Small RNA-Mediated Quiescence of Transposable Elements in Animals

Kuniaki Saito1,2,* and Mikiko C. Siomi1,3,*

1Department of Molecular Biology, Keio University School of Medicine, Tokyo 160-8582, Japan
2JST, PRESTO, Saitama 332-0012, Japan
3JST, CREST, Saitama 332-0012, Japan
*Correspondence: saito@z6.keio.jp (K.S.), siomim@sc.itc.keio.ac.jp (M.C.S.)

DOI 10.1016/j.devcel.2010.10.011

Transposable elements (TEs) are major components of the intergenic regions of the genome. However, TE transposition has the potential to threaten the reproductive fitness of the organism; therefore, organisms have evolved specialized molecular systems to sense and repress the expression of TEs to stop them from jumping to other genomic loci. Emerging evidence suggests that Argonaute proteins play a critical role in this process, in collaboration with two types of cellular small RNAs: PIWI-interacting RNAs (piRNAs) of the germline and endogenous small interfering RNAs (endo-siRNAs) of the soma, both of which are transcribed from TEs themselves.

Introduction

Transposable elements (TEs) were discovered in maize by Barbara McClintock in the 1940s. Since then, various TEs and TE-like sequences have been identified in many species. World-wide genome projects have revealed that TEs are actually major genomic components. For example, TEs comprise 45% of the human genome (Lander et al., 2001) and 15%–22% of the genome of Drosophila melanogaster (Kapitonov and Jurka, 2000; Biémont and Vieira, 2006). TEs can be separated into Class 1 and Class 2, based on their structures and modes of integration (Gogvadze and Buzdin, 2009). Class 1 comprises retrotransposons and retrotransposon-like elements, such as LINEs and SINEs, which replicate via transcribed mRNA intermediates that are converted back into DNA through multiple rounds of reverse transcription. The retrotransposon RNAs thus transpose into the genome by a mechanism analogous to “copy and paste.” Class 2 TEs consist of DNA transposons only. Unlike retrotransposons, DNA transposons move directly from place to place through a “cut and paste” process.

The significance and biological implications of TEs and TE-like sequences have been thoroughly reviewed (Goodier and Kazazian, 2008; Gogvadze and Buzdin, 2009). TEs can occasionally play beneficial roles in organisms. Transposition of LINE-1 elements contributes to genetic diversity in neuronal progenitor cells during the development of the nervous system (Coulal et al., 2009). LINE-1 expression may contribute to the establishment of X inactivation in ES cells (Chow et al., 2010). The human endogenous retrovirus (HERV-W) encodes an envelope protein named syncytin, which plays important roles in trophoblast cell fusion and placental morphogenesis (Mi et al., 2000). The telomeric ends of linear chromosomes are protected by TEs in Drosophila (Pardue et al., 2005). And in many organisms, heterochromatic centromeres, which are crucial for the segregation of chromosomes during cell division, consist of repeats and TEs (e.g., Volpe et al., 2002). However, the mobilization of TEs is often considered deleterious because it leads to genomic structural rearrangements, such as deletions, duplications, and inversions. Organisms have thus developed effective strategies that constrain TE transposition.

One such mechanism is mediated by small RNAs, in a process that is highly analogous to RNA interference (RNAi) or RNA silencing (Ghildiyal and Zamore, 2009; Kim et al., 2009; Malone and Hannon, 2009; Siomi et al., 2008; Siomi and Siomi, 2009; Thomson and Lin, 2009). RNA silencing is triggered by small RNAs 20–30 nucleotides (nt) in length, which are often, if not universally, processed from double-stranded RNAs (dsRNAs) by the enzyme Dicer. The small RNAs are then loaded onto members of the Argonaute family of proteins. Argonaute proteins bind target transcripts by means of base pairing between the small RNA and the target; this interaction confers target specificity on Argonaute regulation of transcription, transcript stability, and/or translation. Accumulating evidence indicates that TEs are silenced by similar (transcriptional or posttranscriptional) mechanisms, in which small RNAs and Argonaute proteins play key roles.

The numbers and types of Argonaute family members differ between species (Cerutti et al., 2000). In the paradigmatic example of Drosophila, five Argonaute proteins are expressed. Among these, the AGO proteins (AGO1 and AGO2) are ubiquitous, whereas the others (AGO3, Aubergine [Aub], and Piwi) are gonad specific (Williams and Rubin, 2002) and are generically referred to as PIWI proteins (Carmell et al., 2002). PIWI-interacting small RNAs (piRNAs) are mainly derived from intergenic loci, thus forming small interfering RNAs (siRNAs) of the germline and endogenous small interfering RNAs (endo-siRNAs) of the soma, both of which are transcribed from TEs themselves.
that contain TE fragments and repeats, and are now referred to as endogenous small interfering RNAs (endo-siRNAs or esiRNAs). AGO/endo-siRNA components and PIWI/piRNA components are highly conserved in vertebrates and invertebrates, and they all function in TE silencing.

In this review, we discuss how piRNAs and endo-siRNAs are produced and how these small RNAs silence TEs, focusing on studies in Drosophila melanogaster and then comparing and contrasting with other organisms such as mice, frogs, and zebrafish.

**PIWI Proteins in Drosophila**

The piwi gene was originally discovered as an essential factor for germline stem cell self-renewal in Drosophila (Lin and Spradling, 1997; Cox et al., 1998). Further investigation showed that mutations in piwi cause hyperactivation of retrotransposons (Vagin et al., 2006; Klenov et al., 2007). Likewise, the paralogous aub gene is necessary for pole cell formation (Harris and Macdonald, 2001) and transposon silencing (Vagin et al., 2006), and PIWI genes are commonly required for female fertility (Cox et al., 1998; Schmidt et al., 1999; Li et al., 2009).

The PIWI proteins show different subcellular localization and expression patterns, suggesting distinct and discrete roles. Among the PIWI proteins, only Piwi is localized in the nucleus (Cox et al., 2000), and its primary function in TE silencing appears to be in the nucleus (Saito et al., 2010). In ovariies, Piwi can be detected in both somatic and germ cells (Cox et al., 2000; Saito et al., 2006; Brennecke et al., 2007; Nishida et al., 2007). In the testes, however, Piwi is barely detectable in germ cells but is strongly expressed in somatic niche cells, which associate with and maintain the germline stem cells. Aub and AGO3 are expressed in the germ cells of both males and females, but are absent in gonadal somatic cells. Unlike Piwi, Aub and AGO3 are exclusively detected in the cytoplasm at steady-state levels and accumulate in the nuage (Brennecke et al., 2007; Gunawardane et al., 2007; Nishida et al., 2007; Li et al., 2009), an electron-dense perinuclear structure characteristic of germline cells in diverse animals (al-Mukhtar and Webb, 1971; Eddy, 1975). These observations indicate that the PIWI proteins possess distinct functions and utilize different systems for expression and subcellular localization.

**piRNAs in Drosophila**

The first direct link between PIWI proteins and TEs was the discovery of piRNAs, defined as small RNAs associated with PIWI proteins in any organism. Typically, piRNAs are 23–30 nt long. Sequencing of piRNAs in Drosophila has revealed that they mainly originate from TEs and TE-related genomic elements (Saito et al., 2006; Brennecke et al., 2007; Yin and Lin, 2007). Whereas miRNAs are processed by Dicer from double-stranded primary precursors (Lee et al., 2004), piRNAs are produced in a Dicer-independent manner from single-stranded precursors (Vagin et al., 2006). piRNAs are considerably more heterogeneous than miRNAs. More than 70% of piRNAs have been cloned only once, indicating that the piRNA population is very complex, in clear contrast to the miRNA population, which consists of approximately 170 species in Drosophila (Brennecke et al., 2007; Yin and Lin, 2007). Mapping of piRNAs on the Drosophila genome revealed that several hundred genomic regions could be referred to as piRNA clusters. These clusters contribute to the generation of piRNAs, most of which encode TEs and TE remnants and are mainly localized in the pericentromeric and subtelomeric heterochromatin regions.

The piRNA clusters fall into two groups: dual-strand and uni-strand clusters (Figure 1). The dual-strand clusters produce piRNAs from both genomic strands, with the representative being 42AB on chromosome 2. The uni-strand clusters are transcribed in only one direction and are represented by the flamenco (flam) locus on chromosome X. The flam locus was first discovered as a regulator of the activity of the retroviral gypsy, idefix, and ZAM elements (Pelisson et al., 1994; Desset et al., 2008). Antisense-oriented copies of these retro-elements in the flam locus result in the production of mainly antisense
piRNAs (Figure 2; Brennecke et al., 2007). Although the mechanisms of transcription from the uni-strand piRNA clusters remain unclear, recent chromatin immunoprecipitation analysis revealed that the heterochromatin protein 1 (HP1) homolog Rhino associates with DNA from this locus and is required for its transcription. The production of 42AB-derived piRNAs occurs in germline cells in ovaries. These piRNAs are later amplified by the ping-pong cycle.

(B) The flam piRNA locus on chromosome X gives rise to transcripts in one direction. The flam-derived piRNAs are loaded onto Piwi and are not amplified. These processes occur only in the ovarian somatic cells.

(C) Amplification of piRNAs derived from the primary processing pathway. The piRNAs deposited from the mother to the offspring are also subject to the ping-pong cycle, specifically in the germline cells of the ovary.

(D) Aub and Piwi associated with piRNAs accumulate at the posterior pole of the oocytes in egg chambers. In developing embryos, the Aub- and Piwi-piRNA complexes are incorporated into the primordial germ cells. AGO3 does not seem to be passed from the mother to the offspring.

In the ping-pong cycle, the Aub complex associated with an antisense piRNA specifically cleaves TE transcripts in the sense orientation. Therefore, the ping-pong cycle not only produces piRNAs but also degrades TE mRNAs at the same time. This model was confirmed by a recent study on ago3 mutant flies (Li et al., 2009). In ago3 mutants, the number of Aub-bound piRNAs was markedly decreased and their antisense strand bias was weakened from 71% to 41%. Some TEs were derepressed in ago3 mutants, as in aub mutants. These observations support the notion that TE silencing and piRNA biogenesis are coupled (Li et al., 2009). Thus, the seeds to initiate the ping-pong cycle are supplied by the primary processed or maternally deposited piRNAs (see below).

**Figure 2. Models for piRNA Biogenesis in Drosophila Ovaries**

(A) The piRNA 42AB locus on chromosome 2 is transcribed in both directions and the sense and antisense transcripts serve as piRNA precursors. Rhino associates with DNA from this locus and is required for its transcription. The production of 42AB-derived piRNAs occurs in germline cells in ovaries. These piRNAs are later amplified by the ping-pong cycle.

(B) The flam piRNA locus on chromosome X gives rise to transcripts in one direction. The flam-derived piRNAs are loaded onto Piwi and are not amplified. These processes occur only in the ovarian somatic cells.

(C) Amplification of piRNAs derived from the primary processing pathway. The piRNAs deposited from the mother to the offspring are also subject to the ping-pong cycle, specifically in the germline cells of the ovary.

(D) Aub and Piwi associated with piRNAs accumulate at the posterior pole of the oocytes in egg chambers. In developing embryos, the Aub- and Piwi-piRNA complexes are incorporated into the primordial germ cells. AGO3 does not seem to be passed from the mother to the offspring.

Bias for adenosine (A) at the tenth nucleotide from their 5’ end (10-A). In fact, the first 10 nucleotides of antisense piRNAs frequently overlap with the sense piRNAs. These findings (including the 1-U/10-A relationship) are referred to as the ping-pong signature; they imply that AGO3-associated piRNAs may pair with Aub- (and Piwi-) associated piRNAs through their first 10 nucleotides (Brennecke et al., 2007; Gunawardane et al., 2007).

Moreover, recombinant Piwi, Aub, and AGO3 cleave complementary target RNAs between nucleotides 10 and 11 of small guiding RNAs (Gunawardane et al., 2007), suggesting an intriguing model in which PIWI proteins generate 5’ ends of piRNAs and amplify piRNAs from longer (e.g., conventional) transcripts in a Slicer-dependent manner (Figure 2). This model is called the ping-pong cycle or amplification loop and explains why piRNAs are generated in a Dicer-independent manner, in contrast to siRNAs and miRNAs (Vagin et al., 2006).

In the ping-pong cycle, the Aub complex associated with an antisense piRNA specifically cleaves TE transcripts in the sense orientation. Therefore, the ping-pong cycle not only produces piRNAs but also degrades TE mRNAs at the same time. This model was confirmed by a recent study on ago3 mutant flies (Li et al., 2009). In ago3 mutants, the number of Aub-bound piRNAs was markedly decreased and their antisense strand bias was weakened from 71% to 41%. Some TEs were derepressed in ago3 mutants, as in aub mutants. These observations support the notion that TE silencing and piRNA biogenesis are coupled (Li et al., 2009). Thus, the seeds to initiate the ping-pong cycle are supplied by the primary processed or maternally deposited piRNAs (see below).
piRNA Biogenesis in Drosophila: The Primary Processing Pathway

Although Aub- and AGO3-associated piRNAs are mainly produced by the ping-pong cycle, Piwi-associated piRNAs show few ping-pong signatures, suggesting a ping-pong-independent piRNA biogenesis pathway (Brennecke et al., 2007). Piwi is required for the silencing of a subset of TEs in gonadal somatic cells. In fact, the majority of Piwi-associated piRNAs in ovaries are derived from the flam uni-strand cluster, which is expressed only in the soma (Brennecke et al., 2007). Further studies suggested the idea that flam piRNAs are loaded onto Piwi by a mechanism distinct from the amplification cycle (Malone et al., 2009; Li et al., 2009). In the ovarian somatic cell (OSC) line, which contains only somatic cells of fly ovaries, Piwi is expressed, but Aub and AGO3 are not, suggesting that the piRNAs in OSCs are produced by the primary processing pathway but are not amplified further by the amplification cycle (Saito et al., 2009). In OSCs, Piwi is associated with abundant pathway but are not amplified further by the amplification cycle (Malone et al., 2009; Li et al., 2009). In the ovarian somatic cell (OSC) line, which contains only somatic cells of fly ovaries, Piwi is expressed, but Aub and AGO3 are not, suggesting that the piRNAs in OSCs are produced by the primary processing pathway but are not amplified further by the amplification cycle (Saito et al., 2009). In OSCs, Piwi is associated with abundant piRNAs, including flam-derived piRNAs, confirming that flam-piRNAs are produced only by the primary pathway (Figure 2).

Mechanisms of PIWI-piRNA-Mediated Silencing

Aub and AGO3 show Slicer activity (the RNaseH-like endonuclease activity of their PIWI domains) in vitro (Gunawardane et al., 2007) and localize in the cytoplasm in vivo (Nishida et al., 2007; Li et al., 2009). Therefore, Aub-piRNA and AGO3-piRNA complexes most probably silence TEs posttranscriptionally by cleaving their transcripts. The Aub-piRNA complex may also induce mRNA decay to silence target genes such as nanos, the embryonic posterior morphogen, by recruiting deadenylation enzyme CCR4 and the RNA-binding protein Smaug to the target mRNA (Rouget et al., 2010). How then do Piwi and the somatic primary piRNAs act to silence TEs? One model is Slicer-mediated RNA cleavage, because recombinant Piwi produced from E. coli possesses Slicer activity (Saito et al., 2006). Another feasible model is that Piwi is involved in heterochromatic gene silencing. This model was suggested from the observations that Piwi influences position-effect variegation (PEV) (Pal-Bhadra et al., 2004). PEV is a probabilistic variegation in gene silencing that depends on the genomic locus; for example, if an euchromatic gene were relocated in close proximity to a heterochromatin through a genetic rearrangement such as inversion or translocation, there is a high chance that this gene will be silenced by the heterochromatin. It has also been shown that HP1a and Piwi associate specifically with heterochromatin-containing TEs, such as 1360 and F element. Piwi may direct HP1a on heterochromatin to silence TEs in early-stage embryos (Pal-Bhadra et al., 2004). However, piwi mutations resulted in an increase of HP1 association with the 3R-TAS subtelomeric region in ovaries and subsequently increased transcriptional silencing, suggesting that Piwi might act as an activator of TEs at 3R-TAS (Yin and Lin, 2007). In addition, a recent study showed that HP1 recruitment on heterochromatin does not depend on Piwi (Moshkovich and Lei, 2010). The mechanism by which the piRNA pathway contributes to heterochromatin regulation thus remains controversial.

Figure 3. Phylogenetic Tree of Piwi Proteins in Fleis, Mice, Frogs, and Zebrafish

In mice, three PIWI proteins, mouse Piwi (Miwi), Miwi-like (Mili), and Mili2 (Kuramochi-Miyagawa et al., 2001, 2004; Deng and Lin, 2002; Carmell et al., 2007), have been characterized, and loss of each leads to derepression of TEs and spermatogenic arrest (Kuramochi-Miyagawa et al., 2001, 2004; Aravin et al., 2007a; Carmell et al., 2007). Mili expression begins in primordial germ cells at embryonic day (E) 12.5 and continues throughout adult spermatogenesis, up to the round spermatid stage (Kuramochi-Miyagawa et al., 2004). Mili2 is transiently expressed in embryonic gonocytes from E15.5, with its expression decreasing after birth, prior to meiosis (Aravin et al., 2008, Kuramochi-Miyagawa et al., 2008). Mili is specific to meiotic and postmeiotic spermatogenesis of the adult; it is expressed from pachytene to round spermatid stages (Kuramochi-Miyagawa et al., 2001; Deng and Lin, 2002). piRNAs isolated from PIWI proteins at each developmental stage can be divided into three groups: prenatal piRNAs, prepachytene piRNAs, and pachytene piRNAs. Mili and Mili2 associate with prenatal piRNAs, largely corresponding to TEs in fetal gonads (Aravin et al., 2007a, 2008; Kuramochi-Miyagawa et al., 2008). At the prepachytene stage, Mili, the only Piwi protein expressed at this stage, associates with prepachytene piRNAs containing both TE and non-TE sequences (Aravin et al., 2007a, 2008; Kuramochi-Miyagawa et al., 2008). Mili associates with pachytene piRNAs derived from non-TE elements, with only 12%–17% of pachytene piRNAs corresponding to TEs (Aravin et al., 2008; Girard et al., 2006; Watanabe et al., 2006). Consistent with the distinct temporal profiles of each paralog and its associated piRNAs, the loss of zebrafish Mili (Zili) leads to the expression of TEs in the developing gonad and early failure of germ cell development, while the loss of zebrafish Miwi (Ziwi) causes apoptosis later in the germline, at the premeiotic stage. In zebrafish and Xenopus, the majority of piRNAs correspond to TEs (Houwing et al., 2008; Kirino et al., 2009; Lau et al., 2009), similar to piRNAs in Drosophila. Unlike Drosophila and mice, the

piRNA-Mediated TE Repression in Vertebrates

Piwi proteins are conserved in vertebrates (Figure 3) and have conserved functions in TE silencing and germ cell development.
Xenopus genome contains numerous DNA transposons (Jurka et al., 2005), and thus the proportion of piRNAs targeting DNA transposons is high (Lau et al., 2009). These studies indicate that piRNA content differs with development and the genomic context.

Accumulating evidence has demonstrated that piRNA biogenesis occurs through the primary piRNA processing and the amplification loop in mice, frogs, and zebrafish (Aravin et al., 2007b; Ghildiyal and Zamore, 2009; Houwing et al., 2007; Kim et al., 2009; Malone and Hannon, 2009). In gonocytes of fetal mice, where both Mili and Miwi2 are expressed, the ping-pong signature is obvious between Mili-associating sense 10-A piRNAs and Miwi2-associating antisense 1-U piRNAs. Moreover, the ping-pong signature is found in zebrafish and Xenopus (Kirino et al., 2009; Lau et al., 2009; Houwing et al., 2008), which strongly supports the existence of conserved mechanisms of biogenesis in vertebrates. Interestingly, piRNA amplification is not limited to cycles between distinct PIWI proteins. At the pre-pachytene stage, Mili shows a tendency to associate with both 1-U and 10-A piRNAs during a certain period (Aravin et al., 2007a). Similar to Mili, Ziwí may be able to amplify piRNAs without Zili expression (Houwing et al., 2008), suggesting that PIWI proteins self-amplify piRNAs via the amplification loop. Taken together, these studies indicate that, as in Drosophila, TE mRNAs in vertebrates are degraded by PIWI proteins via a ping-pong cycle.

Intriguingly, Mili and Miwi2 mutant mice exhibit loss of DNA methylation in TEs. Further studies have clarified that the Mili- and Miwi2-piRNA complexes are involved in de novo DNA methylation of TE s (Aravin et al., 2008; Kuramochi-Miyagawa et al., 2008). In general, CpG DNA methylation causes stable repression of transcription (Bourc’his and Bestor, 2004). However, it remains unclear how piRNAs are involved in DNA methylation.

**Scaffolds and Modifiers of piRNA Pathways**

Tudor (Tud) was identified as a nuage component required for piRNA biogenesis in vertebrates. Interestingly, piRNA biogenesis is apparently unnecessary for producing primary piRNAs spn-E (Lim and Kai, 2007), zucchini (zuc) (Pane et al., 2007), spn-E (Vagin et al., 2006), and maelstrom (mael) (Findley et al., 2003; Lim and Kai, 2007) also affect piRNA accumulation, and TEs are derepressed in the ovaries of mutants for all these genes (Table 1; Vagin et al., 2006; Lim and Kai, 2007; Pane et al., 2007). Deep-sequencing of piRNAs in these mutants has shown distinct requirements: for example, spn-E mutations reduce not only piRNA amplification but also the production of primary piRNAs originating from cluster 2, one of ping-pong-independent uni-strand clusters (Malone et al., 2009). However, spn-E is apparently unnecessary for producing primary piRNAs from flam in ovar y (Malone et al., 2009). On the other hand, zuc, which encodes a putative nuclease, is necessary for somatic primary piRNA processing but not for the ping-pong cycle (Haase et al., 2010; Malone et al., 2009; Olivieri et al., 2010; Saito et al., 2010). These observations suggest that piRNA pathway components perform different and specific roles in piRNA biogenesis in Drosophila. Krimper and Spn-E contain Tud domains; however, whether they interact with PIWI proteins through sDMAs remain unknown.

Krimper, Zuc, and Mael localize to the nuage, suggesting that nuage might be an important site of piRNA processing in germ cells. Recent studies indicate that Yb bodies consisting of Yb and Armitage (Armi) are the sites for somatic primary piRNA processing (Olivieri et al., 2010; Saito et al., 2010). Both Armi and Yb contain an RNA helicase domain and therefore are implicated in RNA metabolism. In fact, mutations in Armi cause a severe decrease in piRNA accumulation in ovaries (Vagin et al., 2006). In Armi- or Yb-depleted cells of the ovarian soma, Piwi is not associated with piRNAs and mislocalizes to the cytoplasm, resulting in a general loss of piRNAs, and suggesting that Armi and Yb are piRNA biogenesis modulators. Yb contains a Tud domain; however, its function remains undetermined. In mice, piRNA pathway components exhibit differential localization. Miwi2/Tdrd9/Mael colocalizes with the components of Processing bodies referred to as piP-bodies, whereas Miili/Tdrd1/MVH/GASZ associates with the intermitochondrial cement or pi-bodies (Aravin et al., 2009; Ma et al., 2009; Shoji et al., 2009; Kuramochi-Miyagawa et al., 2010). piP-body formation depends on pi-bodies, and Miwi2 fails to associate with piRNAs when pi-bodies are not formed. On the other hand, pi-body formation is not affected by piP-body disruption, suggesting a possible progression of piRNA processing from pi-body to piP-body. Thus, both piP-bodies and pi-bodies are implicated as the sites for piRNA biogenesis in mice. It will be interesting
to determine what the specificity of these structures might be, for example, whether primary and ping-pong processes are differentially localized between them.

**Germline Transmission and Hybrid Dysgenesis**

The complex relationship between nuage architecture, germline development, and piRNA pathways suggests that the broader implications of piRNA-mediated germline integrity control are worth considering. During ovary development, Aub and Piwi accumulate in specialized cytoplasm at the posterior pole of the oocyte, where germ cell determinants are localized (Rongo and Lehmann, 1996). Maternally deposited Aub and Piwi are thus incorporated into the primordial germ cells that inherit this “posterior pole plasm” and are transmitted to the next generation directly. In contrast, AGO3 does not seem to be passed on from mother to offspring, suggesting that antisense piRNAs may be selectively inherited, whereas sense piRNAs are not. Brennecke et al. revealed that piRNAs are truly transmitted from mother to offspring (Brennecke et al., 2008). This maternal inheritance of piRNAs was shown to explain a long-standing but poorly understood phenomenon of “hybrid dysgenesis,” in which the introduction of new genetic material via the male germline induces genetic instability in female progeny (Kidwell and Kidwell, 1976). The relevant genetic material was known to include TEs for some time (e.g., Rubin et al., 1982), and a maternally transmitted epigenetic suppressor, called a cytotype, had

| Table 1. Features of TE-Derived Small RNAs in Animals |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| **Small RNA**   | Mouse           | Frog            | Zebrafish       | Fly             | Function/Motif  |
| Size (nt)       | 25–30           | 26–31           | 27–30           | 24–30           | reviewed in Klattenhoff and Theurkauf, 2008; Siomi and Kuramochi-Miyagawa, 2009; Ghildiyal and Zamore, 2009; Kim et al., 2009; Malone and Hannan, 2009; Thomson and Lin, 2009; Lau, 2010; Senti and Brennecke, 2010 |
| Nucleotide preference | 5’U; 10th A | 5’U; 10th A | 5’U; 10th A | 5’U; 10th A | see entry above |
| Associated Argonaute protein | Miwi; Mili; Miwi2 | Xiwi1a,1b; Xili; Xiwi2 | Ziwi; Zili | Piwi; Aub; AGO3 | see entry above |
| Amplification of small RNA (5’U-10thA) | Mili-Miwi2; Mili-Mili | not determined | Ziwi-Zili | Piwi-AGO3; Aub-AGO3 | see entry above |
| Other factors required for small RNA-mediated TE silencing | Tdrd-1 | Tdrd-1* | Tdrd-1* | Tudor, MYND | Vagin et al., 2009; Reuter et al., 2009 |
| Tdrd-9 | Tdrd-9* | Tdrd-9* | Spindle E | Tudor, DEXDc | Shoji et al., 2009; Vagin et al., 2006 |
| Maelstrom | Maelstrom* | Maelstrom* | Maelstrom | HMG | Soper et al., 2008; Lim and Kai, 2007 |
| Tdrd-6* | Tdrd-6* | Tdrd-6* | Krimper | Tudor | Lim and Kai, 2007 |
| Mov10I1 | Mov10I1* | Mov10I1* | Armitage | Helicase? | Zheng et al., 2010; Frost et al., 2010; Vagin et al., 2006 |
| Tdrd-7*/Tdrd-5* | Tdrd-7*/Tdrd-5* | Tdrd-7*/Tdrd-5* | Tejas | Tudor | Patil and Kai, 2010 |
| MVH | DDX4* | Vasa* | Vasa | DEAD | Kuramochi-Miyagawa et al., 2010; Lim and Kai, 2007 |
| PLD6* | PLD6* | PLD6* | Zucchini | Nuclease? | Pane et al., 2007 |
| Squash | Nuclease? | Pane et al., 2007 |
| PRMT5* | PRMT5* | PRMT5* | diPRMT5 | Methyltransferase | Kirino et al., 2009 |
| Fs(1)Yb | Tudor | Olivieri et al., 2010; Saito et al., 2010 |
| GASZ | GASZ* | GASZ* | CG2183* | ANK repeat | Ma et al., 2009 |

Genes marked with an asterisk show that the involvement on piRNA-mediated TE silencing has not been confirmed experimentally.
been proposed (Engels, 1979), but it took several decades before piRNA characterization progressed to a point where the realization could be made that the cytotype corresponds to maternally inherited PIWI/piRNA complexes (Figure 2; Brennecke et al., 2008). These results extended the paradigm of piRNA-dependent germine integrity, to provide molecular insight into the epigenetics of incompatibility between diverging animal subpopulations.

Interestingly, fertility is enhanced by two environmental factors, age and high temperature (Bucheton, 1978). Heat shock protein 90 (Hsp90), a chaperone involved in several cellular and developmental pathways, has been implicated in TE control mediated by piRNAs (Specchia et al., 2010). Mutations affecting Hsp90, or treatment with the specific Hsp90 inhibitor geldanamycin, decreased piRNA biogenesis and increased the expression levels and the mobilization rates for all types of TEs. These observations suggest that Hsp90 may act as a piRNA factor and couple the control of TEs to environmental inputs.

### Endogenous Small Interfering RNAs: Biogenesis

Recent studies have indicated that TE silencing by small RNAs is not a gonad-specific event. Rather, TE transcripts are converted into a subset of small RNAs in nongonadal somatic cells, which display different characteristics from those of piRNAs. These small RNAs are designated endo-siRNAs or esiRNAs. In *Drosophila*, endo-siRNAs specifically bind to AGO2, the effector of RNAi (Figure 4). AGO2 was originally demonstrated to associate with exogenous siRNAs (exo-siRNAs) derived from external dsRNAs in RNAi. The exo-siRNAs produced from viral dsRNAs have been shown to serve as a defense mechanism against viral infection (Galiana-Arnoux et al., 2006; van Rij et al., 2006; Wang et al., 2006). However, the recent discovery of endo-siRNAs clarified that AGO2 does not always wait for “external” siRNAs, but rather plays particular roles in cellular RNAi with endo-siRNAs.

Like piRNAs, endo-siRNAs are mainly derived from TEs, heterochromatic regions, and intergenic elements of the genome. The endo-siRNAs occasionally arise from protein-coding genes transcribed from overlapping 3′ untranslated regions (UTRs) or long hairpin RNA genes (Czech et al., 2008; Ghildiyal et al., 2008; Kawamura et al., 2008; Okamura et al., 2008). Sequencing of small RNAs in *Drosophila* embryos, ovaries, and the S2 cell line has revealed that endo-siRNAs are 21 nt long and broadly map across the locations of most TEs in both the sense and antisense orientations (Figure 4). However, they do not show an obvious nucleotide preference at their 5′ end. These characteristics, together with the observation that endo-siRNAs show a typical phasing pattern when mapped on the genome, suggest Dicer2 involvement in the processing pathway. In fact, loss of *dicer2* function abolished the production of endo-siRNAs in vivo (Czech et al., 2008; Kawamura et al., 2008; Okamura et al., 2008). It had been previously shown that Dicer1 partners with the dsRNA-binding partner Locquacious (Loqs) to generate miRNAs (Förstemann et al., 2005; Jiang et al., 2005; Saito et al., 2005), whereas Dicer2 uses the related protein R2D2 during exo-siRNA production (Liu et al., 2003). It therefore came as a surprise to find that endo-siRNA biogenesis occurs by an unusual combination of Dicer2 and Loqs (Czech et al., 2008; Okamura et al., 2008). This confusion was resolved when it was discovered that the loqs gene gives rise to four isoforms, *loqs-PA-loqs-PD* ( Förstemann et al., 2005; Hartig et al., 2009; Zhou et al., 2009; Miyoshi et al., 2010), such that Loqs-PB and Loqs-PD are involved in the miRNA and endo-siRNA processing pathways, respectively (Hartig et al., 2009; Miyoshi et al., 2010; Zhou et al., 2009). Indeed, Loqs-PB associates with Dicer1, whereas Loqs-PD associates with Dicer2. These specific associations of Loqs isoforms with Dicer proteins most probably confer the ability of Dicer proteins to discriminate small RNA precursor substrates. The involvement of Loqs-PD in TE silencing might be modest, because endo-siRNA production depends less on Loqs than Dicer2 and AGO2 (Chung et al., 2008; Czech et al., 2008).

### Endo-siRNA-Mediated TE Silencing

In mouse oocytes, the endo-siRNA pathway plays important roles in TE silencing (Tam et al., 2008; Watanabe et al., 2008; Murchison et al., 2007). However, the effects of endo-siRNAs are subtler in *Drosophila*, where TEs are derepressed (1.5- to 9-fold) when *dicer2* or *ago2* gene functions are lost, but only a small subset of specific TEs are affected, and the mutant flies are viable and fertile (Chung et al., 2008; Czech et al., 2008; Ghildiyal et al., 2008; Kawamura et al., 2008). It is plausible that the absence of strong phenotypes may be due in part to redundancy with piRNAs: Endo-siRNA factors are expressed in germine cells where piRNA-mediated silencing occurs, and sequence analysis has shown that AGO2-associated endo-siRNAs often map to developmentally critical loci. In this case, the TE silencing is mediated by a distinct set of endo-siRNAs that might not necessarily reflect piRNA activity.
piRNA clusters. Thus, these regions of the genome are able to produce both endo-siRNAs and piRNAs. This raises important questions about how sense and antisense strands from these piRNA clusters form dsRNAs to produce endo-siRNAs and how they stay single stranded as piRNA precursors.

Conclusions

Despite their sequence diversity, piRNAs and endo-siRNAs do not disturb the expression of protein-coding genes. Thus, there must be strict cellular systems that discriminate TEs from protein-coding genes. How does this operate in vivo? One possibility is that TEs are identified at the transcription stage. Rhino contributes specifically to piRNA production from dual-strand piRNA clusters but not from uni-strand clusters, suggesting that Rhino can discriminate between the two clusters. Because Rhino is an HP1 homolog, it is most likely that Rhino functions at the transcriptional level, although there is no direct evidence to support this. The second possibility is that TEs are identified at the posttranscriptional level. TEs are transcribed by RNA polymerase II and are poly(A)-tailed, suggesting that piRNA precursors and mRNAs may contain similar structural characteristics. This would make it difficult to discriminate between TE transcripts and protein-coding transcripts. However, there should be cellular mechanisms that can separate the two species. It would be of interest to identify these mechanisms in Drosophila and in other species, such as mice.

ACKNOWLEDGMENTS

We thank H. Siomi and other members of the Siomi laboratory for discussions and comments on the manuscript. This work was supported by Ministry of Education, Culture, Sports, Science, and Technology of Japan grants to K.S. and M.C.S. K.S. is supported by Precursory Research for Embryonic Science and Technology from Japan Science and Technology Agency. M.C.S. is supported by Core Research for Evolutional Science and Technology from the Japan Science and Technology Agency.

REFERENCES


