicing signals, differ substantially among species [43,44], the horizontal transfer of a transposon may insert it into a host, like *C. neoformans*, with incompatible signal preferences. The limited coevolution of transposons with their hosts would thus make transposons in general more

poorly equipped than host genes to undergo the many steps of eukaryotic gene expression. This potential disparity provides opportunities for the evolution of genome defense pathways, and may even contribute to the evolutionary emergence of host gene expression steps themselves [45].

RNA processing signals affect RNA silencing in plants
Several findings raise the possibility that RNA silencing specificity in plants is influenced, as in *C. neoformans*, by the RNA processing signals encoded in foreign elements. In plants, two endo-siRNA pathways – each with distinct Dicer, Argonaute, and RdRP proteins – act to generate siRNA against transposons. In the post-transcriptional gene silencing (PTGS) pathway, transposon transcripts produced by RNA Pol II are used as substrates for siRNA production in a manner thought to be initiated by RNA-dependent RNA polymerase 6 (RDR6) [46]. In the RNA-directed DNA methylation (RdDM) pathway, transcription of transposon loci by the specialized RNA polymerases Pol IV and Pol V enables siRNA production and DNA methylation, respectively, at these loci [47]. DNA methylation not only represses transcription but also recruits additional Pol IV activity. Therefore, this pathway leads to lasting, heritable silencing, such that most transposons in plant genomes are not normally expressed [48]. Newly encountered transposons can be targeted for methylation if they exhibit sequence similarity to existing RdDM targets, presumably because of the *in trans* action of RdDM pathway siRNA [48]. How can a transposon of novel sequence be identified? Recent studies suggest that the initial recognition of transposon transcripts by the PTGS pathway can subsequently engage the RdDM pathway to effect lasting DNA methylation [48,49]. These findings highlight RDR6 activity as a key factor in transposon recognition and RNA silencing specificity in plants.

Promoter sequences are required for the detection of foreign elements by PTGS, suggesting that RNA signals guide the specificity of this pathway [50]. However, these signals remain largely unknown, in part due to the challenges of manipulating and analyzing repetitive, endogenous transposons. Nevertheless, studies of transgene silencing have provided several clues. First, PTGS can recognize even single-copy transgenes that are not predicted to form dsRNA, supporting the model that RDR6 action on single-stranded, non-self transcripts is an initiating event in this silencing pathway [51,52]. Second, screens for *Aribidopsis thaliana* mutants that enhance PTGS yielded mutations affecting the 5′-3′ RNA exonucleases XRN2/3/4 as well as 3′ end formation factors that mediate transcript cleavage and polyadenylation [53–55]. These mutants accumulate uncapped or nonpolyadenylated transgene transcripts, suggesting a model in which aberrant transcripts with one or more incompletely processed ends may be particularly good substrates for RDR6. Consistent with this model, siRNA production from a transgene transcript is promoted by the absence of transgene termination signals, and is suppressed by the addition of strong terminators [52]. Interestingly, an XRN4 mutation does not simply increase silencing of appropriate PTGS targets, but also stimulates the production of siRNA from host transcripts not normally targeted by PTGS,

further suggesting that 5′ end formation acts to generally suppress RDR6 activity [56].

Several questions regarding the initiation of PTGS by foreign elements remain to be answered. First, it will be important to determine the precise substrate of RDR6 activity *in vivo*. The results described above highlight transcripts with bare 5′ or 3′ ends as an attractive possibility, and are further supported by the observation that host transcripts, if separated from their 5′ and 3′ end modifications by miRNA-mediated cleavage events, become RDR6 substrates [57]. However, it is unclear whether this model is sufficient to explain the observation that perturbations of several other RNA processing steps– including splicing [53,58], nonsense mediated decay [59], and exosome-mediated decay [59]– also affect PTGS targeting. It is possible that all of these perturbations act by indirectly influencing a single preferred RDR6 substrate, whose identity remains hidden. Second, it will be important to understand how RDR6 substrate preferences enable PTGS to distinguish self from non-self. Anecdotal evidence suggests that repetitive sequences tend to produce poorly polyadenylated transcripts [52], but whether transposons in plant genomes generally exhibit poor polyadenylation or cap formation has not yet been systematically tested.

Genomic memory of foreign sequences enables transposon detection

RNA silencing pathways of the piRNA class silence transposons and other foreign elements in animal germlines, where they are required for fertility [60]. In these pathways, small RNAs are produced in a Dicer-independent manner, then act with Argonaute proteins of the PIWI clade to effect transcriptional and post-transcriptional silencing of complementary target transcripts [61]. Although the molecular mechanisms of piRNA biosynthesis and target silencing vary among species, piRNA pathways share a common 'adaptive immune system' strategy for foreign element recognition. Specifically, these pathways generate a large diversity of piRNA sequences, which survey the transcriptome for complementary targets. Upon target recognition, an amplification step generates additional target-specific small RNAs, which enforce long-term silencing, often in a heritable manner [61]. Thus, piRNA pathway specificity for transposons is determined in large part by the initial selection of piRNA sequences, a process that in *Drosophila* involves the restriction of piRNA production to specialized genomic loci called piRNA clusters (Figure 3A). Regulation of secondary piRNA amplification provides an additional opportunity to impose silencing specificity, and recent studies suggest that this is a significant source of piRNA specificity in *C. elegans* (Figure 3B).

piRNA in D. melanogaster
In *Drosophila*, germline piRNAs are produced from specialized loci that contain clusters of transposon fragments [62,63]. Clusters thus represent a genomically encoded memory of prior transposon insertions. Single-stranded transcripts from these clusters undergo processing to generate primary piRNA, which act with PIWI family proteins (Piwi, Aub) to guide cleavage of complementary transposon transcripts originating throughout the genome. This
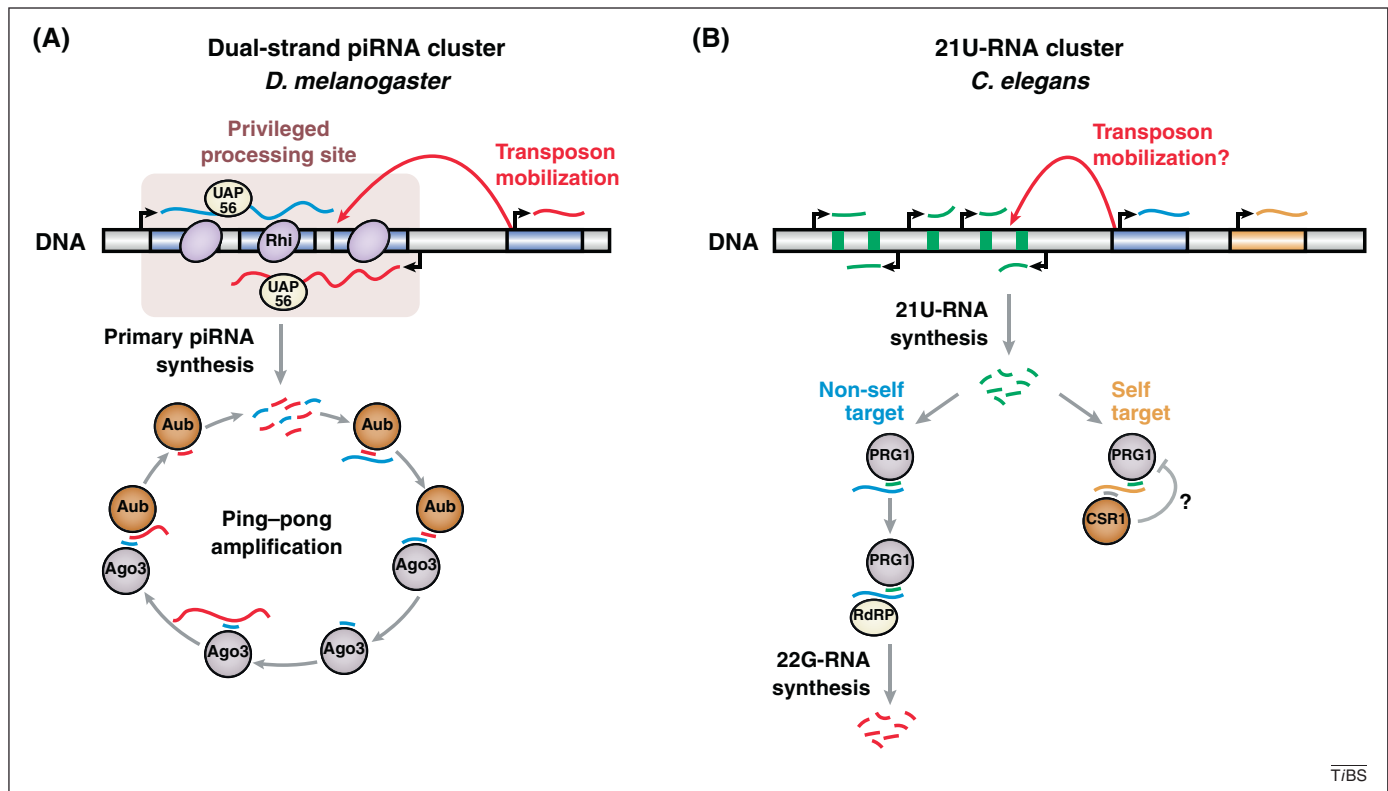
Figure 3. PIWI-interacting RNA (piRNA) production occurs at specialized genomic loci. **(A)** In *Drosophila melanogaster*, the chromatin environment at dual-strand piRNA clusters enables piRNA production from these loci. This chromatin is associated with H3K9 methylation and the specialized heterochromatin protein (HP)1 homolog Rhino (Rhi). Primary transcripts produced from piRNA clusters are bound by UAP56, a DEAD box protein required for piRNA processing specificity [87]. The ping–pong amplification cycle subsequently promotes the generation of piRNAs complementary to active transposons. Mobilization of transposons into piRNA clusters contributes to the enrichment of foreign sequences at these loci. **(B)** In *Caenorhabditis elegans*, 21U-RNAs are encoded in genomic clusters, but each is expressed by its own promoter. 21U-RNA sequences are diverse, and not strongly enriched in transposon sequences, suggesting that transposon mobilization into 21U-RNA clusters is not a major mechanism by which 21U-RNA specificity is achieved. 21U-RNAs loaded in the PIWI protein PRG-1 are potentially complementary to both non-self and self transcripts. In the former case, PRG-1 triggers the production of repressive 22G-RNAs to silence the non-self transcript. In the latter case, PRG-1-mediated silencing appears to be less efficient, potentially due to a protective mechanism. This protective mechanism is hypothesized to involve targeting of self transcripts by 22G-RNAs loaded in the Argonaute protein CSR-1.

cleavage defines the 5′ end of a secondary piRNA, which is then loaded into a distinct PIWI protein (Ago3) and undergoes 3′ trimming to proper length. Subsequently, mature Ago3 complexes act to bind and cleave complementary piRNA cluster transcripts repeatedly, thereby promoting precursor conversion into mature piRNAs [62,64,65]. This amplification step, termed the ping–pong cycle, biases piRNA production towards transcriptionally active transposons, and ensures that most piRNAs are antisense to their targets, and therefore active [65,66]. These abundant antisense piRNAs then promote both post-transcriptional and transcriptional transposon silencing, with the latter being mediated by H3K9 methylation of target loci [61].

A central question in *Drosophila* piRNA specificity is to understand how particular genomic regions are marked as sources of piRNA production in such a way that ensures their enrichment for transposon sequences. In one model, piRNA clusters act as passive transposon traps. In this model, the repetitiveness and mobility of transposons make them more likely than host genes to integrate at a defined genomic locus, such as a piRNA cluster [67]. Each integration event would enable piRNA-mediated silencing of the integrated transposon and its relatives, and tend to be fixed in the population. This model is supported by the finding that piRNA clusters can produce piRNAs from a wide variety of experimentally inserted sequences [68],

and by the observation that, among *Drosophila* species, the genomic location of a piRNA cluster is often conserved even when the constituent transposon fragments differ [69]. Furthermore, the initiation of silencing of a transposon by its mobilization into piRNA clusters has been observed experimentally [70].

Other observations raise the possibility that the selection of sequences for piRNA biogenesis involves additional licensing steps. For instance, transgenes inserted into piRNA clusters template the production of piRNAs in a nonuniform manner [68,71], suggesting that piRNA precursor processing preferences may dictate selection of particular sequences that are best suited for genome defense. Furthermore, some transposon families appear to mobilize to piRNA cluster loci more often than would be expected by chance [70,72], suggesting that an active mechanism promotes insertion into a cluster. One possibility is that the specialized chromatin landscape at piRNA clusters, which includes H3K9 methylation and a heterochromatin protein (HP)1 homolog (Rhino), promotes local transposon insertions [73,74]. Another intriguing possibility, suggested by recent work [75,76], is that the silencing of piRNA target loci leads to the birth of new piRNA clusters at these loci, potentially in a mechanism involving the H3K9 methyl marks deposited during Piwi-mediated silencing [61]. Such a mechanism would enrich transposon

Table 1. Proposed mechanisms for licensing of RNA silencing pathways

RNA silencing pathway	Organism	Proposed licensing mechanism	Refs.
Endogenous siRNA	*H. sapiens*	Formation of dsRNA substrates for Dicer	[12,17–20,23]
	M. musculus		
	D. melanogaster		
	C. elegans		
	Z. mays		
	S. castellii		
Quelling	*N. crassa*	Unusual DNA repair structure at tandem repeat locus	[32]
Meiotic silencing of unpaired DNA	*N. crassa*	Unpaired DNA in meiosis I	[33,38]
	C. elegans		
Endogenous siRNA	*C. neoformans*	Stalled splicing of target transcript	[40]
Post-transcriptional gene silencing	*A. thaliana*	Improper target transcript 5′ or 3′ end formation	[52–56]
piRNA	*D. melanogaster*	Privileged chromosomal loci for piRNA production (e.g., specialized chromatin)	[62,74]
	C. elegans	Absence of self-protective signals in target (e.g., CSR-1 targeting?)	[82,83]

and other piRNA target sequences within piRNA clusters. Further testing of these hypotheses awaits a more complete understanding of the specialized chromatin at piRNA clusters, and how its constituent proteins, such as Rhino, enable piRNA production from these loci.

21U-RNA in C. elegans
In *C. elegans*, the PIWI protein PRG-1 associates with small RNAs of the 21U class [60]. Like *Drosophila* piRNAs, 21U-RNAs are encoded largely in genomic clusters [77], but they differ in that 21U-RNAs are expressed as autonomous units instead of from long precursor transcripts [78,79]. Recent findings indicate that 21U-RNAs can trigger repression of transposon and transgene sequences [11,80–83]. Recognition of a target transcript by a 21U-RNA leads to RdRP-dependent production of 22G siRNA of the worm-specific Argonaute (WAGO) class, which are antisense to the target. Production of 22G siRNA not only serves to amplify the RNA silencing response, but also engages transcriptional silencing mechanisms, which depend on H3K9 methylation and can be stably inherited. However, the pool of 21U-RNAs is vast (>16 000 distinct species) and complex in sequence, raising the question of how genome surveillance by these RNAs can specifically distinguish self from non-self.

The sequences of 21U-RNAs do not appear sufficient to establish specificity for foreign elements. For instance, although 21U-RNAs are depleted in sequences that would target host genes, they do not show a strong enrichment for transposon sequences [11]. Furthermore, 21U-RNAs appear to direct PRG-1 to complementary targets even when the targets contain as many as four mismatches with the 21U-RNA sequence [11,82]. Given such relaxed sequence specificity, the 21U-RNA population would be expected to target most host genes [11,82]. To explain the lack of self-reactivity in 21U-RNA pathways, it has been proposed that an additional mechanism protects self transcripts from 21U-mediated silencing by antagonizing the amplification of repressive 22G siRNA. This mechanism is proposed to involve CSR-1, an Argonaute protein involved not in gene silencing, but rather in facilitating chromosome segregation via a class of 22G siRNA that targets host genes expressed in the germline [84]. Consistent with such a

model, 21U-RNAs are less efficient in silencing a target transcript that is also targeted by 22G siRNAs in CSR-1 [82]. Furthermore, *C. elegans* transgenes that escape 21U-mediated silencing can establish a stably active state that dominantly activates silenced, homologous loci *in trans*, consistent with the idea that diffusible factors, such as CSR-1-bound small RNAs, act to license host gene expression [83]. Together, these findings suggest that piRNA pathway specificity in worms relies not only on a comparison of transcript and 21U-RNA sequences, but also on features of transcript expression history, as potentially indicated by CSR-1 22G RNAs.

Concluding remarks
We have summarized evidence that RNA silencing pathways recognize several distinguishing features of transposons, including their tendency to produce dsRNA, exist in unusual chromosomal arrangements, exhibit suboptimal gene expression properties, and occupy specialized chromatin contexts (Table 1). The extent to which these features are sufficient to distinguish transposons from host genes is unknown. In this regard, it is interesting to note that distinct RNA silencing pathways may act combinatorially to identify non-self elements, as has been suggested for the MSUD and piRNA pathways in *C. elegans* [81]. Recent observations further suggest that the specificity of RNA silencing pathways for transposons may involve not only distinguishing signals in the transposons themselves, but also protective signals possessed by host genes. Efforts to understand the nature of these signals should reveal fundamental principles by which organisms detect and silence genomic parasites while avoiding inappropriate targeting of beneficial genomic sequences.

References
1 Feschotte, C. (2008) Transposable elements and the evolution of regulatory networks. *Nat. Rev. Genet.* 9, 397–405
2 Nekrutenko, A. and Li, W.H. (2001) Transposable elements are found in a large number of human protein-coding genes. *Trends Genet.* 17, 619–621

3 Jordan, I.K. *et al.* (2003) Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet.* 19, 68–72

4 Kapitonov, V.V. and Jurka, J. (2005) RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. *PLoS Biol.* 3, e181

5 Burns, K.H. and Boeke, J.D. (2012) Human transposon tectonics. *Cell* 149, 740–752

6 Hedges, D.J. and Deininger, P.L. (2007) Inviting instability: transposable elements, double-strand breaks, and the maintenance of genome integrity. *Mutat. Res.* 616, 46–59

7 O'Donnell, K.A. and Burns, K.H. (2010) Mobilizing diversity: transposable element insertions in genetic variation and disease. *Mob. DNA* 1, 21

8 Shabalina, S.A. and Koonin, E.V. (2008) Origins and evolution of eukaryotic RNA interference. *Trends Ecol. Evol.* 23, 578–587

9 Ghildiyal, M. and Zamore, P.D. (2009) Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.* 10, 94–108

10 Castel, S.E. and Martienssen, R.A. (2013) RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat. Rev. Genet.* 14, 100–112

11 Bagijn, M.P. *et al.* (2012) Function, targets, and evolution of *Caenorhabditis elegans* piRNAs. *Science* 337, 574–578

12 Drinnenberg, I.A. *et al.* (2009) RNAi in budding yeast. *Science* 326, 544–550

13 Kaneko, H. *et al.* (2011) DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. *Nature* 471, 325–330

14 Tabara, H. *et al.* (1999) The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* 99, 123–132

15 Ketting, R.F. *et al.* (1999) Mut-7 of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell* 99, 133–141

16 Plasterk, R.H. (2002) RNA silencing: the genome's immune system. *Science* 296, 1263–1265

17 Sijen, T. and Plasterk, R.H. (2003) Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature* 426, 310–314

18 Yang, N. and Kazazian, H.H., Jr (2006) L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat. Struct. Mol. Biol.* 13, 763–771

19 Slotkin, R.K. *et al.* (2005) Heritable transposon silencing initiated by a naturally occurring transposon inverted duplication. *Nat. Genet.* 37, 641–644

20 Tam, O.H. *et al.* (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* 453, 534–538

21 Kim, S.H. and Walbot, V. (2003) Deletion derivatives of the MuDR regulatory transposon of maize encode antisense transcripts but are not dominant-negative regulators of mutator activities. *Plant Cell* 15, 2430–2447

22 Okamura, K. *et al.* (2008) Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in *Drosophila melanogaster*. *Nat. Struct. Mol. Biol.* 15, 998

23 Ghildiyal, M. *et al.* (2008) Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* 320, 1077–1081

24 Bzymek, M. and Lovett, S.T. (2001) Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* 98, 8319–8325

25 Ahmed, M. and Liang, P. (2012) Transposable elements are a significant contributor to tandem repeats in the human genome. *Comp. Funct. Genomics* 2012, 947089

26 Chicas, A. *et al.* (2004) RNAi-dependent and RNAi-independent mechanisms contribute to the silencing of RIPed sequences in *Neurospora crassa*. *Nucleic Acids Res.* 32, 4237–4243

27 Chang, S-S. *et al.* (2012) RNA interference pathways in fungi: mechanisms and functions. *Annu. Rev. Microbiol.* 66, 305–323

28 Romano, N. and Macino, G. (1992) Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. *Mol. Microbiol.* 6, 3343–3353

29 Lee, H.C. *et al.* (2010) The DNA/RNA-dependent RNA polymerase QDE-1 generates aberrant RNA and dsRNA for RNAi in a process requiring replication protein A and a DNA helicase. *PLoS Biol.* 8, e1000496

30 Cogoni, C. *et al.* (1996) Transgene silencing of the al-1 gene in vegetative cells of *Neurospora* is mediated by a cytoplasmic effector and does not depend on DNA–DNA interactions or DNA methylation. *EMBO J.* 15, 3153–3163

31 Nolan, T. *et al.* (2008) The RNA-dependent RNA polymerase essential for post-transcriptional gene silencing in *Neurospora crassa* interacts with replication protein A. *Nucleic Acids Res.* 36, 532–538

32 Zhang, Z. *et al.* (2013) Homologous recombination as a mechanism to recognize repetitive DNA sequences in an RNAi pathway. *Genes Dev.* 27, 145–150

33 Shiu, P.K. *et al.* (2001) Meiotic silencing by unpaired DNA. *Cell* 107, 905–916

34 Hammond, T.M. *et al.* (2013) Identification of small RNAs associated with meiotic silencing by unpaired DNA. *Genetics* 194, 279–284

35 Hammond, T.M. *et al.* (2013) Novel proteins required for meiotic silencing by unpaired DNA and siRNA generation in *Neurospora crassa*. *Genetics* 194, 91–100

36 Lee, D.W. *et al.* (2004) Properties of unpaired DNA required for efficient silencing in *Neurospora crassa*. *Genetics* 167, 131–150

37 Pratt, R.J. *et al.* (2004) DNA methylation affects meiotic trans-sensing, not meiotic silencing, in *Neurospora*. *Genetics* 168, 1925–1935

38 Maine, E.M. *et al.* (2005) EGO-1, a putative RNA-dependent RNA polymerase, is required for heterochromatin assembly on unpaired dna during *C. elegans* meiosis. *Curr. Biol.* 15, 1972–1978

39 She, X. *et al.* (2009) Regulation of heterochromatin assembly on unpaired chromosomes during *Caenorhabditis elegans* meiosis by components of a small RNA-mediated pathway. *PLoS Genet.* 5, e1000624

40 Dumesic, P.A. *et al.* (2013) Stalled spliceosomes are a signal for RNAi-mediated genome defense. *Cell* 152, 957–968

41 Wang, X. *et al.* (2010) Sex-induced silencing defends the genome of *Cryptococcus neoformans* via RNAi. *Genes Dev.* 24, 2566–2582

42 Chang, W. *et al.* (2013) BARE retrotransposons are translated and replicated via distinct RNA pools. *PLoS ONE* 8, e72270

43 Kupfer, D.M. *et al.* (2004) Introns and splicing elements of five diverse fungi. *Eukaryot. Cell* 3, 1088–1100

44 Irimia, M. *et al.* (2007) Coevolution of genomic intron number and splice sites. *Trends Genet.* 23, 321–325

45 Madhani, H.D. (2013) The frustrated gene: origins of eukaryotic gene expression. *Cell* 155, 744–749

46 Voinnet, O. (2008) Use, tolerance and avoidance of amplified RNA silencing by plants. *Trends Plant Sci.* 13, 317–328

47 Haag, J.R. and Pikaard, C.S. (2011) Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. *Nat. Rev. Mol. Cell Biol.* 12, 483–492

48 Nuthikattu, S. *et al.* (2013) The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21-22 nucleotide small interfering RNAs. *Plant Physiol.* 162, 116–131

49 Mari-Ordonez, A. *et al.* (2013) Reconstructing de novo silencing of an active plant retrotransposon. *Nat. Genet.* 45, 1029–1039

50 Vaucheret, H. *et al.* (1997) A transcriptionally active state is required for post-transcriptional silencing (cosuppression) of nitrate reductase host genes and transgenes. *Plant Cell* 9, 1495–1504

51 Dalmay, T. *et al.* (2000) 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

52 Luo, Z. and Chen, Z. (2007) Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by RDR6-mediated RNA silencing in *Arabidopsis*. *Plant Cell* 19, 943–958

53 Herr, A.J. *et al.* (2006) Defective RNA processing enhances RNA silencing and influences flowering of *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14994–15001

54 Gazzani, S. *et al.* (2004) A link between mRNA turnover and RNA interference in *Arabidopsis*. *Science* 306, 1046–1048

55 Gy, I. *et al.* (2007) *Arabidopsis* FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. *Plant Cell* 19, 3451–3461

56 Gregory, B.D. *et al.* (2008) A link between RNA metabolism and silencing affecting *Arabidopsis* development. *Dev. Cell* 14, 854–866

57 Axtell, M.J. *et al.* (2006) A two-hit trigger for siRNA biogenesis in plants. *Cell* 127, 565–577

58 Christie, M. *et al.* (2011) Intron splicing suppresses RNA silencing in *Arabidopsis*. *Plant J.* 68, 159–167

59 Moreno, A.B. *et al.* (2013) Cytoplasmic and nuclear quality control and turnover of single-stranded RNA modulate post-transcriptional gene silencing in plants. *Nucleic Acids Res.* 41, 4699–4708

60 Ishizu, H. *et al.* (2012) Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. *Genes Dev.* 26, 2361–2373

61 Luteijn, M.J. and Ketting, R.F. (2013) PIWI-interacting RNAs: from generation to transgenerational epigenetics. *Nat. Rev. Genet.* 14, 523–534

62 Brennecke, J. *et al.* (2007) Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128, 1089–1103

63 Senti, K.A. and Brennecke, J. (2010) The piRNA pathway: a fly's perspective on the guardian of the genome. *Trends Genet.* 26, 499–509

64 Gunawardane, L.S. *et al.* (2007) A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. *Science* 315, 1587–1590

65 Li, C. *et al.* (2009) Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell* 137, 509–521

66 Kelleher, E.S. and Barbash, D.A. (2013) Analysis of piRNA-mediated silencing of active TEs in *Drosophila melanogaster* suggests limits on the evolution of host genome defense. *Mol. Biol. Evol.* 30, 1816–1829

67 Girard, A. and Hannon, G.J. (2008) Conserved themes in small-RNA-mediated transposon control. *Trends Cell Biol.* 18, 136–148

68 Muerdter, F. *et al.* (2012) Production of artificial piRNAs in flies and mice. *RNA* 18, 42–52

69 Malone, C.D. *et al.* (2009) Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* 137, 522–535

70 Khurana, J.S. *et al.* (2011) Adaptation to P element transposon invasion in *Drosophila melanogaster*. *Cell* 147, 1551–1563

71 de Vanssay, A. *et al.* (2013) Profiles of piRNA abundances at emerging or established piRNA loci are determined by local DNA sequences. *RNA Biol.* 10, 1233–1239

72 Karpen, G.H. and Spradling, A.C. (1992) Analysis of subtelomeric heterochromatin in the *Drosophila* minichromosome Dp1187 by single P element insertional mutagenesis. *Genetics* 132, 737–753

73 Rangan, P. *et al.* (2011) piRNA production requires heterochromatin formation in *Drosophila*. *Curr. Biol.* 21, 1373–1379

74 Klattenhoff, C. *et al.* (2009) The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* 138, 1137–1149

75 Olovnikov, I. *et al.* (2013) De novo piRNA cluster formation in the *Drosophila* germ line triggered by transgenes containing a transcribed transposon fragment. *Nucleic Acids Res.* 41, 5757–5768

76 de Vanssay, A. *et al.* (2012) Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus. *Nature* 490, 112–115

77 Ruby, J.G. *et al.* (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* 127, 1193–1207

78 Gu, W. *et al.* (2012) CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as *C. elegans* piRNA precursors. *Cell* 151, 1488–1500

79 Billi, A.C. *et al.* (2013) A conserved upstream motif orchestrates autonomous, germline-enriched expression of *Caenorhabditis elegans* piRNAs. *PLoS Genet.* 9, e1003392

80 Das, P.P. *et al.* (2008) Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Mol. Cell* 31, 79–90

81 Ashe, A. *et al.* (2012) piRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*. *Cell* 150, 88–99

82 Lee, H-C. *et al.* (2012) *C. elegans* piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell* 150, 78–87

83 Shirayama, M. *et al.* (2012) piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. *Cell* 150, 65–77

84 Claycomb, J.M. *et al.* (2009) The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. *Cell* 139, 123–134

85 Wicker, T. *et al.* (2007) A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8, 973–982

86 Kazazian, H.H., Jr (2004) Mobile elements: drivers of genome evolution. *Science* 303, 1626–1632

87 Zhang, F. *et al.* (2012) UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* 151, 871–884