The epigenetic regulation of transposable elements by PIWI-interacting RNAs in Drosophila

Kuniaki Saito*

Department of Molecular Biology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan

(Received 15 January 2013, accepted 26 February 2013)

A mechanism is required to repress the expression and transposition of transposable elements (TEs) to ensure the stable inheritance of genomic information. Accumulating evidence indicates that small non-coding RNAs are important regulators of TEs. Among small non-coding RNAs, PIWI-interacting RNAs (piRNAs) serve as guide molecules for recognizing and silencing numerous TEs and work in collaboration with PIWI subfamily proteins in gonadal cells. Disruption of the piRNA pathway correlates with loss of proper genomic organization, gene expression control and fertility. Moreover, recent studies on the molecular mechanisms of piRNA biogenesis and on piRNA function have shown that piRNAs act as maternally inherited genic elements, transferring information about repressed TEs to progeny. These findings enable a molecular explanation of mysterious epigenetic phenomena, such as hybrid dysgenesis and TE adaptation with age. Here, I review our current knowledge of piRNAs derived from biochemical and genetic studies and discuss how small RNAs are utilized to maintain genome organization and to provide non-DNA genetic information. I mainly focus on Drosophila but also discuss comparisons with other species.

Key words: Drosophila, hybrid dysgenesis, piRNAs, PIWI proteins, transposable elements

INTRODUCTION

In eukaryotes, large proportions of genomes consist of various transposable elements (TEs) and their remnant sequences. For instance, 45% of the human genome and 15–22% of the Drosophila genome contain TE sequences (Biomet and Vieira, 2006; Lander et al., 2001). A major class of TEs is the retrotransposons, which are abundant because of their “copy and paste” mode of amplification (Finnegan, 2012; Gogvadze and Buzdin, 2009). When TEs are inserted into a genomic location, gene expression patterns may become altered, which may compromise cell viability (Callinan and Batzer, 2006; Hancks and Kazazian, 2012). Some TEs seem to play beneficial roles in regulating gene expression as distal enhancers, chromatin boundaries and binding sites of many transcription factors (see the review by Ichiyanagi, 2013), while others TEs are recognized and repressed to maintain genome integrity. In various species, TEs are remarkably compartmentalized in the heterochromatin, which is accomplished by histone modification and heterochromatin proteins (Levin and Moran, 2011; Slotkin and Martienssen, 2007). In Schizosaccharomyces pombe (S. pombe), a small RNA-mediated system is involved in heterochromatin formation at the centromere region where small RNAs are utilized to recognize the centromere region (Moazed, 2009). These studies reveal an unsuspected role for small RNAs in the regulation of chromatin structure and genome stability.

Until recently, little was known about how Drosophila and vertebrate TEs are remarkably compartmentalized in heterochromatin and specifically silenced in cells regardless of their heterogeneity and diversity. However, it is now proposed that in germ cells, PIWI-interacting RNAs (piRNAs) act as guides for recognition of TEs by PIWI subfamily proteins (Ishizu et al., 2012; Juliano et al., 2011; Khurana and Theurkauf, 2010; Senti and Brennecke, 2010). Recent studies have shown that piRNAs contain numerous species that cover the entire sequence of a target TE (Ghildiyal and Zamore, 2009; Malone and Hannon, 2010; Senti and Brennecke, 2010). Moreover, accumulating evidence indicates the existence of functional links between these small RNAs and silencing of TEs and explains epigenetic phenomena such as hybrid dysgenesis, defects in interspecies crosses and genome imprinting, indicating the importance of small RNAs as epigenetic regulators (Pillai and Chuma, 2012; Saito and Siomi, 2010).
In this article, I begin by providing a brief overview of piRNA expression and function, followed by a summary of the mechanism of piRNA biogenesis and TE silencing. With this background, I then describe the epigenetic roles of *Drosophila* piRNAs and their impact on genome organization.

**EXPRESSION AND FUNCTION OF ARGONAUTE PROTEINS**

Argonaute family members feature a PAZ and a PIWI domain and associate with small RNAs to act as effector molecules (Sashital and Doudna, 2010). In the well-established small RNA or RNA interference (RNAi) pathway, double-stranded RNAs (dsRNAs) are processed into small interfering RNAs (siRNAs) of approximately 21 nucleotides by the enzyme Dicer. The resulting siRNAs associate with Argonaute (Ago) proteins to form the RNA-induced silencing complex (RISC). The siRNA bound Ago protein cleaves the target transcript depending upon base pairing between the siRNA and the target; this is known as slicer activity and results in gene silencing. In an analogous way to RNAi, microRNAs (miRNAs) negatively regulate translation of target mRNAs (Ghildiyal and Zamore, 2009).

In *Drosophila*, Argonaute family members include five proteins that are further divided into two classes depending upon the tissues in which they are expressed: the Argonaute subfamily proteins, Ago1 and Ago2, are ubiquitously expressed and associate with piRNAs; Piwi, Aubergine (Aub) and Ago3, are expressed specifically in gonadal cells (Williams and Rubin, 2002). SiRNAs are bound by Ago2, whereas Ago1 binds to miRNAs (Okamura et al., 2004). In S2 cells, a cell-line derived from embryonic soma, Ago2 associates with endo-siRNAs (endogenous siRNAs) consisting of TE derived sequences and silences TE expression in an analogous way to RNAi (Czech et al., 2008; Kawamura et al., 2008; Okamura et al., 2008). However, endo-siRNAs are unlikely to be the primary factor in the system repressing TEs in somatic cells because disruption of Ago2 does not cause a dramatic effect on the viability of *Drosophila* compared with heterochromatin protein mutants (Deshpande et al., 2005; Eisenberg et al., 1990).

In contrast to somatic cells, the small RNA pathway is the major system for silencing TEs in germ cells and gonadal somatic cells where PIWI subfamily proteins are expressed and associate with piRNAs derived from TEs. Piwi was the first PIWI subfamily protein to be well characterized in animals (Cox et al., 1998). Piwi has been shown to be expressed in both germline stem cells and ovarian somatic cells, such as the follicle cells, and to be crucial for maintenance of the germline stem cells (Cox et al., 2000). The other PIWI subfamily members, Aub and Ago3, are exclusively expressed in germline cells and are required for the formation of pole cells in offspring (Brennecke et al., 2007; Li et al., 2009; Nishida et al., 2007). Sequencing of piRNAs in *Drosophila* has revealed that they are 23–30 nucleotide single-stranded RNAs that mainly originate from TEs and TE-related genomic elements (Brennecke et al., 2007; Li et al., 2009; Saito et al., 2006). Whereas retrotransposon expression is increased less than 10-fold in ago2 mutant flies, the disruption of Piwi genes can cause a more than 100-fold overexpression of TEs and sterility in females, indicating the crucial roles of PIWI subfamily genes in TE silencing as well as in germline development (Vagin et al., 2006).

**PIRNA BIOGENESIS AND TRANSPOSON SILENCING**

**PiRNA clusters in Drosophila** PiRNAs are defined as small RNAs that associate with PIWI subfamily proteins (Aravin et al., 2006; Girard et al., 2006; Lau et al., 2006). In *Drosophila*, piRNAs were originally described as rasiRNAs (repeat-associated small interfering RNAs) that were defined as a subset of small non-coding RNAs discovered in testis and derived from repetitive genomic regions, such as suppressor of stellate, satellite DNA and TEs (Aravin et al., 2003). This suggested roles for the regulation of repetitive elements. In *Drosophila*, piRNAs have several characteristics that are different from those of miRNAs and siRNAs, such as their length and sources (Brennecke et al., 2007; Gunawardane et al., 2007; Saito et al., 2006; Vagin et al., 2006; Yin and Lin, 2007). The length of piRNAs is slightly longer than miRNAs, ranging from 24 to 31 nucleotides (Brennecke et al., 2007; Gunawardane et al., 2007; Saito et al., 2006). Although some piRNAs are derived from genomic region of protein coding mRNAs, a large proportion of piRNAs is mapped on TEs (Saito et al., 2009; Robine et al., 2009). Strikingly, cloning analyses show the considerable heterogeneity of piRNAs (Saito et al., 2006; Brennecke et al., 2007; Gunawardane et al., 2007). To date, hundreds of thousands of piRNA species have been cloned and more than 70% of piRNAs have been cloned only once, indicating considerable piRNA heterogeneity and comprehensive coverage of TE sequences (Brennecke et al., 2007; Yin and Lin, 2007). PiRNAs are mapped on several hundred genomic regions, referred to as piRNA clusters where TEs and TE remnants are enriched (Fig. 1). The majority of piRNAs are generated from two types of cluster: monodirectional clusters, such as the flamenco (*flam*) locus, where piRNAs map to only one DNA strand; and dual-strand clusters, such as the 42AB region, where piRNAs are mapped on both DNA strands (Fig. 1) (Brennecke et al., 2007). The *flam* locus is composed of antisense-oriented retrotransposon copies, such as gypsy, idefix, ZAM and mdg1, resulting in the production of antisense piRNAs (Brennecke et al., 2007; Desset et al., 2008; Pelisson et al., 2008).
1994). 42AB is composed of a line of TEs, such as GATE, Rt1b and roo. In both clusters, the mapped piRNAs cover tens of thousands of DNA base pairs, suggesting that piRNAs are generated from long precursor transcripts (Brennecke et al., 2007; Klattenhoff et al., 2009).

Supporting this idea, most flam-mapped piRNAs are lost in fruit fly lines with an inserted P element at the distal end of flam (Brennecke et al., 2007; Klattenhoff et al., 2009).

Many copies of over 100 different kinds of TEs are pres-
ent in the *Drosophila* genome (Biemont and Vieira, 2006; Kapitonov and Jurka, 2003). In the absence of *flam*-derived piRNAs, *gypsy*, *idefix* and *ZAM* elements are highly expressed, resulting in defects of oogenesis, whereas disruption of piRNA biogenesis from dual-strand clusters causes derepression of TEs, such as *GATE* and *HeT-A*, suggesting that piRNAs are involved in the silencing of TEs (Brennecke et al., 2007; Desset et al., 2008; Klattenhoff et al., 2009; Pelisson et al., 1994).

**The role of Piwi** In *Drosophila* ovarian somatic cells, Piwi mainly associates with piRNAs derived from monodirectional clusters, such as *flam* (Lau et al., 2009; Saito et al., 2009). The disruption of Piwi causes desilencing of *flam*-related TEs and defects in the proliferation of ovarian follicle cells and in the maintenance of germ line stem cells (Cox et al., 1998; Saito et al., 2010; Vagin et al., 2006). This phenotype is also seen in *flam* mutant flies, suggesting the cooperative roles of Piwi and piRNAs (Desset et al., 2008; Pelisson et al., 1994). Interestingly, mutational analyses and functional studies indicate that the nuclear localization of Piwi is required for TE silencing in fruit flies and ovarian somatic cell line, whereas its slicer activity is not, suggesting that the silencing step occurs in the nucleus and is not a consequence of the direct cleavage of TE mRNAs (Cox et al., 2000; Darricarrere et al., 2013; Klénov et al., 2011; Saito et al., 2010).

Loss of Piwi expression results in a decreased abundance of the stable repressive histone mark (trimethylated H3K9) and Pol II density for some TEs, suggesting transcriptional gene silencing roles of Piwi (Fig. 2) (Klénov et al., 2011; Shpiz et al., 2011; Sienski et al., 2012; Wang and Elgin, 2011). TEs are remarkably compartmentalized in the heterochromatin, which is accomplished by histone modification, such as H3K9 and H3K27 trimethylation in *Drosophila* (Negre et al., 2011). These observations raise the question of whether the Piwi-piRNA complex causes H3K9 trimethylation by directly recruiting histone modifiers on TEs. Remarkably, in Piwi knocked down cells, H3K9 trimethylation marks are also lost in the euchromatic genes adjacent to TEs, indicating that Piwi-mediated epigenetic marks spread and affect the gene expression patterns of the genomic regions that flank TEs (Fig. 2) (Sienski et al., 2012). In addition, it is surprising that H3K9 trimethylation by Piwi-piRNAs seems to depend upon the transcription of TE mRNAs, despite TEs being tightly associated with the repressed chromatin state, H3K9 trimethylation (Sienski et al., 2012). This is analogous to the well-characterized RNA silencing system of heterochromatin formation in *S. pombe*. In *S. pombe*, siRNAs recognize the centromere region through complementarity between small RNAs and centromere-derived nascent RNA transcripts, thereby recruiting the histone modifiers to achieve heterochromatinization (Moazed, 2009). It is therefore possible that the piRNA-Piwi complex recruits histone modifiers on TEs and TE-flanking regions. However, it remains unclear whether Piwi directly induces H3K9 trimethylation on TEs and what molecular mechanisms and related factors are involved.

**The role of Aub and AGO3** In contrast to Piwi, Aub and Ago3 are exclusively expressed in germ cells and localized in the cytoplasm (Harris and Macdonald, 2001; Li et al., 2009; Nishida et al., 2007). Unlike Piwi, Aub- and AGO3-associated piRNAs are mainly mapped on bidirectional clusters (Brennecke et al., 2007; Li et al., 2009). Examination of the nucleotide preferences of piRNAs has shown that Aub-associated piRNAs have a strong bias for 5′ uracil and an antisense TE transcript orientation and that Ago3-associated piRNAs have a strong bias for adenosine at the tenth nucleotide from their 5′ end (10th A) and a sense TE transcript orientation (Brennecke et al., 2007; Gunawardane et al., 2007). Both Aub and Ago3 can cleave target RNA between nucleotides 10 and 11 from the 5′ end of their associated piRNAs *in vitro* (Gunawardane et al., 2007). Moreover, Aub-bound piRNAs are markedly decreased and TEs are derepressed in *ago3* and *aub* mutants, indicating that the silencing of TEs is coupled with piRNA biogenesis by the slicer-dependent mechanism in which Aub cleaves TE mRNA and creates the 5′ end of Ago3-associated piRNAs, and Ago3 cleaves antisense TE transcripts and creates the 5′ end of Aub-associated piRNAs, thereby amplifying piRNAs and enhancing TE silencing. This model is referred to as the ‘ping-pong cycle’, and was first described in *Drosophila* and later reported in mouse (Aravin et al., 2007; Brennecke et al., 2007; Carmell et al., 2007; Gunawardane et al., 2007; Kuramochi-Miyagawa et al., 2008). In contrast to the Piwi-mediated silencing mechanism, in which the extent of silencing by Piwi-associated piRNAs spreads to nearby genes, ping-pong-mediated silencing is unlikely to spread because the cleavage reactions depend on complementarity between the piRNAs and the targets.

**Commonalities and differences between *Drosophila* and mouse** While there are some similarities between mouse and *Drosophila* piRNAs, there are also several differences between the two species. Similar to *Drosophila*, mouse piRNAs are 24-32 nucleotides in length and are longer than miRNAs and siRNAs (Aravin et al., 2006, 2007; Carmell et al., 2007; Girard et al., 2006; Grivna et al., 2006a, 2006b; Watanabe et al., 2006). However, it seems that the character of piRNAs changes corresponding to the stage of spermatogenesis in mouse. In the fetal and neonatal testis, a large proportion of piRNAs known as pre-pachytene piRNAs originated from TEs is mapped many times in the genome (Aravin et al., 2007). In contrast to the pre-pachytene piRNAs, the proportion of pachytene-piRNAs originated from TEs is low.
in the adult testis (Aravin et al., 2006; Girard et al., 2006). In an analogous way to Drosophila, the ping-pong cycle operates in mouse. In the fetal and neonatal testis, Mili and Miwi2, both of which are PIWI-subfamily proteins, associate with pre-pachytene piRNAs and are responsible for the ping-pong cycle and TE silencing (Aravin et al., 2008; Kuramochi-Miyagawa et al., 2008). In the male gametogenesis, de novo CpG DNA methylation occurs at the imprinted genes and repetitive elements in mice fetal testes, which is essential for fertility. DNA methylation is one of several epigenetic mechanisms that cells use to repress gene expression. The epigenetic mark of DNA methylation, which is not present in Drosophila, might be established by piRNAs in fetal mice testes. Cells lacking either Mili or Miwi2 showed a marked reduction in CpG DNA methylation across TEs, such as IAP and LINE, and an increase of TE expression. Interestingly, a proportion of Miwi2 is localized in the nucleus, suggesting that Miwi2 mediates de novo DNA methylation on TEs in fetal mouse testis (Aravin et al., 2008; Kuramochi-Miyagawa et al., 2008). More recently, disruption of piRNA biogenesis has been shown to cause the loss of DNA methylation of Rasgrf1, a paternally imprinted gene, suggesting a link between piRNAs and genomic imprinting (Watanabe et al., 2011). Overall, the biogenesis mechanism and function of piRNAs show many similarities, but also important differences, between Drosophila and mouse.

**HYBRID DYSGENESIS AND ADAPTATION TO TE INVASION**

In Drosophila, the determinants specifying germ cell fate are inherited in pole cells from pole plasm at the posterior of the oocyte (Rongo and Lehmann, 1996). Pole cells then become the primordial germ cells in offspring. Aub but not Ago3 is continuously detected at the posterior pole of the oocyte and embryo, suggesting that piRNAs as well as Aub are incorporated into the germ cells of the next generation (Fig. 3A) (Brennecke et al., 2007; Nishida et al., 2007). This maternal deposition of piRNAs explains the molecular basis of the mysterious genetic phenomenon, hybrid dysgenesis, which was discovered 30 years ago (Fig. 3B) (Kidwell et al., 1977).

Hybrid dysgenesis is a sterility syndrome reported in the early 1970s (Hiraizumi, 1971; Kidwell et al., 1977). Laboratory strains of Drosophila melanogaster, separated about 100 years ago, allowed the isolation of strains carrying the P element, which are referred to as P strains, while strains lacking it are referred to as M strains (Brookfield et al., 1984; Daniels et al., 1990). The sterility syndrome is observed in progeny of crosses between P strain males and M strain females, whereas the reciprocal crosses do not show the sterility phenotype, regardless of genetic identity (Fig. 3B) (Kidwell et al., 1977; Rubin et al., 1982). Similar observations have been reported for the I element, one of the LINEs in Drosophila. I element-carrying strains are divided into two categories, one is the I strains, which possess active I elements that can transpose into genomic regions, and the other is the R strains, which contain defective copies of I elements (Bucheton et al., 1984). In the offspring of dysgenic crosses, P or I elements are overexpressed and frequent transpositions are detected. Interestingly, the repressed state of a P or I element, referred to as P cytotype or I cytotype, respectively, is maternally but not paternally inherited, suggesting that maternal components determine repression of TEs in offspring. However, it was not until 2008 that the maternally inherited piRNA-Aub complex was identified as the molecular entity of the cytotype (Brennecke et al., 2008). Regarding P-M hybrid dysgenesis, P elements do not transpose efficiently in P strains, because of the presence of high amounts of piRNAs that are amplified by the ping-pong cycle between hetero- and euchromatic P element transcripts in the female germline. Importantly, the amplification of piRNAs is not observed without maternal piRNA deposition even if the P element is paternally inherited, suggesting that maternal deposition of piRNAs is required for the accumulation of piRNAs in progeny. In other words, maternally deposited piRNAs act as the trigger that starts the piRNA amplification in offspring germ cells (Fig. 3) (Brennecke et al., 2008). Similar mechanism was proposed for I element, indicating that piRNAs serve as transgenerational information of the repressed TEs (Chambeyron et al., 2008).

As described above, maternal deposition of piRNAs is required for the ping-pong cycle and TE silencing in offspring. In addition, a supply of piRNA precursor is also required for the sustained ping-pong cycle in the offspring (Brennecke et al., 2008). Although the sense piRNAs, namely TE mRNAs, are produced from both bidirectional piRNA clusters, such as 42AB, and euchromatic TEs, the antisense piRNA precursors are only transcribed from bidirectional clusters, which are highly enriched for transposon fragments (Klattenhoff et al., 2009). Therefore, the repression of TEs is determined by not only the maternally inherited piRNAs acting as triggers but also by the TEs in bidirectional clusters acting as enhancers. From these characteristics of piRNAs, they can be considered as part of an ‘immune system’ for TEs. Therefore, it is not surprising that loss of TEs in bidirectional clusters raises the possibility of new TE invasion and that, in contrast, acquisition of new TE sequence in bidirectional clusters raises the immunity against new TE invasion. Interestingly, the observation that P element insertion into bidirectional piRNA clusters occurs within a single generation and restores fertility in dysgenic crosses indicates the existence of an adaptive system, allowing the dysgenic genome to overcome invasive TEs (Khurana et al., 2011). This adaptation efficiency increases as the hybrids of P element dysgenic
crosses age, which might contribute to genomic evolution as well as diversity.

CONCLUSION AND PERSPECTIVES

Characterization of the biogenesis and functions of piRNAs demonstrates their unique roles in maintaining genomic stability and organization. In this regard, small RNAs have numerous advantages in regulating TEs, as well as defective TE sequences. First, PIWI-piRNA complexes can act in trans, allowing the repression of interspersed TEs on chromosomes. Second, the fragmentation of TEs into small RNA sequences might contribute to the repression of diverse TEs and to tolerance of small numbers of mutations of TE sequences.

Accumulating evidence has started to emerge concerning the roles of protein factors in piRNA biogenesis; however, how piRNA-mediated silencing is achieved and what molecules are involved is largely unknown. In this regard, knowledge of the molecular basis of the small RNA pathway in other species, such as *S. pombe* and *tetrahymena* provides some hints for small RNA mediated...
mechanisms (Grewal, 2010; Mochizuki, 2012). For example, studying the potential roles of the *Drosophila* homologues of *S. pombe* I3K9 methyltransferase and heterochromatin proteins, which have key functions in determining the heterochromatin structure of the *S. pombe* centromere, will be of great interest. Similarly, the role of the scnRNA-mediated pathway for genomic rearrangement in tetrahymena is well characterized, and determining whether homologues of this machinery are required in *Drosophila* will also be an important line of future research. In addition, the small RNA-dependent TE silencing is also reported in plants (see the review by Ito, 2013). Therefore, comparing the silencing systems among organisms as well as TE species might provide clues to TE-mediated genome expansion.

In the context of hybrid dysgenesis and P element adaptation with age in *Drosophila*, the mechanisms how piRNA clusters are determined will be of great interest. One important observation is that bidirectional piRNA clusters are associated with Rhino, which is required for the production of piRNA precursors (Klattenhoff et al., 2009). In particular, it will be of great interest to address what factors Rhino recognizes and how Rhino affects piRNA precursor transcription. Given that piRNA clusters contribute to the euchromatic TE content in the genome, understanding the mechanisms that determine piRNA clusters will shed light on the molecular basis of how higher eukaryotes obtained and amplified DNA elements during evolution.

I thank Haruhiko Siomi for helpful comments. This work was supported by the Funding Program for Next Generation World-Leading Researchers (LS1009).

**REFERENCES**


Girard, A., Sachidanandam, R., Hannon, G. J., and Carmell, M.


