Review

Cell PRESS

The piRNA pathway: a fly's perspective on the guardian of the genome

Kirsten-André Senti and Julius Brennecke

Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Dr. Bohrgasse 3, 1030 Vienna, Austria

Throughout the eukaryotic lineage, small RNA silencing pathways protect the genome against the deleterious influence of selfish genetic elements such as transposons. In animals an elaborate small RNA pathway centered on PIWI proteins and their interacting piRNAs silences transposons within the germline. In contrast to other small RNA silencing pathways, we lack a mechanistic understanding of this genome defense system. However, genetic and molecular studies have uncovered a fascinating conceptual framework for this pathway that is conserved from sponges to mammals. We discuss our current understanding of the piRNA pathway in *Drosophila* with an emphasis on origin and biogenesis of piRNAs.

The piRNA pathway in the genome–transposon conflict

Eukaryotic genomes harbor a variety of selfish genetic elements, stretches of DNA that gain a transmission advantage relative to the rest of the genome, although not increasing the organism's fitness [1]. The best-understood and most widespread selfish elements are mobile elements called transposons [2]. The success of these 'genome parasites' derives from their ability to multiply within the genome by transposition to new sites. This ultimately affects host fitness owing to insertional mutagenesis and ectopic chromosomal recombination. Throughout the eukaryotic lineage, the threat posed by transposons is met by host defense systems that selectively silence them. Although early genetic studies pointed to the existence of such defense systems [3,4], for a long time their molecular nature remained mysterious. This changed abruptly when small RNA pathways were discovered which govern RNA-mediated silencing [5–8]. Over the past ten years it has become increasingly evident that small RNA silencing pathways protect the genomes of plants, fungi and animals against transposons and other selfish elements [9,10].

Here we discuss a small RNA silencing pathway that is selectively active in animal gonads where it safeguards the genome of reproductive cells against transposons. This so-called piRNA pathway centers on PIWI family proteins and their bound PIWI-interacting RNAs (piRNAs). We focus on the piRNA pathway of the *Drosophila* ovary where a long history of genetic research, combined with recent studies of small RNAs, has revealed the conceptual framework of this genome-surveillance system which is molecularly conserved from sponges to mammals.

Concepts of small RNA silencing pathways

A silencing machine called the RNA-induced silencing complex (RISC) is common to all small RNA pathways. Its central components are an Argonaute family protein and a bound small RNA [11]. Complementary base pairing with the small RNA guides RISC to specific target RNAs, and this typically results in target silencing. The remarkable elegance of small RNA pathways is their inherent simplicity. Because mRNAs are key intermediates of all gene expression programs, loading an Argonaute protein with a small RNA can allow the inhibition of essentially every cellular process [12,13]. Argonaute family proteins have diversified during evolution, and three small RNA pathways can be distinguished in most animals. These are the ubiquitous microRNA pathway, the ubiquitous small-interfering (siRNA) pathway, and the generally germline-specific piRNA pathway (Box 1).

A key advantage of small RNA pathways in the defense against foreign genetic elements is that the target sequence can also act to trigger small RNA biogenesis. Small RNAs are thereby inevitably coupled to their target, even if target sequences evolve rapidly. In most animals the siRNA and piRNA pathways implement this principle. Within the siRNA pathway, intra- or inter-molecular double-stranded RNAs (dsRNA) originating from transposons or viral transcripts trigger siRNA production via Dicer. Loaded into RISC, these siRNAs guide target silencing *in trans*, and provide a protective layer in somatic cells [12,13] (Box 2). The real battle, however, takes place in the germline where transposons are particularly active due to their predominantly vertical transmission strategy. In the germline, the piRNA pathway silences selfish elements to ensure genetic stability across generations [10,14]. piRNA biology differs considerably from that of other small RNA pathways, and very little is known about piRNA biogenesis and their mode of action.

The *Drosophila* ovary: evidence for distinct piRNA modules

The *Drosophila* ovary consists of germline cells and somatic support cells (predominantly follicle cells) that have key roles in maintaining and protecting the germline cells (Figure 1a). Within the germline, nurse cells and oocytes share a common syncytial cytoplasm. Intricate connections also exist between germline and somatic cells, for example via exchange of developmental signals or nutrient flow into the germline. This exposes the oocyte genome to both internal and external threats. On the one hand, several transposons are highly active in germline cells and exploit cellular machineries to maximize access to the oocyte genome [15]. On the other hand, some retroelements of the *gypsy* family form viral particles in somatic support cells, and these particles can invade the germline,

Corresponding author: Brennecke, J. (julius.brennecke@imba.oeaw.ac.at)

0168-9525/$ – see front matter doi:10.1016/j.tig.2010.08.007 Trends in Genetics, December 2010, Vol. 26, No. 12

Box 1. Small RNA pathways

Genetic studies identified several basic concepts of small RNA-mediated regulation years before RNA interference and related small RNA pathways were described [28,85,86]. However, key discoveries in the field came much later with the finding that dsRNA is the trigger for RNA interference and the identification of small RNAs and Argonaute proteins as the key components of all small RNA pathways [5–8].

Evolution has shaped a diverse array of pathways based on the common principle of target repression via small complementary RNA guides. This is best illustrated by the radiation of Argonaute proteins, the universal binding partners for small RNAs, which in many cases are able to cleave (slice) the target upon successful recognition [87,88]. In most animals, two classes of Argonaute proteins can be distinguished: the AGO subfamily and the PIWI subfamily (a third subfamily, the so-called WAGO proteins, has been identified only in nematodes [89]). AGO proteins are expressed ubiquitously and are loaded with microRNAs or endogenous siRNAs in response to specific dsRNA triggers. Whereas microRNAs guide the regulation of endogenous gene expression programs [90], siRNAs are mostly involved in the suppression of foreign gene expression [13] from viruses or selfish genetic elements. In flies the siRNA pathway is much more elaborate than in mammals, presumably because insects lack an adaptive immune system.

Most animals possess two or three PIWI family proteins that are typically expressed in gonads. Although both flies and mice express three PIWI proteins, pairwise orthologies cannot be determined. In fact, the three *Drosophila* PIWI proteins Piwi, Aubergine and AGO3 are more closely related to each other than to the mouse PIWI family members MILI, MIWI and MIWI2. This suggests that PIWI proteins radiated independently in both lineages from a single ancestral protein. Mutations affecting PIWI family lead to sterility and severe defects in gametogenesis in all animals examined so far [10,14].

Box 2. Transposon silencing by endogenous siRNAs

In flies, deep sequencing of small RNAs bound to AGO2 from somatic and germline tissues identified a large fraction of endogenous siRNAs (endo-siRNAs) with sequences corresponding to transposons and other genomic repeats [33,34,36,91]. Presumably, dsRNA from sense and antisense transcripts triggers their production. In ovaries, endo-siRNA profiles therefore overlap with those of piRNAs, but they lack a similar antisense bias. Significant derepression of several transposons at the RNA level has been observed in ovaries and flies mutant for AGO2, Dicer-2 or Loquacious, the three key factors for the endo-siRNA pathway. Thus, in gonads piRNAs and siRNAs collaborate to silence transposons, although the extent of repression (at least for several elements) appears to be much higher for the piRNA pathway [26,32]. The *Penelope* element of *Drosophila virilis*, however, is largely controlled via endo-siRNAs with only very few piRNAs targeting this element being identified [92]. In non-gonadal tissues, the endo-siRNA pathway appears to be the only line of defense against transposons, although additional repression at the chromatin level cannot be excluded.

In mammals, endo-siRNAs have so far been only identified in oocytes where they have an important role in transposon control [93,94]. Studies in the nematode *Caenorhabditis elegans* indicate a much more pronounced role of endo-siRNAs in defense against transposons [95–97]. In this case, however, the nematode-specific WAGO family of Argonaute proteins makes direct comparisons to other siRNA pathways more difficult.

In flies and mice, some endo-siRNAs are also generated from piRNA clusters [34,35]. This might simply reflect the ability of piRNA cluster transcripts to form low levels of dsRNA with complementary transposon transcripts. However, a significant mechanistic connection between the two pathways seems unlikely because piRNA pathway mutations have little impact on siRNA populations and siRNA pathway mutations do not affect piRNA pools [31,32,34].

presumably via cellular transport vesicles [16–19]. In both cell types the piRNA pathway is the major line of defense against transposons. How silencing is achieved, however, differs significantly. In fact, whereas germline cells express three PIWI family members [Piwi, Aubergine (Aub) and AGO3], somatic support cells exclusively express Piwi [20–22]. We first describe the linear piRNA module that acts in ovarian somatic support cells. We then build on these concepts to introduce the more complex scenario of the germline where a piRNA amplification module based on PIWI proteins is active.

A linear piRNA module in somatic gonadal cells

piRNA biology in somatic cells of the ovary exhibits an overall simple architecture (Figure 1b). All somatic support cells express Piwi, the only protein of the PIWI family in flies to localize in the nucleus [23,24]. Piwi binds to a spectrum of predominantly transposon-derived piRNAs [20,25–27] and silences transposon expression by an unknown mechanism. The only other factor with an understood function is the X-chromosomal *flamenco* locus that serves as a major source of piRNAs [20,28–31]. To a large extent, deep-sequencing of piRNA populations has provided our current insight into somatic piRNA biogenesis, their spectrum of targets, and their biological roles.

Somatic piRNAs and their origin

The purest datasets of somatic piRNAs were obtained from an ovarian somatic sheet cell line (OSS cells) [21,22]. We will use this dataset to illustrate key features of somatic piRNAs. We note that populations of Piwi-bound piRNAs from entire ovaries, and the population of piRNAs that is selectively found in ovaries, but not in eggs (during the final stages of oocyte development, germline cells release their contents into the growing oocyte and somatic support cells undergo apoptosis), strongly suggest that OSS cells accurately reflect the *in vivo* situation [31,32].

Somatic Piwi-bound piRNAs are ~22–30 nt long and ~75% carry a uridine at their 5′ ends, a pattern found for several Argonaute family proteins [20,25,27]. Over 60% of somatic piRNAs map to multiple genomic loci. This comes as no surprise given that 70% map to annotated transposons or transposon fragments, a strong enrichment compared to the average transposon content of ~10% over the assembled genome (Box 3). The remaining 30% of piRNAs map to non-annotated regions and protein-coding genes. Given the observed piRNA composition, the cell evidently selects specific RNAs for piRNA processing. Significant insight into this process, and therefore also piRNA biogenesis, has been extracted from piRNAs that map to transposons and gene exons.

Transposon-derived piRNAs

Over 90% of transposon annotated piRNAs in the soma are antisense to active transposon transcripts [21], in clear contrast to the siRNA pathway where sense and antisense small RNA populations are equally abundant [26,33–36]. If mapped across transposon transcripts, piRNAs typically cover the entire sequence [31,32]. No obvious patterns suggest preferential processing from certain regions, indicating that RNA structure does not trigger piRNA biogenesis. However, in several cases piRNA profiles exhibit

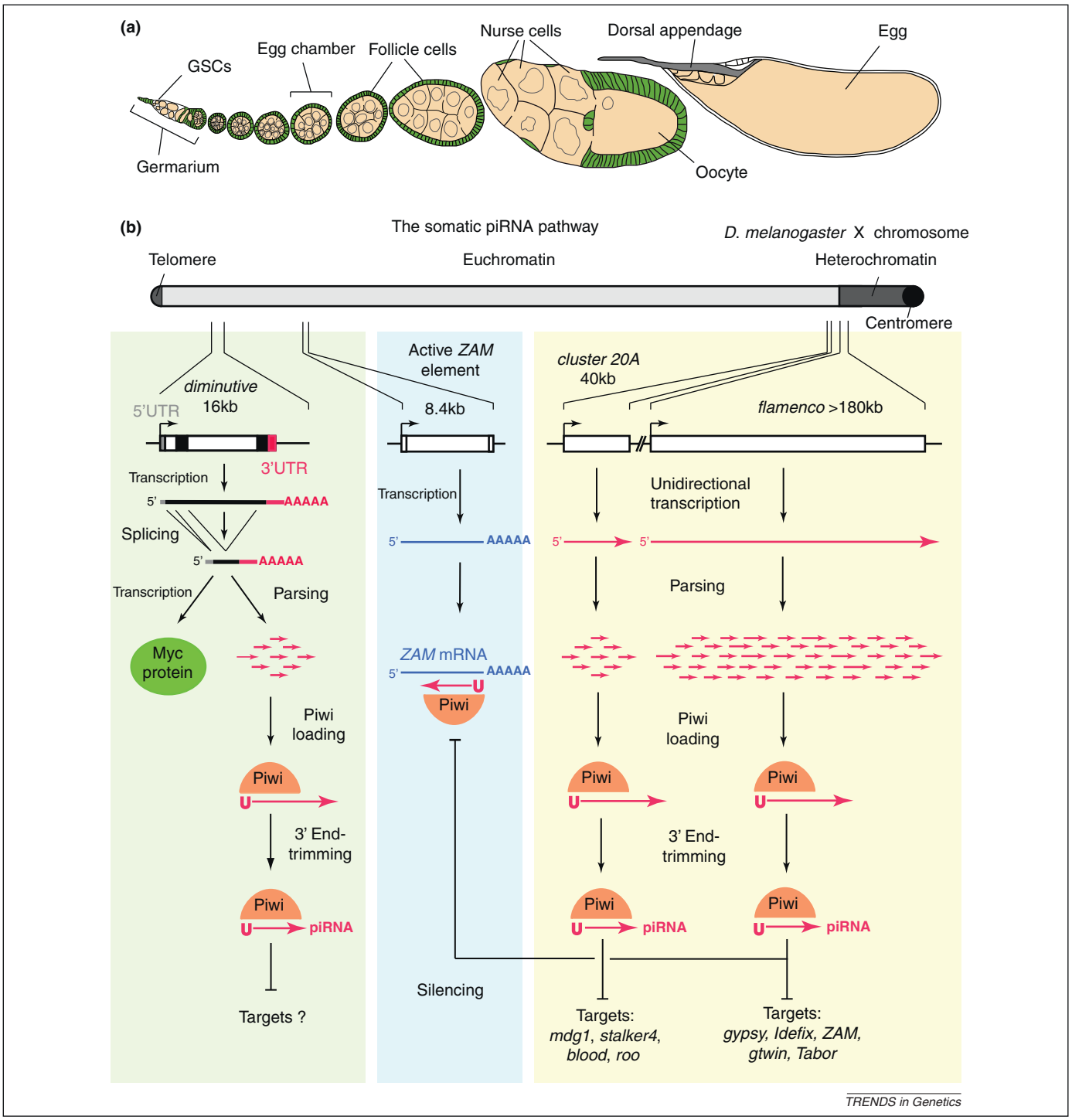

Figure 1. A primary piRNA pathway is active in somatic cells of the *Drosophila* ovary. **(a)** The *Drosophila* oocyte is in direct contact with germline-derived cells (beige) and is surrounded by cells of somatic origin (green). This cartoon depicts an ovariole, the functional unit of the ovary (reproduced with kind permission from A. Spradling [110]). Development proceeds from left (germarium) to right (mature egg). In the germarium, germline stem cells (GSCs) divide asymmetrically into GSCs and differentiating cystoblasts. Four mitotic cystoblast cell divisions produce 15 nurse cells and an oocyte that remain connected by cytoplasmic bridges. Each of these germline cell clusters is surrounded by an epithelium of somatic follicle cells (green) to form an egg chamber that continuously grows until the oocyte matures into an egg. Follicle cells finally undergo apoptosis after depositing the eggshell. The deposited egg therefore lacks somatic cells. **(b)** A schematic representation of the somatic piRNA pathway (primary piRNA module). For illustrative purposes, piRNA source and target loci from the X chromosome (drawn at the top) are displayed. Colored boxes summarize primary piRNA biogenesis from piRNA clusters (yellow) and from 3′ UTRs of protein coding genes (green). The blue box shows expression and silencing of *ZAM*, a prototypical LTR-retrotransposon, active in follicle cells. Yellow box: the piRNA clusters *20A* and *flamenco* are located at the boundary between euchromatin and heterochromatin. Both contain almost exclusively transposon fragments oriented antisense to the unidirectional promoter. piRNA cluster transcripts (red) therefore give rise to antisense piRNAs. Unknown mechanisms parse piRNA cluster transcripts into shorter fragments that are loaded onto Piwi. At this step, Piwi probably selects precursors with a 5′ uridine (1U RNA). Subsequently, the 3′ tail of Piwi-bound RNAs is trimmed and 2′-OH-methylated to generate mature piRNAs. Blue box: the sequence of mature piRNAs defines their target. Displayed is an active copy of the *ZAM* LTR-retrotransposon and its sense transcript being silenced by a piRNA. Green box: piRNA biogenesis from genes (here *diminutive*, *dm*). The spliced *dm* transcript with 5′ UTR, coding sequence and 3′ UTR is shown. Mature *dm* mRNAs are either translated to generate Myc protein or act as piRNA precursors. piRNAs are preferentially processed from 3′ UTR sequences, presumably by a similar mechanism as for piRNA cluster transcripts. Genic piRNAs are in sense orientation to the host gene and their targets (if any) remain to be identified.

Box 3. Transposable elements in *Drosophila*

The 180 Mb *Drosophila melanogaster* genome harbors over 100 transposon families, and members of all major classes (LINE- and LTR-type retroelements as well as DNA-type elements) are represented [98,99]. Release 5 of the assembled genome contains 117 Mb of euchromatic and ~24 Mb of heterochromatic sequence. Assembly of the heterochromatic portion was an important prerequisite for identifying piRNA clusters [100,101]. Annotated transposon sequences make up 7% of euchromatin and ~75% of heterochromatin. Most heterochromatic insertions, however, are sequence fragments and around two thirds of the 560 full-length (and thus presumably active) elements are located in euchromatin. The genomes of other drosophilid species contain similar transposon families. Both transposon load and their individual sequences are, however, typically species-specific [102]. Consequently, piRNA clusters differ in their content and serve as species-specific repositories of transposons that are, or have been, active in a population [31].

pronounced boundaries, and some transposon regions do not give rise to piRNAs. Insights into the genomic origin of piRNAs offer a coherent explanation for both the antisense bias of piRNAs and the irregularities of piRNA profiles across certain elements [20,31].

About 15% of somatic transposon-derived piRNAs map to unique sites in the genome, thus allowing the confident identification of their genomic origins [20]. This analysis has led to two conclusions. First, piRNAs appear to originate predominantly from inactive transposon copies, or transposon fragments, rather than from active copies. Second, piRNA-generating transposon sequences are densely packed in a few genomic loci. These so-called piRNA clusters span dozens to hundreds of kilobases in length, encompassing regions with the highest density of broken, mutated, and therefore immobile, transposon fragments in the entire genome [20,37]. piRNA clusters are a conserved hallmark of piRNA pathways although their repeat content varies widely among species [38–41]. Two piRNA clusters dominate in somatic ovarian cells; both are located on the X chromosome, and roughly at the euchromatin/heterochromatin boundary (Figure 1b). The larger one is the genetically identified *flamenco* locus, the smaller is referred to as *cluster 20A* according to its cytology [20,21,28–31]. piRNAs from both clusters are derived only from one genomic DNA strand, arguing for unidirectional transcription. Moreover, ~90% of the transposon fragments in *flamenco*, and all of those in *cluster 20A*, are oriented antisense to the direction of transcription. This immediately explains the massive antisense bias of transposon-derived piRNAs. The *flamenco* cluster appears to be expressed and processed only in somatic ovarian cells [31]. Interestingly, most transposon fragments in *flamenco* belong to the *gypsy* family of retrotransposons, precisely those that invade the germline via the somatic niche [16–19,31].

These observations have led to a model in which the somatic piRNA pathway stores transposon sequence information in specialized genomic regions. Their unidirectional transcripts are parsed into piRNAs, which – after loading into Piwi – allow *trans* silencing of complementary transposons [30,31]. In this scenario, insertions of *gypsy*-type elements antisense to the direction of cluster transcription were positively selected during evolution. Strong support for this model stems from an analysis of fragments of the *ZAM* retroelement that are located within *flamenco*: only the sequence regions of *ZAM* that are found within *flamenco* give rise to abundant piRNAs [31]. piRNA production in the soma should therefore be independent of the expression of active elements. Indeed, levels of *gypsy*-derived somatic piRNAs are not influenced by the presence of active *gypsy* elements in the genome [42].

In addition to *flamenco* and *cluster 20A*, several other, often smaller, piRNA clusters have been cataloged based on OSS piRNAs [21]. These are, however, not yet assembled into chromosomal contigs. We note that *flamenco* ends in a genomic gap of unknown size and that some of these cluster fragments might therefore correspond to pieces of *flamenco*.

piRNAs from exons

Based on the suspicious localization of piRNA clusters at the euchromatin/heterochromatin boundary, one might infer that certain chromatin marks might flag cluster transcripts for piRNA biogenesis. This model was challenged with the surprising discovery that transcripts from several hundred genes are substrates for piRNA biogenesis and are the source of nearly 10% of somatic piRNAs [22,43]. About 95% of genic piRNAs are in the sense orientation to the host transcript and typically originate from exons, indicating that mature mRNAs are the substrates for processing.

Only a subset of cellular mRNAs give rise to piRNAs, and there is no apparent correlation between transcript abundance and piRNA levels [43]. Because exonic sense piRNAs will typically have no fully complementary targets within the cell, their function is unknown. It has been suggested that some exonic piRNAs target cellular transcripts via incomplete pairing [22]. The predicted target sites, however, are located in the intron of the target gene. It remains to be shown whether this allows significant target regulation and how tolerant target recognition is towards incomplete complementarity between small RNA and target. Alternatively, the cell modulates the expression of host genes because some mRNA transcripts are consumed during piRNA biogenesis [43]. Interestingly, the gene giving rise to most piRNAs encodes the transcription factor Traffic jam which is required for Piwi expression in somatic support cells, suggestive of a classic negative-feedback loop [22,43].

Primary piRNA biogenesis

The linear biogenesis of piRNAs, from precursor transcripts to complexes with PIWI proteins, has been termed 'primary piRNA biogenesis' [20]. piRNAs in somatic support cells of the gonad seem to be exclusively generated via primary processing. Little is known about this process at the mechanistic level. The above-mentioned features of piRNAs strongly suggest that single-stranded transcripts (originating from piRNA clusters and genes) are substrates for the processing machinery. A P-element insertion at the beginning of *flamenco* abrogates piRNA production over the entire 180 kb cluster, strongly arguing for a long, single-stranded transcript [20,30,31]. Moreover, piRNA biogenesis is independent of Dicer

[26], the key enzyme in the miRNA and siRNA pathways where dsRNAs serve as trigger molecules. piRNA profiles across exons or clusters do not correlate with any obvious RNA secondary structure features, although pronounced peaks of 'genome-unique' piRNAs across clusters are apparent. The genomic spread of transposons dictates that most sequence fragments within a piRNA cluster often have identical copies elsewhere in the genome, and are thus repetitive. By comparison, areas of genome-unique sequence are rare and locally dispersed within piRNA clusters. This distribution of genome-unique and repetitive areas correlates well with the peaks and valleys in piRNA profiles across clusters, arguing against local differences in piRNA biogenesis from cluster transcripts.

It is entirely unclear how the cell selects cluster transcripts and those from a subset of genes for piRNA biogenesis. Are these transcripts marked in any special way for biogenesis? A starting point to answer this question might well reside in the pool of genic piRNAs. In some cases, the level of genic piRNAs per kilobase approaches that of *flamenco*-derived piRNAs, indicating that genic piRNAs are not merely noisy by-products of cellular RNA metabolism. These transcripts are somehow special, perhaps because of the sequence motifs they contain or by virtue of features such as RNA half-life or translation efficiency; it will be important to decipher the underlying molecular basis for this.

The precise subcellular location for piRNA biogenesis is unknown. Although Piwi is enriched in the nucleus, accumulating evidence suggests that primary piRNA biogenesis takes place in the cytoplasm. First, an overwhelming proportion of genic piRNAs originate from the 3′ untranslated region (3′ UTR), and the first piRNAs map almost directly downstream of the stop codon [22,43]. This indicates that ribosomal association precedes piRNA processing. Second, an N-terminally truncated Piwi protein that cannot localize to the nucleus is loaded efficiently with piRNAs [22].

Although variable in length, piRNAs with the same core sequence typically share the same 5′ end, but differ at their 3′ ends. The first base shows a strong bias for uridine. 5′ nucleotide preferences are common among Argonaute proteins [44–46], and a recent study indicated that Argonaute proteins can read the identity of the first base of the bound RNA [47]. Taken together, these findings suggest a random-processing model in which initially longer piRNA precursor transcripts are first stochastically parsed into smaller RNA fragments. It seems likely that Piwi selectively binds or stabilizes those RNA fragments (or pre-piRNAs) that start with a 5′ uridine. This might well explain observed local irregularities in piRNA profiles. In a final step, 3′ trimming of the pre-piRNA would generate the final piRNA. The footprint of Piwi on the pre-piRNA would determine piRNA length, explaining why piRNA populations bound to different PIWI family proteins differ in length [20]. According to this model, piRNA precursors are loaded onto Piwi as single-stranded RNAs. This is in contrast to siRNAs and miRNAs, which are loaded as small RNA duplexes into Argonaute proteins, after which one strand is removed [48–52].

A recent study in *Schizosaccharomyces pombe*, however, indicates that the proposed piRNA biogenesis model might not be that exotic – although fission yeast Argonaute is primarily loaded with Dicer products (small RNA duplexes), it appears that initially it is loaded with so-called primal RNAs [53]. Strikingly, primal RNAs are preferentially derived from the 3′ UTRs of cellular transcripts and centromeric repeats in a Dicer-independent manner and appear to be trimmed at their 3′ ends, potentially by the exosome [53]. Following target interaction and cleavage, an RNA-dependent RNA polymerase (RdRP) converts the target transcript into dsRNA, providing the substrate for Dicer processing into the much more abundant siRNAs. The resemblance of primal RNAs to primary piRNAs is provacative, although RdRP-dependent amplification does not seem to participate in the piRNA pathway.

The proteins involved in primary piRNA biogenesis are unknown, with the exception of Zucchini, a predicted nuclease with a phospholipase D domain [22,31,54]. The step at which Zucchini acts during biogenesis is, however, unclear. Although multiple other proteins have been identified as essential piRNA pathway members, genetic studies indicate that these are selectively involved in the more complex biology of the germline piRNA pathway [31,32,55]. Undoubtedly, several yet-to-be-identified factors must participate in biogenesis, loading and function of primary piRNAs, and the availability of the OSS cell line [56] promises rapid progress towards their identification and characterization.

The germline piRNA pathway and piRNA ping–pong

Considerable evidence indicates that the linear primary piRNA pathway feeding into Piwi is also active in germline cells [31,32]. piRNA biology in the germline is, however, much more complex. In addition to Piwi, ovarian germline cells also express Aubergine and AGO3, two related PIWI family proteins [20,32,57,58]. Sequence analysis of piRNAs selectively bound to Piwi, Aub and AGO3 revealed the existence of a sophisticated piRNA amplification loop that acts in parallel or in addition to the primary piRNA pathway described above (Figure 2). The central players in this so-called ping–pong cycle are Aub and AGO3, which localize to the cytoplasm of germline cells with an accumulation around the nucleus [20,58]. In the prevailing model, Aub is guided via an antisense piRNA to a sense transcript from an active transposon. Subsequent slicer cleavage of the target transcript triggers production of a novel sense piRNA, which is loaded into AGO3. The AGO3–piRNA complex in turn cleaves complementary piRNA cluster transcripts. This prompts biogenesis of a novel antisense and Aub-bound piRNA, whose sequence is identical to that of the initiator piRNA. Because Aub and AGO3 presumably act catalytically [58], the ping–pong cycle amplifies silencing-competent piRNAs, with the loop acting efficiently only in the presence of a target transcript (active transposon message). Indeed, ping–pong piRNAs are the most abundant cellular piRNAs [20]. A key conceptual difference from the primary piRNA pathway is that piRNA biogenesis in the germline depends, in part, on target expression. Elegant genetic experiments hinted at

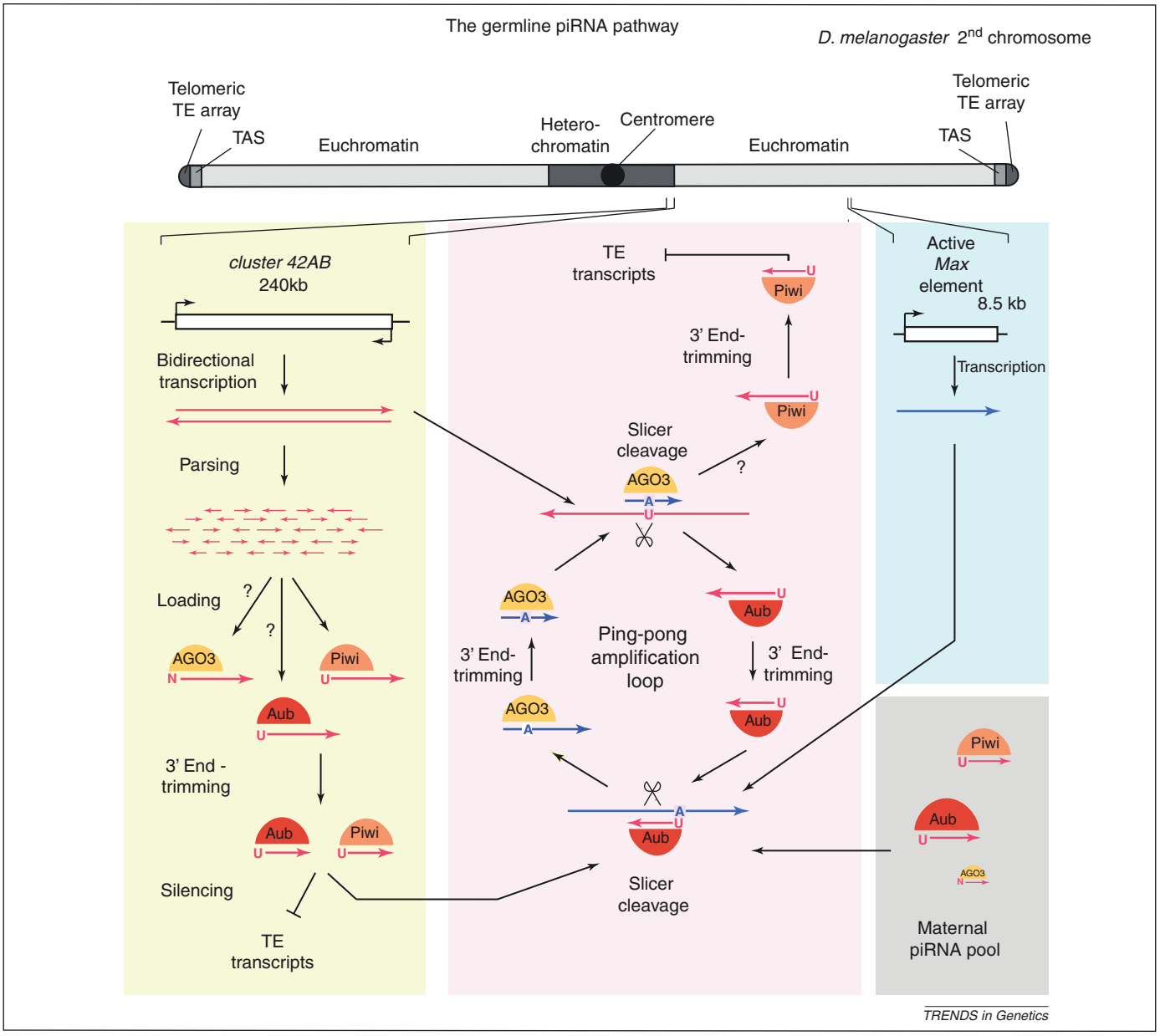

Figure 2. The primary piRNA pathway and the ping–pong amplification loop are active in germline cells. Representative examples of germline piRNA sources and targets originating from the 2nd chromosome are shown. Colored boxes show primary piRNA biogenesis from *cluster 42AB* (yellow), the adaptive module of the target-dependent ping–pong amplification loop (red), expression and silencing of a typical active LTR-retrotransposon (*Max* element; blue), and the contribution of maternally inherited piRNAs (gray). Yellow box: *cluster 42AB* contains transposon fragments in both orientations and is bidirectionally transcribed. During primary piRNA biogenesis, cluster transcripts (red) presumably generate sense and antisense piRNAs. Unknown mechanisms parse the long piRNA precursor transcripts into shorter fragments that are loaded onto PIWI family proteins (Piwi, probably Aub and potentially AGO3). Piwi and Aub probably select RNA fragments with a 5′ uridine (1U RNA). Subsequently, the 3′ tails of pre-piRNAs are trimmed and 2′-OH-methylated to generate mature piRNAs. piRISCs with antisense piRNAs are competent to silence sequence complementary transcripts of active transposons (TE). Primary piRNA biogenesis in the germline is probably similar to the one in somatic cells. Blue box: an active copy of the *Max* element (LTR-retrotransposon) and its transcribed sense transcript are silenced by complementary piRISCs. Red box: schematic representation of the ping–pong cycle. An Aub-complexed piRNA (red) that is antisense to an active sense *Max* transcript (dark blue) guides slicing (scissors) of the transposon RNA precisely 10 nt downstream of its 5′ uridine. The sliced *Max* transcript is predicted to be loaded onto AGO3 and typically has a profound bias for an adenine at position 10 (10A RNA). The AGO3-bound pre-piRNA is 3′-trimmed and 2′-OH-methylated. This mature AGO3–piRNA complex in turn cleaves complementary cluster transcripts and triggers production of an Aub-loaded antisense piRNA whose sequence is identical to the initiating piRNA. It is currently impossible to experimentally distinguish between Aub–piRNA complexes generated via primary piRNA biogenesis or via ping–pong. Weak ping–pong signatures are shared between AGO3 and Piwi and could indicate that Piwi (in addition to primary biogenesis) also receives piRNAs via AGO3-mediated target slicing. Grey box: at the end of oogenesis, mature Aub– and Piwi–piRNA complexes (to a lesser extent also AGO3 complexes) are efficiently loaded into the oocyte. Maternal Aub, and to a lesser extent also Piwi, localize to the posterior pole of the mature egg where the future germline will form. Maternal piRNAs might serve important roles in the starting phase of the ping–pong cycle.

this even before piRNAs were discovered [59,60]. Signatures of the ping–pong cycle have been found in sponges, planaria, moths, fish, frogs and mammals [40,61–65] (Box 4). It is therefore one of the hallmarks of the piRNA pathway. Target-dependent amplification of a small RNA response is common among small RNA pathways in fungi, plants and nematodes [9]. Here, however, cleavage of the target triggers dsRNA synthesis by RdRP enzymes, generating Dicer substrates. Most animals lack RdRP enzymes, and the piRNA pathway instead utilizes sense and antisense RNAs from different transcripts and couples these via reciprocal Slicer cleavage.

Box 4. Commonalities and differences among animal piRNA pathways

Based on primary sequence analyses, PIWI proteins are found throughout the animal kingdom. They are typically expressed in germline cells but are also found in cells responsible for regeneration in lower invertebrates (e.g. neoblasts in planaria [62,103]). Primary piRNA biogenesis and signatures of the ping–pong amplification cycle are found in species ranging from sponges to mammals, indicating an ancient origin of the key features of the pathway [61]. A notable exception is *C. elegans*, where the two PIWI family proteins PRG1 and PRG2 are expressed in gonads, but bind 21U RNAs, a different class of small RNAs [45,95–97]. These appear to have distinct biogenesis features, do not exhibit ping–pong signatures, and have a largely unclear function and target spectrum.

In mammals, three PIWI family proteins are expressed in testes and only one in ovaries [104–107]. Interestingly, the pathway seems largely dispensable for oogenesis, where an endo-siRNA pathway centered on AGO proteins cooperates in silencing transposable elements [93,94]. During mouse spermatogenesis, the three PIWI family proteins MIWI2, MILI and MIWI are expressed in different, but overlapping, temporal domains. MIWI2 and MILI are the key players in the genome defense pathway and their bound piRNAs show signatures of ping–pong and primary piRNA biogenesis [108,109]. MIWI2, the only nuclear mouse PIWI family protein, is suggested to guide *de novo* DNA methylation at transposon loci, a process so far only reported for the mammalian piRNA pathway [108,109]. MIWI is expressed only after the pachytene stage of meiosis and binds primary piRNAs derived from a distinct set of piRNA clusters. Pachytene piRNAs are not enriched in transposon sequences, accumulate to very high levels, and have an unclear function, but this is presumably distinct from transposon silencing [38,39,41].

Germline piRNA clusters

RNAs in early embryos (before the onset of zygote genome transcription) by and large reflect the pool of germline transcripts generated during oogenesis [31,66]. piRNAs from young embryos are thus the best proxy for the germline-specific piRNA pool. Germline piRNAs originate predominantly from several piRNA clusters, but also arise from transcripts of active elements, in accordance with the ping–pong model. Evidence for this is best documented for the *I*-element [a retroelement of the long interspersed nuclear element (LINE) family], where cluster-resident fragments and active elements have diverged sufficiently at the sequence level to allow them to be distinguished [66].

Similar to somatic clusters, germline clusters are strongly enriched in transposon fragments; the most prominent map to euchromatin-heterochromatin boundaries [20]. Because a much broader spectrum of transposable elements [LINE-, long terminal repeat (LTR)- and DNA-type elements] is highly active in the germline, it comes as no surprise that germline piRNA clusters contain a more diverse collection of transposon fragments [31]. Figure 3 depicts the soma-specific *flamenco* cluster, the germline-specific *cluster 42AB*, and the shared *cluster 20A*; *cluster 20A* is the only germline piRNA cluster that resembles *flamenco* in that it is unidirectionally transcribed and contains only antisense transposon fragments [20,31]. In the germline it preferentially loads Piwi, and to a lesser extent, also Aub. All other germline clusters spawn piRNAs from both strands, indicating bidirectional transcription and alleviating the pressure for transposons to integrate in a biased orientation as observed for *flamenco* [31]. piRNAs originating from these clusters are loaded into all three PIWI family proteins, but absolute numbers cannot be derived because only a minority of piRNAs can unambiguously be mapped to clusters. Finally, germline piRNA clusters are also found at telomeres, where abundant piRNAs are derived from the telomeric arrays of *HeT-A*, *TART* and *TAHRE* elements as well as from subtelomeric satellite repeats [20].

It is unclear whether piRNA cluster transcripts are essential for ping–pong or whether any sense and antisense transcripts could engage in it. Similarly unknown is how the cell prevents auto-amplification of piRNAs derived from bidirectionally transcribed clusters. The best evidence for this stems from the analysis of *I*-element piRNAs: *I*-element fragments within *cluster 42AB* give rise to high piRNA levels only in strains with active elements [66]. It is possible that cluster transcripts are physically isolated in specific subcellular domains and are only presented to specific protein complexes, thereby guaranteeing accurate progression through the biogenesis cycle.

The connections between primary piRNA biogenesis and the ping–pong cycles are also only poorly understood. The ping–pong signature is mostly confined to Aub-AGO3 and Aub-Aub pairs [20,32]. Ping–pong is almost entirely lost in *aub* mutants [31]. Whereas Aub-Aub ping–pong prevails in *AGO3* mutants, the resulting piRNA levels are severely diminished [32]. Piwi, by contrast, is dispensable for the ping–pong cycle [31]. Nevertheless, a weak but significant ping–pong interaction takes place between Piwi and AGO3 [20,32], suggesting that Piwi – although not providing input – could be a recipient in the cycle (Figure 2). This might explain the antisense bias of Piwi-bound piRNAs originating from germline clusters [20] and could also explain the loss of nuclear Piwi in late-stage late-stage egg chambers lacking AGO3 [32]. In such a model, primary piRNA biogenesis must also load Aub or AGO3. In fact, germline cells in *Drosophila* testes express only Aub and AGO3, whereas Piwi is expressed in somatic support cells only [67]. It therefore remains to be determined how primary piRNA biogenesis feeds into the ping–pong cycle.

The maternal piRNA pool

During oogenesis the oocyte is connected to the 15 nurse cells via cytoplasmic bridges. Ultimately, nurse cells transfer their cytoplasm to the maturing oocyte (Figure 1a). Piwi and Aub localize to the posterior pole of the oocyte where the primordial germ cells of the embryo will form [57,66,68]. Maternal piRNAs thus provide a protective layer against transposons even before transcription initiates in future germ cells. Genetic experiments have suggested the existence of a maternal factor with essential roles in transposon silencing [3,4]. In this so-called hybrid dysgenesis phenomenon, crosses between naïve females and males carrying a novel transposon produce sterile offspring, whereas the reciprocal cross does not. Small RNA sequencing approaches have shown that inheritance of maternal piRNAs is required for an efficient ping–pong response in the F1 generation [66]. Three possibilities could explain this observation: (i) maternal piRNAs are required to kick-start the ping–pong cycle; (ii) maternal piRNAs have an essential influence on the chromatin status of piRNA clusters and/or transposons; and

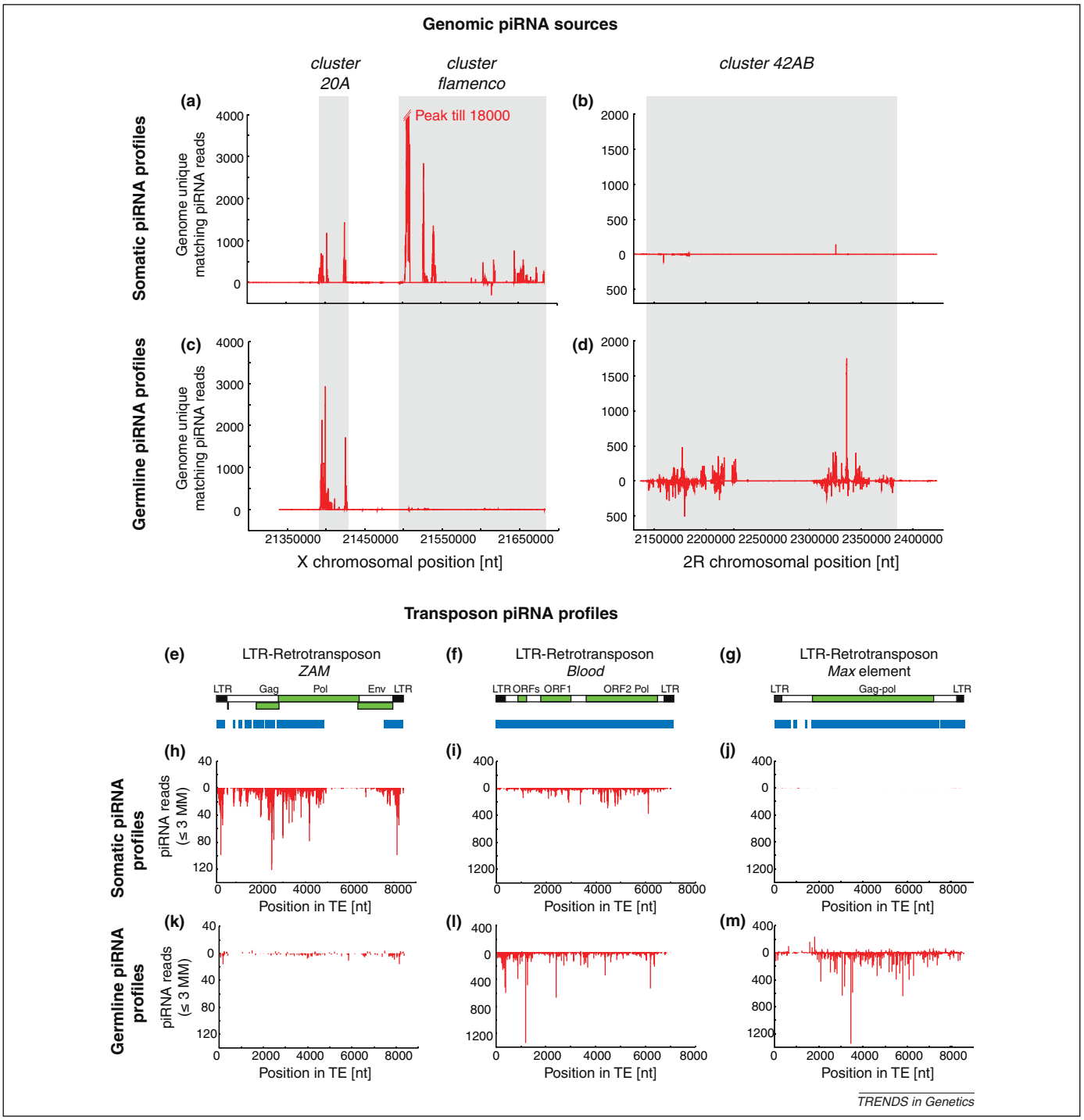

Figure 3. piRNA profiles along clusters and transposons are tissue-specific. There are pronounced differences between the piRNA pools found in the somatic and germline cells of the *Drosophila* ovary. Bases for these differences are tissue-specific transcription of piRNA clusters and the presence of the ping–pong cycle in germline cells only. Somatic graphs are based on the OSS cell data from [43], and germline graphs are based on early embryo libraries from [66]. To enable comparison of these populations, profiles were normalized to 1 million sequenced repeat-derived 23–30 nt small RNAs. **(a–d)** These profiles indicate that *flamenco* is a soma-specific piRNA cluster whereas *cluster 42AB* is germline-specific (cluster coordinates are shaded in light grey). *cluster 20A* is processed into piRNAs in both cell types. Also apparent is the unidirectional transcription of *flamenco* and *cluster 20A* whereas *cluster 42AB* is transcribed in both orientations. In each panel, only genome-unique piRNAs were used and a 200 nt sliding window with step size of 20 nt was applied. Sense and antisense piRNAs are displayed as upwards and downwards peaks, respectively. **(e–g)** Schematics of the LTR-retrotransposons *ZAM*, *Blood* and the *Max* element. Blue bars display the respective transposon fragments found in piRNA clusters (antisense *ZAM* fragments within *flamenco*, a complete antisense *Blood* element in *cluster20A*, and *Max* fragments in *cluster 42AB*). Transposon cartoons and cluster fragments are length-matched to the piRNA profiles shown below. **(h–m)** Profiles of somatic and germline piRNAs mapping to *ZAM*, *Blood* and *Max*. For each graph, piRNAs mapping with up to three mismatches to the indicated element were pooled. *ZAM* is a prototypical element expressed and silenced in somatic cells, whereas *Max* is apparently only silenced (and presumably transcribed) in germline cells. *Blood* silencing is active in both cell types. The *ZAM* fragments present in the *flamenco* piRNA cluster (blue) are in striking agreement with the observed piRNA profiles, suggesting that they are the major source of *ZAM* piRNAs. Similarly, piRNA profiles for *Blood* and *Max* are consistent with their respective fragments in piRNA *clusters 20A* and *42AB*. Ping–pong signatures are significant only for *Blood* and *Max* in the germline samples (not shown). ORF: open reading frame.

(iii) the cellular Aub and AGO3 pools are limiting and, in the absence of maternal piRNAs for a certain element, the low levels of primary piRNAs are unable to compete with piRNAs abundantly inherited maternally.

In summary, the germline piRNA pathway is considerably more elaborate than the linear somatic piRNA pathway. Here we also largely lack insight into the molecular and cellular details. Genetic studies have identified multiple proteins involved in the piRNA pathway [23,31,32,54,55,57,69–78], and several are specifically required for the germline piRNA pathway [31,32,55]. Among these are several RNA helicases, but they also include proteins of unknown function. Recent studies have linked the piRNA pathway to Tudor biology [79]. Tudor domains bind symmetrically methylated arginines in Aub, AGO3, and potentially Piwi [80–82]. The *Drosophila* genome encodes at least 20 proteins containing Tudor domains and many of these proteins are selectively expressed in the germline (K-A.S. and J.B., unpublished observations). Given this, and considering the complexity of piRNA biology, we expect the number of proteins with essential roles in this genome defense system to increase considerably.

Concluding remarks

Research in the piRNA field is in a paradoxical situation where we understand a lot about conceptual frameworks, but lack almost every in terms of mechanistic and molecular insight. A great deal of understanding from other small RNA pathways has emerged from *in vitro* assays. No such attempt has been reported for the piRNA pathway, probably reflecting the complexity of this approach. Without a doubt, *in vitro* systems, coupled with genetics and structural approaches, will be essential for understanding the order and dynamics of the molecular events during piRNA biogenesis and silencing.

A second challenge will be to understand the nature of piRNA cluster transcripts and to decipher the protein–RNA network that forms the basis of the piRNA pathway. Here we expect rapid progress by combinations of next-generation sequencing approaches with technological advances in determining RNA–protein interactions [83,84]. Overall, these are exciting days for everybody working on this fascinating genome defense system, and, if recent years are a measure, many surprises are yet to come.

Acknowledgements

This work was supported by a European Commission Marie Curie postdoctoral fellowship to K.A.S. A European Research Council (ERC) Starting Grant from the European Commission supports work in the Brennecke laboratory.

References

1 Hurst, G.D. and Werren, J.H. (2001) The role of selfish genetic elements in eukaryotic evolution. *Nat. Rev. Genet.* 2, 597–606
2 Slotkin, R.K. and Martienssen, R. (2007) Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* 8, 272–285
3 Bucheton, A. *et al.* (1984) The molecular basis of I–R hybrid dysgenesis in *Drosophila melanogaster*: identification, cloning, and properties of the I factor. *Cell* 38, 153–163
4 Rubin, G.M. *et al.* (1982) The molecular basis of P–M hybrid dysgenesis: the nature of induced mutations. *Cell* 29, 987–994
5 Fire, A. *et al.* (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811
6 Hamilton, A.J. and Baulcombe, D.C. (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286, 950–952
7 Zamore, P.D. *et al.* (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* 101, 25–33
8 Hammond, S.M. *et al.* (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293–296
9 Girard, A. and Hannon, G.J. (2008) Conserved themes in small-RNA-mediated transposon control. *Trends Cell Biol.* 18, 136–148
10 Malone, C.D. and Hannon, G.J. (2009) Small RNAs as guardians of the genome. *Cell* 136, 656–668
11 Tomari, Y. and Zamore, P.D. (2005) Perspective: machines for RNAi. *Genes Dev.* 19, 517–529
12 Okamura, K. and Lai, E.C. (2008) Endogenous small interfering RNAs in animals. *Nat. Rev. Mol. Cell Biol.* 9, 673–678
13 Ghildiyal, M. and Zamore, P.D. (2009) Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.* 10, 94–108
14 Klattenhoff, C. and Theurkauf, W. (2008) Biogenesis and germline functions of piRNAs. *Development* 135, 3–9
15 Bucheton, A. (1990) I transposable elements and I–R hybrid dysgenesis in *Drosophila*. *Trends Genet.* 6, 16–21
16 Chalvet, F. *et al.* (1999) Proviral amplification of the *Gypsy* endogenous retrovirus of *Drosophila melanogaster* involves env-independent invasion of the female germline. *EMBO J.* 18, 2659–2669
17 Leblanc, P. *et al.* (2000) Life cycle of an endogenous retrovirus, *ZAM*, in *Drosophila melanogaster*. *J. Virol.* 74, 10658–10669
18 Pelisson, A. *et al.* (1994) *Gypsy* transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the *Drosophila flamenco* gene. *EMBO J.* 13, 4401–4411
19 Brasset, E. *et al.* (2006) Viral particles of the endogenous retrovirus *ZAM* from *Drosophila melanogaster* use a pre-existing endosome/exosome pathway for transfer to the oocyte. *Retrovirology* 3, 25
20 Brennecke, J. *et al.* (2007) Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128, 1089–1103
21 Lau, N.C. *et al.* (2009) Abundant primary piRNAs, endo-siRNAs, and microRNAs in a *Drosophila* ovary cell line. *Genome Res.* 19, 1776–1785
22 Saito, K. *et al.* (2009) A regulatory circuit for *piwi* by the large Maf gene *traffic jam* in *Drosophila*. *Nature* 461, 1296–1299
23 Cox, D.N. *et al.* (1998) A novel class of evolutionarily conserved genes defined by *piwi* are essential for stem cell self-renewal. *Genes Dev.* 12, 3715–3727
24 Cox, D.N. *et al.* (2000) *piwi* encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. *Development* 127, 503–514
25 Saito, K. *et al.* (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev.* 20, 2214–2222
26 Vagin, V.V. *et al.* (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* 313, 320–324
27 Yin, H. and Lin, H. (2007) An epigenetic activation role of Piwi and a Piwi-associated piRNA in *Drosophila melanogaster*. *Nature* 450, 304–308
28 Prud'homme, N. *et al.* (1995) *Flamenco*, a gene controlling the *gypsy* retrovirus of *Drosophila melanogaster*. *Genetics* 139, 697–711
29 Desset, S. *et al.* (2003) *COM*, a heterochromatic locus governing the control of independent endogenous retroviruses from *Drosophila melanogaster*. *Genetics* 164, 501–509
30 Mevel-Ninio, M. *et al.* (2007) The *flamenco* locus controls the *gypsy* and *ZAM* retroviruses and is required for *Drosophila* oogenesis. *Genetics* 175, 1615–1624
31 Malone, C.D. *et al.* (2009) Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* 137, 522–535
32 Li, C. *et al.* (2009) Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell* 137, 509–521
33 Ghildiyal, M. *et al.* (2008) Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* 320, 1077–1081
34 Czech, B. *et al.* (2008) An endogenous small interfering RNA pathway in *Drosophila*. *Nature* 453, 798–802

35 Chung, W.J. *et al.* (2008) Endogenous RNA interference provides a somatic defense against *Drosophila* transposons. *Curr. Biol.* 18, 795–802
36 Kawamura, Y. *et al.* (2008) *Drosophila* endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature* 453, 793–797
37 Bergman, C.M. *et al.* (2006) Recurrent insertion and duplication generate networks of transposable element sequences in the *Drosophila melanogaster* genome. *Genome Biol.* 7, R112
38 Lau, N.C. *et al.* (2006) Characterization of the piRNA complex from rat testes. *Science* 313, 363–367
39 Aravin, A. *et al.* (2006) A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* 442, 203–207
40 Aravin, A.A. *et al.* (2007) Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* 316, 744–747
41 Girard, A. *et al.* (2006) A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* 442, 199–202
42 Pelisson, A. *et al.* (2007) A novel repeat-associated small interfering RNA-mediated silencing pathway downregulates complementary sense *gypsy* transcripts in somatic cells of the *Drosophila* ovary. *J. Virol.* 81, 1951–1960
43 Robine, N. *et al.* (2009) A broadly conserved pathway generates 3′UTR-directed primary piRNAs. *Curr. Biol.* 19, 2066–2076
44 Mi, S. *et al.* (2008) Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5′ terminal nucleotide. *Cell* 133, 116–127
45 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
46 Ghildiyal, M. *et al.* (2010) Sorting of *Drosophila* small silencing RNAs partitions microRNA* strands into the RNA interference pathway. *RNA* 16, 43–56
47 Frank, F. *et al.* (2010) Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature* 465, 818–822
48 Matranga, C. *et al.* (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* 123, 607–620
49 Miyoshi, K. *et al.* (2005) Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. *Genes Dev.* 19, 2837–2848
50 Rand, T.A. *et al.* (2005) Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell* 123, 621–629
51 Leuschner, P.J. *et al.* (2006) Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Rep.* 7, 314–320
52 Maiti, M. *et al.* (2007) QIP, a putative exonuclease, interacts with the *Neurospora* Argonaute protein and facilitates conversion of duplex siRNA into single strands. *Genes Dev.* 21, 590–600
53 Halic, M. and Moazed, D. (2010) Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. *Cell* 140, 504–516
54 Pane, A. *et al.* (2007) *zucchini* and *squash* encode two putative nucleases required for rasiRNA production in the *Drosophila* germline. *Dev. Cell* 12, 851–862
55 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
56 Niki, Y. *et al.* (2006) Establishment of stable cell lines of *Drosophila* germ-line stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 103, 16325–16330
57 Harris, A.N. and Macdonald, P.M. (2001) *Aubergine* encodes a *Drosophila* polar granule component required for pole cell formation and related to eIF2C. *Development* 128, 2823–2832
58 Gunawardane, L.S. *et al.* (2007) A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. *Science* 315, 1587–1590
59 Jensen, S. *et al.* (1999) Taming of transposable elements by homology-dependent gene silencing. *Nat. Genet.* 21, 209–212
60 Jensen, S. *et al.* (2002) Regulation of I-transposon activity in *Drosophila*: evidence for cosuppression of nonhomologous transgenes and possible role of ancestral I-related pericentromeric elements. *Genetics* 162, 1197–1209
61 Grimson, A. *et al.* (2008) Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* 455, 1193–1197
62 Friedlander, M.R. *et al.* (2009) High-resolution profiling and discovery of planarian small RNAs. *Proc. Natl. Acad. Sci. U. S. A.* 106, 11546–11551
63 Kawaoka, S. *et al.* (2009) The *Bombyx* ovary-derived cell line endogenously expresses PIWI/PIWI-interacting RNA complexes. *RNA* 15, 1258–1264
64 Houwing, S. *et al.* (2008) Zili is required for germ cell differentiation and meiosis in zebrafish. *EMBO J.* 27, 2702–2711
65 Lau, N.C. *et al.* (2009) Systematic and single cell analysis of *Xenopus* Piwi-interacting RNAs and Xiwi. *EMBO J.* 28, 2945–2958
66 Brennecke, J. *et al.* (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 322, 1387–1392
67 Nishida, K.M. *et al.* (2007) Gene silencing mechanisms mediated by Aubergine piRNA complexes in *Drosophila* male gonad. *RNA* 13, 1911–1922
68 Megosh, H.B. *et al.* (2006) The role of PIWI and the miRNA machinery in *Drosophila* germline determination. *Curr. Biol.* 16, 1884–1894
69 Anne, J. *et al.* (2007) Arginine methyltransferase Capsuleen is essential for methylation of spliceosomal Sm proteins and germ cell formation in *Drosophila*. *Development* 134, 137–146
70 Boswell, R.E. and Mahowald, A.P. (1985) *tudor*, a gene required for assembly of the germ plasm in *Drosophila melanogaster*. *Cell* 43, 97–104
71 Chen, Y. *et al.* (2007) *Cutoff* and *aubergine* mutations result in retrotransposon upregulation and checkpoint activation in *Drosophila*. *Curr. Biol.* 17, 637–642
72 Cook, H.A. *et al.* (2004) The *Drosophila* SDE3 homolog *armitage* is required for *oskar* mRNA silencing and embryonic axis specification. *Cell* 116, 817–829
73 Gillespie, D.E. and Berg, C.A. (1995) *Homeless* is required for RNA localization in *Drosophila* oogenesis and encodes a new member of the DE-H family of RNA-dependent ATPases. *Genes Dev.* 9, 2495–2508
74 Horwich, M.D. *et al.* (2007) The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr. Biol.* 17, 1265–1272
75 Lim, A.K. and Kai, T. (2007) Unique germ-line organelle, nuage, functions to repress selfish genetic elements in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 104, 6714–6719
76 Patil, V.S. and Kai, T. (2010) Repression of retroelements in *Drosophila* germline via piRNA pathway by the Tudor domain protein Tejas. *Curr. Biol.* DOI:10.1016/j.cub.2010.02.046
77 Saito, K. *et al.* (2007) Pimet, the *Drosophila* homolog of HEN1, mediates 2′-O-methylation of Piwi- interacting RNAs at their 3′ ends. *Genes Dev.* 21, 1603–1608
78 Vagin, V.V. *et al.* (2004) The RNA interference proteins and vasa locus are involved in the silencing of retrotransposons in the female germline of *Drosophila melanogaster*. *RNA Biol.* 1, 54–58
79 Siomi, M.C. *et al.* (2010) How does the royal family of Tudor rule the PIWI-interacting RNA pathway? *Genes Dev.* 24, 636–646
80 Kirino, Y. *et al.* (2009) Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. *Nat. Cell Biol.* 11, 652–658
81 Nishida, K.M. *et al.* (2009) Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in *Drosophila* germlines. *EMBO J.* 28, 3820–3831
82 Kirino, Y. *et al.* (2010) Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization. *RNA* 16, 70–78
83 Hafner, M. *et al.* (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141, 129–141
84 Chi, S.W. *et al.* (2009) Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps. *Nature* 460, 479–486
85 Lee, R.C. *et al.* (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843–854
86 Wightman, B. *et al.* (1993) Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855–862
87 Carmell, M.A. *et al.* (2002) The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev.* 16, 2733–2742
88 Tolia, N.H. and Joshua-Tor, L. (2007) Slicer and the argonautes. *Nat. Chem. Biol.* 3, 36–43
89 Gu, W. *et al.* (2009) Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the *C. elegans* germline. *Mol. Cell* 36, 231–244

90 Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297
91 Okamura, K. *et al.* (2008) The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* 453, 803–806
92 Rozhkov, N.V. *et al.* (2010) Small RNA-based silencing strategies for transposons in the process of invading *Drosophila* species. *RNA* 16, 1634–1645
93 Tam, O.H. *et al.* (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* 453, 534–538
94 Watanabe, T. *et al.* (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 453, 539–543
95 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
96 Batista, P.J. *et al.* (2008) PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility *in C. elegans*. *Mol. Cell* 31, 67–78
97 Wang, G. and Reinke, V. (2008) A C. elegans Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. *Curr. Biol.* 18, 861–867
98 Kaminker, J.S. *et al.* (2002) The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol.* 3, RESEARCH0084
99 Celniker, S.E. and Rubin, G.M. (2003) The *Drosophila melanogaster* genome. *Annu. Rev. Genomics Hum. Genet.* 4, 89–117
100 Smith, C.D. *et al.* (2007) The Release 5.1 annotation of *Drosophila melanogaster* heterochromatin. *Science* 316, 1586–1591
101 Hoskins, R.A. *et al.* (2007) Sequence finishing and mapping of *Drosophila melanogaster* heterochromatin. *Science* 316, 1625–1628
102 Clark, A.G. *et al.* (2007) Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450, 203–218
103 Reddien, P.W. *et al.* (2005) SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science* 310, 1327–1330
104 Kuramochi-Miyagawa, S. *et al.* (2004) *Mili*, a mammalian member of *piwi* family gene, is essential for spermatogenesis. *Development* 131, 839–849
105 Kuramochi-Miyagawa, S. *et al.* (2001) Two mouse *piwi*-related genes: *miwi* and *mili*. *Mech. Dev.* 108, 121–133
106 Deng, W. and Lin, H. (2002) *miwi*, a murine homolog of *piwi*, encodes a cytoplasmic protein essential for spermatogenesis. *Dev. Cell* 2, 819–830
107 Carmell, M.A. *et al.* (2007) Miwi2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev. Cell* 12, 503–514
108 Aravin, A.A. *et al.* (2008) A piRNA pathway primed by individual transposons is linked to *de novo* DNA methylation in mice. *Mol. Cell* 31, 785–799
109 Kuramochi-Miyagawa, S. *et al.* (2008) DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev.* 22, 908–917
110 Frydman, H.M. and Spradling, A.C. (2001) The receptor-like tyrosine phosphatase lar is required for epithelial planar polarity and for axis determination within *Drosophila* ovarian follicles. *Development* 128, 3209–3220