Piwi-interacting RNAs: biological functions and biogenesis

Kaoru Sato*1 and Mikiko C. Siomi†1,2

*Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, Tokyo, Japan
†Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Saitama, Japan

Abstract
The integrity of the germline genome must be maintained to achieve successive generations of a species, because germline cells are the only source for transmitting genetic information to the next generation. Accordingly, the germline has acquired a system dedicated to protecting the genome from ‘injuries’ caused by harmful selfish nucleic acid elements, such as TEs (transposable elements). Accumulating evidence shows that a germline-specific subclass of small non-coding RNAs, piRNAs (piwi-interacting RNAs), are necessary for silencing TEs to protect the genome in germline cells. To silence TEs post-transcriptionally and/or transcriptionally, mature piRNAs are loaded on to germline-specific Argonaute proteins, or PIWI proteins, to form the piRISC (piRNA-induced silencing complex). The present chapter will highlight insights into the molecular mechanisms underlying piRISC-mediated silencing and piRNA biogenesis, and discuss a possible link with tumorigenesis, particularly in Drosophila.

Keywords:
Drosophila, germline, piwi-interacting RNA, transposable element.

1Correspondence may be addressed to either author (email kaoru@biochem.s.u-tokyo.ac.jp or siomim@biochem.s.u-tokyo.ac.jp).
**Introduction**

RNA silencing is achieved by gene silencing pathways that are mediated by small ncRNAs (non-coding RNAs) of 20–30 nt in length. The effector complex in RNA silencing is the RISC (RNA-induced silencing complex), which consists of a member of the Argonaute family of proteins and a small RNA that guides the RISC to its targets to be silenced [1–6]. Upon RISC recognition of targets, Argonaute protein inhibits their expression by either cleaving them with its Slicer endonuclease activity, or by inducing translational inhibition, RNA destabilization and chromatin remodelling through DNA methylation and/or histone modification.

piRNAs (piwi-interacting RNAs) are a subset of small RNAs that trigger RNA silencing in the gonads. piRNAs are typically 23–30 nt long and associate specifically with germline-specific Argonaute proteins of the PIWI subfamily [7]. piRNAs were discovered during an investigation into how the tandemly repeated Ste (Stellate) gene is silenced in the male germ-line of Drosophila melanogaster [8]. Without Ste silencing, spermatogenesis does not proceed properly and males become infertile. Ste silencing was linked to a Y-chromosomal locus, Su(Ste) (Suppressor of Stellate), which consists of tandem repeats showing strong similarity to Ste at the nucleotide sequence level [6,8]. The functional importance of small RNAs arising from Su(Ste) was argued in early 2000 [8] when the mechanism underlying Ste silencing was still unclear. Later, a small RNA profiling study on Drosophila testes and early embryos revealed endogenous 23–30 nt small RNAs derived from repetitive intergenic elements scattered across the genome, including from Su(Ste) and TEs (transposable elements) [9]. These small RNAs were originally termed rasiRNAs [repeat-associated siRNAs (small interfering RNAs)], but they are currently known as piRNAs because they associate with PIWI proteins to function in invertebrate and vertebrate RNA silencing [10–15].

In the present chapter, we summarize our current understanding of piRNA functions and biogenesis, mainly focusing on studies using D. melanogaster. We also review recent progress in our understanding of the biological involvement of the piRNA pathway in non–gonadal cells, such as in brain tumour development.

**Targets of piRNA-mediated silencing**

**TE silencing by the piRISC (piRNA-induced silencing complex)**

In Drosophila, piRNAs associate with three PIWI proteins, Piwi, Aub and AGO3, to form the piRISC and guide the silencing of RNAs that accommodate complementary sequences. The main targets of piRNAs are TEs and this is highly conserved across animal species. TEs are genomic parasites that move from one chromosomal location to another by either a cut-and-paste (transposition) or copy-and-paste (insertion) mode of action [16]. In this way, they potentially modify and disrupt the functions of other genes and often threaten the integrity of the host genome [6,16]. Studies using Drosophila oogenesis have provided critical insights into the molecular mechanisms underlying piRNA biogenesis and piRNA-mediated TE silencing.

Mutations in piwi, aub or ago3 lead to TE derepression in the germline, indicating the non-redundancy of these genes [6,17–23]. Indeed, all PIWI genes have crucial roles in gonadal development: both piwi and aub are required for male and female fertility [24,25], whereas
ago3 is required for female fertility, but only partially required for male fertility [18]. The non-redundancy of PIWI proteins may also be explained by the fact that, unlike germline-specific Aub and Ago3, piwi is expressed in both germline and somatic cells [19–22,26].

PIWI proteins exhibit Slicer activity in vitro; thus piRISCs probably silence their targets post-transcriptionally, by cleaving them, as does Ago2 associated with siRNAs. Aub and AGO3 are involved in Slicer-mediated silencing in the cytoplasm (see below for more detail). Unlike Aub and Ago3, Piwi is localized in the nucleus. Also, nuclear localization, but not Slicer activity of Piwi, is required for silencing TEs [26,27]. In mice, DNA methylation influences chromatin structures and is strongly related to piRNA-mediated RNA silencing in testes [28], leading to speculation that Piwi in flies would also induce DNA methylation or covalent histone modifications in the nucleus. Yet, DNA methylation activity has not been conclusively detected in Drosophila gonads.

**Protein-coding gene silencing by the piRISC**

Some piRNAs are proposed to target protein-coding genes, such as FasIII (Fasciclin III) and nos (nanos) [20,27,29]. The mRNA level of FasIII is up-regulated in tj (traffic jam) mutants, which show a failure of intermingling of germline cells with surrounding somatic cells in the larval ovary, leading to infertility [30]. The tj gene encodes a large Maf transcription factor. Interestingly, piRNAs are produced from the 3′-UTR (untranslated region) of tj mRNAs, and these form a piRISC with Piwi. In piwi mutants, FasIII is up-regulated, implying that tj-piRNAs silence FasIII in collaboration with Piwi [27]. piRNAs may also induce degradation of maternally deposited mRNAs. nos encodes a posterior morphogen important for Drosophila germline development. nos mRNA deadenylation and decay are likely to be regulated by piRNAs derived from roo and 412 transposons, which show complementarity to the nos 3′-UTR. Smaug, an RNA-binding protein that provides translational repression of unlocalized nos mRNA, may assist by recruiting the CCR4 deadenylation complex to the target during maternal-zygotic transition in the embryo [29].

**piRNA biogenesis**

In animals, endogenous siRNAs also silence TEs [1,3]; however, in contrast with non-gonadal somatic cells where endogenous siRNAs are the main trigger of TE silencing, piRNAs in the germline function at the forefront of the defence against transposons. The current model of piRNA biogenesis involves two spatially and mechanistically distinct pathways: the primary processing pathway and a secondary amplification pathway (ping-pong amplification loop) (Figure 1). In the germline, piRNA biogenesis involves both pathways, whereas piRNAs in the gonadal somas are generated solely via the primary processing pathway.

**The primary processing pathway**

Primary piRNAs are produced from the piRNA clusters that act as sources of piRNAs and that are often located in pericentromeric or subtelomeric regions on the genome [6,26]. Each cluster spans several to more than 200 kb and contains multiple transposon fragments [6,26]. Most of the piRNA clusters produce piRNAs from both genomic strands, suggesting bidirectional transcription [6,26]. Other clusters, such as flam (flamenco), produce piRNAs almost exclusively from one genomic strand [26]. Investigation of piRNA biogenesis in Drosophila ovarian
Somas, in which only Piwi is expressed, revealed the molecular mechanism of the primary processing pathway. Representatives of primary piRNAs are鱼类-piRNAs, which are predominantly expressed in ovarian somatic cells (Figure 1) [19].鱼类-piRNAs are mostly antisense to active transposons, and thus act as trans-silencers of TEs. Studies have identified two putative RNA helicases, Armi and FS(1)Yb [Female Sterile (1) Yb; also known as Yb], and a nuclease, Zuc (Zucchini), as primary piRNA factors (Figure 2). A lack of Armi or Yb eliminates Piwi-associated primary piRNAs from somatic follicle cells of ovaries and cell cultures, the ovarian somatic sheet and ovarian somatic cells [31–34]. Armi and Yb are components of Yb bodies, non-membranous high-density structures adjacent to mitochondria [31,34]. Depletion of Yb causes the disappearance of Yb bodies (Figure 2), which interferes with primary piRNA production and piRISC formation, causing mislocalization of Piwi in the cytoplasm, leading to derepression of TEs [31–34]. Therefore Yb bodies can be considered as the cytoplasmic centre for piRNA production, piRISC formation and inspection. Zuc was originally identified as a gene required for axis determination during oogenesis, and was also found to be required for piRNA biogenesis in ovaries and testes. Zuc encodes an endoribonuclease which localizes on the surface of mitochondria (Figure 2). A lack of Zuc eliminates Piwi-associating primary piRNAs from somatic follicle cells of ovaries and cell cultures, the ovarian somatic sheet and ovarian somatic cells [31–34]. Armi and Yb are components of Yb bodies, non-membranous high-density structures adjacent to mitochondria [31,34]. Depletion of Yb causes the disappearance of Yb bodies (Figure 2), which interferes with primary piRNA production and piRISC formation, causing mislocalization of Piwi in the cytoplasm, leading to derepression of TEs [31–34]. Therefore Yb bodies can be considered as the cytoplasmic centre for piRNA production, piRISC formation and inspection. Zuc was originally identified as a gene required for axis determination during oogenesis, and was also found to be required for piRNA biogenesis in ovaries and testes. Zuc encodes an endoribonuclease which localizes on the surface of mitochondria (Figure 2). A lack of Zuc eliminates Piwi-associating primary piRNAs, and leads to the accumulation of piRNA precursor- or intermediate-like molecules, suggesting that piRNA intermediates derived from uni-strand clusters such as鱼类 are cleaved by Zuc to produce piRNAs that bind to Piwi [27,34,35]. In germline cells, primary piRNAs are

Figure 1. Two piRNA biogenesis pathways
In the Drosophila melanogaster primary piRNA processing pathway, antisense transcripts transcribed from piRNA clusters and/or transposons are processed to piRNAs by unknown mechanisms and are loaded on to Aub or Piwi. piRNAs derived from the鱼类 locus are exclusively loaded on to Piwi because鱼类 is active only in ovarian somas where only Piwi is expressed. piRISCs produced through this mechanism act as a ‘trigger’ for the amplification loop. The amplification loop (also known as the ping-pong cycle) most probably involves the Slicer activity of Aub and Ago3, but not that of Piwi. Aub associated with antisense piRNA cleaves complementary piRNA precursors (sense). This determines the 5′-ends of piRNAs that are loaded on to Ago3. Ago3 associated with sense piRNA cleaves complementary piRNA precursors, generating the 5′-end of antisense piRNAs, which are subsequently loaded on to Aub. The 3′-ends of piRNAs are trimmed by an unknown nuclease (or nucleases), which is followed by 2′-O-methylation mediated by HEN1/Pimet. piRNAs that induce the amplification loop may also be maternally deposited.

© The Authors Journal compilation © 2013 Biochemical Society
required to initiate the amplification loop, although the protein factors required for the primary pathway in the germline remain unknown.

The ping-pong amplification loop
As in ovarian somas, in the germline, the primary processing pathway provides an initial pool of piRNAs. Primary piRNAs arising from bidirectional piRNA clusters, such as the 42AB cluster, which contains transposon fragments in both sense and antisense orientations, act as sources for the ping-pong amplification loop. The ping-pong amplification loop further shapes the piRNA population by amplifying sequences that target active transposons. The ping-pong amplification loop is conserved in many animal species [6,14,15].

The ping-pong amplification loop requires pre-existing primary piRNAs, which are mainly antisense against active transposon transcripts (Figure 1). These piRNAs preferentially associate with two PIWI subfamily proteins, Piwi and Aub [6,19,26]. In contrast, sense-strand piRNAs preferentially associate with Ago3 [6,22,26]. Antisense piRNAs bound to Piwi or Aub show a strong bias for 1U-containing piRNAs, whereas sense-strand piRNAs bound to Ago3 tend to have an adenine at position 10 [6,22,26]. piRNAs that are associated with Aub or Ago3 often overlap at their 5'-ends by 10 nt. These findings led to the ping-pong amplification loop
model for piRNA biogenesis, where antisense piRNAs with Aub direct the cleavage of sense-strand transposon transcripts, generating sense piRNAs for Ago3 (Figure 3). The resulting Ago3–piRNA complex then directs cleavage of antisense piRNA precursors, generating antisense piRNAs for Aub. The 5′-ends of amplified secondary piRNAs are determined by Aub and Ago3 Slicer. The factors necessary for 3′ trimming in this cycle remain undetermined. The ping-pong amplification loop thus obviates the need for an RdRP (RNA-dependent RNA polymerase), which is needed to amplify siRNA triggers in plants, nematodes and yeast [1]. In fact, genes encoding RdRP are not found in the Drosophila genome.

piRNA pathway factors in Drosophila
Factors for piRNA cluster transcription
The mechanism underlying transcription of piRNA clusters is still an important question to be answered in the piRNA field, although several proteins have been shown to be involved in this process. Rhi (Rhino), which encodes an HP1 (heterochromatin protein 1) homologue, associates with the 42AB cluster and this association is required for its transcription and for secondary piRNA accumulation specifically in germline cells (Table 1) [36]. cuff (cutoff), a gene related to the yeast transcription termination factor Rai1, physically interacts with Rhi and accumulates at centromeric/pericentromeric regions in the nucleus of germline cells and strongly co-localizes with the major heterochromatic domains [37]. Cuff is required for 42AB cluster transcription, but some regions in the 42AB cluster are transcribed by a Cuff-independent mechanism [37], suggesting that the dual-strand cluster might not produce single full-length transcripts spanning the entire locus, as opposed to the flam locus [26]; rather, the 42AB locus may have multiple internal promoters.

A histone methyltransferase, dSETDB1, responsible for H3K9me3 (histone H3 Lys9 trimethylation), is required for both bi- and uni-directional piRNA cluster transcription in both germline and somatic cells of the gonads (Table 1) [39], indicating that H3K9me3 is a unifying
Table 1. Factors involved in piRNA biogenesis in *Drosophila*

B-box, zinc finger B-box domain; CHROMO, CHRromatin Organization MOdifier domain; CHROMO shadow, a variant of CHROMO; DEAD, DEAD-box helicase; DEXDc, DEAD-like helicase domain; EGFP, enhanced green fluorescent protein; HA, haemagglutinin; HA2, helicase-associated domain; HELICc, helicase-superfamily C-terminal domain; KH, K homology domain, MBD, Methyl-CpG binding domain; MID, middle; MYND, zinc-finger myeloid-nervy-DEAF-1 domain; N.D., not determined; PAZ, PIWI/Argonaute/Zwille; PIWI, P-element-induced whimpy testes; PreSET, N-terminal domain to SET; RING, Really Interesting New Gene finger domain; SAM, S-adenosylmethionine; SET, Su(var)3-9/Enhancer-of-zeste/Trithorax domain; SNase, Staphylococcal nuclease; UBA, ubiquitin associated domain; - , undetected.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Symbol</th>
<th>CG number</th>
<th>TUDOR domains</th>
<th>Other domains</th>
<th>Category</th>
<th>Mammalian homologue</th>
<th>Germ(^*)</th>
<th>Germline-soma(^\dagger)</th>
<th>Interaction with PIWIs</th>
<th>Other binding partners(^\ddagger)</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>piwi</td>
<td>piwi</td>
<td>CG6122</td>
<td>PAZ, MID, PIWI</td>
<td>PIWI subfamily proteins</td>
<td>PIWIL1–4</td>
<td>Nucleus</td>
<td>Nucleus</td>
<td>None</td>
<td>Yb, Armi, Tud, Vas, Spn-E, Tej</td>
<td>[26,57]</td>
<td></td>
</tr>
<tr>
<td>aurbergine</td>
<td>aub</td>
<td>CG6137</td>
<td>PAZ, MID, PIWI</td>
<td>PIWI subfamily proteins</td>
<td>PIWIL1–4</td>
<td>Nuage</td>
<td>-</td>
<td>Ago3</td>
<td>[26,57]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>argonaute 3</td>
<td>ago3</td>
<td>CG40300</td>
<td>PAZ, MID, PIWI</td>
<td>PIWI subfamily proteins</td>
<td>PIWIL1–4</td>
<td>Nuage</td>
<td>-</td>
<td>Aub</td>
<td>[5,18,26]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dPRMT5</td>
<td>dPRMT5</td>
<td>CG3730</td>
<td>SAM-methyltransferase</td>
<td></td>
<td>PRMT5</td>
<td>N.D.</td>
<td>N.D.</td>
<td>N.D.</td>
<td>N.D.</td>
<td>[44]</td>
<td></td>
</tr>
<tr>
<td>valois</td>
<td>vls</td>
<td>CG10728</td>
<td>4 × WD</td>
<td>Arginine methylation</td>
<td>MEP50</td>
<td>Nuage</td>
<td>-</td>
<td>N.D.</td>
<td>dPRMT5, Tud</td>
<td>[58]</td>
<td></td>
</tr>
<tr>
<td>partner of piwis</td>
<td>papi</td>
<td>CG7082</td>
<td>2 × KH</td>
<td>TDRDs</td>
<td>TDRD2</td>
<td>Nuage</td>
<td>-</td>
<td>Piwi, Ago3</td>
<td>Tral, Me31B, TER94</td>
<td>[41]</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Table 1. Factors involved in piRNA biogenesis in *Drosophila*  *(Continued)*

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Symbol</th>
<th>CG number</th>
<th>TUDOR domains</th>
<th>Other domains</th>
<th>Category</th>
<th>Mammalian homologue</th>
<th>Germ*</th>
<th>Germine-soma†</th>
<th>Interaction with PIWI</th>
<th>Other binding partners‡</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>qin</em></td>
<td><em>qin</em></td>
<td>CG14303, CG14306</td>
<td>5</td>
<td>RING, B-box</td>
<td>TDRDs, piRNA cluster transcription</td>
<td>TDRD4</td>
<td>Nuage, nucleus</td>
<td>-</td>
<td>Piwi, Aub</td>
<td>Vas, Spn-E</td>
<td>[40,42]</td>
</tr>
<tr>
<td><em>tejas</em></td>
<td><em>tej</em></td>
<td>CG8589</td>
<td>1</td>
<td>Lotus (Tejas)</td>
<td>TDRDs</td>
<td>TDRD5</td>
<td>Nuage</td>
<td>-</td>
<td>Aub</td>
<td>Vas, Spn-E</td>
<td>[43]</td>
</tr>
<tr>
<td><em>tudor</em></td>
<td><em>tud</em></td>
<td>CG9450</td>
<td>11</td>
<td>TDRDs</td>
<td>TDRDs</td>
<td>TDRD6</td>
<td>Nuage</td>
<td>Cytoplasm</td>
<td>Aub, Ago3</td>
<td>N.D.</td>
<td>[44,57]</td>
</tr>
<tr>
<td><em>spindle-E</em></td>
<td><em>spn-E</em></td>
<td>CG3158</td>
<td>1</td>
<td>DEXDc, HELICc, HA2</td>
<td>TDRDs</td>
<td>TDRD9</td>
<td>Nuage</td>
<td>Aub</td>
<td>Vas, Tej</td>
<td></td>
<td>[38]</td>
</tr>
<tr>
<td><em>FS(1)Yb</em></td>
<td><em>Yb</em></td>
<td>CG2706</td>
<td>1</td>
<td>Helicase-like, DEAD, Helicase, ZF</td>
<td>TDRDs</td>
<td>TDRD12</td>
<td>Yb body</td>
<td>Piwi</td>
<td>Armi, Vret</td>
<td></td>
<td>[31–34]</td>
</tr>
<tr>
<td><em>BoYb</em></td>
<td></td>
<td>CG11133</td>
<td>1</td>
<td>DEAD, Helicase, ZF</td>
<td>TDRDs</td>
<td>TDRD12</td>
<td>Nuage</td>
<td>-</td>
<td>N.D.</td>
<td>N.D.</td>
<td>[45]</td>
</tr>
<tr>
<td><em>SoYb</em></td>
<td></td>
<td>CG31755</td>
<td>2</td>
<td>DEAD, Helicase, ZF</td>
<td>TDRDs</td>
<td>TDRD12</td>
<td>Cytoplasm</td>
<td>Yb body</td>
<td>N.D.</td>
<td>N.D.</td>
<td>[45]</td>
</tr>
<tr>
<td><em>krimp</em></td>
<td><em>krimp</em></td>
<td>CG15707</td>
<td>1</td>
<td>None</td>
<td>TDRDs</td>
<td>None</td>
<td>Nuage</td>
<td>N.D.</td>
<td>N.D.</td>
<td>N.D.</td>
<td>[38,57]</td>
</tr>
<tr>
<td><em>vreteno</em></td>
<td><em>vret</em></td>
<td>CG4771</td>
<td>2</td>
<td>None</td>
<td>TDRDs</td>
<td>None</td>
<td>Nuage</td>
<td>Yb body</td>
<td>Piwi, Aub, Ago3</td>
<td>Yb, Arm</td>
<td>[46]</td>
</tr>
<tr>
<td><em>dSETDB1</em></td>
<td><em>dSETDB1</em></td>
<td>CG12196</td>
<td>2</td>
<td>MBD, PreSET, 2 × SET</td>
<td>TDRDs, piRNA cluster transcription</td>
<td>SETDB1</td>
<td>Nucleus</td>
<td>Nucleus</td>
<td>N.D.</td>
<td>N.D.</td>
<td>[39]</td>
</tr>
<tr>
<td>Gene</td>
<td>Gene Symbol</td>
<td>Gene Name</td>
<td>Gene Classification</td>
<td>Chromosome Location</td>
<td>Localization</td>
<td>Notes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>-----------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>--------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rhino</td>
<td>rhi</td>
<td>CG10683</td>
<td>CHROMO, CHROMO shadow</td>
<td>piRNA cluster transcription</td>
<td>HP1d Nucleus</td>
<td>N.D.</td>
<td></td>
<td>Cuff</td>
<td>[36]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cutoff</td>
<td>cuff</td>
<td>CG13190</td>
<td>RAI1-like piRNA cluster transcription</td>
<td>DOM3Z Nuage (triple HA-tagged), nucleus (EGFP-tagged)</td>
<td></td>
<td>N.D.</td>
<td>Rhi</td>
<td>[37]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vasa</td>
<td>vas</td>
<td>CG43081</td>
<td>DEADc, HELICc</td>
<td>?</td>
<td>VASA Nuage</td>
<td>A Ub</td>
<td>Tej, Spn-E</td>
<td>[38]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>armitage</td>
<td>ami</td>
<td>CG11513</td>
<td>Helicase</td>
<td>?</td>
<td>MOV10 Cytoplasm</td>
<td>Yb body Piwi Yb, Vret</td>
<td>[22–31]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>zucchini</td>
<td>zuc</td>
<td>CG12314</td>
<td>Phospholipase-D/ nuclease family Endoribonuclease</td>
<td>mitoPLD Nuage</td>
<td>N.D.</td>
<td>Aub N.D.</td>
<td>[35,50]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>maelstrom</td>
<td>mael</td>
<td>CG11254</td>
<td>HMG-box Nuclease?</td>
<td>Maelstrom Nuage, nucleus, spindle MT</td>
<td>Cytoplasm, nucleus, spindle MT</td>
<td>none MTOC proteins</td>
<td>[49]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>squash</td>
<td>squ</td>
<td>CG4711</td>
<td>Similar to RNase HII</td>
<td>Nuclease? None</td>
<td>Nuage N.D.</td>
<td>Aub N.D.</td>
<td>[50]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>heat shock</td>
<td>hsp83</td>
<td>CG1242</td>
<td>Hsp90, ATPase Chaperon</td>
<td>Hsp90 Any subcellular fraction</td>
<td>Any subcellular fraction Piwi HOP</td>
<td>[52,53]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Nurse cells
†Follicle cells
‡Only proteins relating to the piRNA or other small ncRNA biogenesis pathways.
mark dictating piRNA cluster transcription in both germline and somatic cells. Rhi recognizes H3K9me3 at bidirectional germline clusters, but there must be an alternative Rhi-independent mechanism directing somatic and unidirectional germline cluster transcription. These observations strongly indicate that repressive marks deposited by dSETDB1 are required for transcription from all major piRNA clusters, although how these marks are targeted to piRNA clusters and how repressive marks act in the transcription of piRNA clusters remain unknown.

Other piRNA pathway factors

Many TDRD (Tud domain-containing) proteins, including Tudor, Spindle-E, Krimper, Tejas, Vreteno, Qin/Kumo and PAPI (Partner of PIWIs) are required for piRNA-mediated TE silencing in germline cells, most likely for the amplification loop (Table 1) [40–46]. Both genetic and biochemical studies have indicated that TDRD proteins associate with PIWI proteins through their sDMAs (symmetric dimethylarginine residues) and that the responsible factor for the sDMA modification of PIWI proteins is DART5/PRMT5 (protein arginine N-methyltransferase 5) [6,47]. Tudor acts as a ‘scaffold’ facilitating the piRNA amplification pathway by recruiting sDMA-modified Aub and Ago3 and their targets in germline cells [44]. Functions of other TDRD proteins remain largely undetermined.

A number of proteins other than TDRD proteins have been genetically implicated in piRNA biogenesis. Mael (Maelstrom), a HMG (high mobility group)-box protein, is a piRNA factor (Table 1) [38,48]. A recent study showed that Mael co-ordinates microtubule organization via interacting with protein components of the MTOC (microtubule-organizing centre) [49]; however, it remains unknown at the molecular level how microtubule organization is physically connected with piRNA biogenesis. Mutations in a helicase, Vasa, and in a putative nuclease, Squash, also have an impact on piRNA populations in ovaries (Table 1) [6,50]. These piRNA factors are mostly localized at the nuage [6,38,50]. Functional alterations of Hsp (heat-shock protein) 90, which have been previously implicated in canalization [51], affect the piRNA biogenesis pathway, leading to transposon activation and the induction of morphological mutants (Table 1) [52]. Hsp90 forms a complex with Piwi and an Hsp70/Hsp90 organizing protein homologue, HOP (Table 1) [53]. Post-translational regulation of Piwi by Hsp90 and HOP may allow Piwi to suppress the generation of new genotypes by transposon-mediated ‘canonical’ mutagenesis [51–53]. The precise molecular functions of many piRNA factors await further investigation.

The possible role(s) of piRNAs in cancer

A growing number of studies have found that PIWI proteins in humans and mice, specifically, HIWI, PIWIL2 and PIWIL2-like proteins, are expressed in various types of tumour cells [54,55]. In addition, piRNAs were also detected in these cells [55]. These results indicate that cancer development may be linked to the piRNA pathway. A study demonstrated that ectopic expression of piRNA pathway genes contributes to the growth and development of malignant brain tumours in Drosophila [56]. Inactivation of the germline genes vasa, piwi or aub suppressed malignant tumour growth, demonstrating that germline traits are necessary for tumour growth, at least in Drosophila [56]. Although the potential role of piRNAs in cancer has just emerged and remains to be investigated, these data highlight the importance of understanding the exact role of the piRNA pathway during tumorigenesis and suggest new possibilities for tumour therapy.
Conclusions

The system that piRNAs utilize to protect the genome in the gonads against harmful TEs is elaborate. Recent detailed investigations of the piRNA machinery have revealed that piRNAs are mostly produced through a very dedicated, rather complex, system that requires many factors. New factors are still emerging, and a full understanding of piRNA biogenesis is still a long way off. Thus we still struggle to see fully how piRNAs, especially those in the nucleus, silence TEs at the molecular level. The results of recent studies suggest that piRNAs might also be involved in regulating stem cell and cancer development outside of the gonads. Further studies with more sophisticated techniques than those used to date will be required to achieve our goal and to potentially expand our use of piRNAs as tools in research and therapy.

Summary

- piRNAs (piwi-interacting RNAs) are a germline-specific class of small non-coding RNAs.
- piRNAs protect the integrity of germline cell genomes from harmful transposons.
- There are two plausible models for piRNA biogenesis, the primary piRNA biogenesis pathway and the ping-pong amplification loop.
- Many proteins, such as Tudor domain-containing proteins and heterochromatin-related proteins, are involved in the piRNA biogenesis pathway.
- piRNAs and PIWI subfamily proteins are expressed in somatic stem cells and several cancer cells, and might regulate stem cell and cancer development.

We apologize to colleagues whose relevant primary publications were not cited because of space constraints. We thank all the members of the Siomi laboratory at Keio University School of Medicine, especially K. Saito and H. Siomi, for useful comments and critical reading of the chapter before submission. Work in our laboratory is supported by grants from the Ministry of Education, Culture, Sports, Science and Technology (MEXT, Japan) and CREST (JST, Japan).

References


58. Anne, J. and Mechler, B.M. (2005) Valois, a component of the nuage and pole plasm, is involved in assembly of these structures, and binds to Tudor and the methyltransferase Capsuléen. Development 132, 2167–2177