NIH Public Access
Author Manuscript
Mol Cell. Author manuscript; available in PMC 2009 August 21.

Published in final edited form as:
Mol Cell. 2008 September 26; 31(6): 785–799. doi:10.1016/j.molcel.2008.09.003.

A piRNA pathway primed by individual transposons is linked to *de novo* DNA methylation in mice

Alexei A. Aravin[1], **Ravi Sachidanandam**[1], **Deborah Bourc'his**[2], **Christopher Schaefer**[2,3], **Dubravka Pezic**[4], **Katalin Fejes Toth**[1], **Timothy Bestor**[3], and **Gregory J. Hannon**[1,*]

[1]Watson School of Biological Sciences, Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724

[2]Institute Jacques Monod, Inserm U741/Paris 7 University, Tour 43-44/2eme etage/couloir 43-44, 2 place Jussieu, 75251 Paris Cedex 05, France

[3]Department of Genetics and Development, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA

[4]Institute of Molecular Biotechnology of the Austrian Academy of Sciences, 1030 Vienna, Dr. Bohr-Gasse, Austria

Abstract

piRNAs and Piwi proteins have been implicated in transposon control and are linked to transposon methylation in mammals. Here, we examined the construction of the piRNA system in the restricted developmental window in which methylation patterns are set during mammalian embryogenesis. We find robust expression of two Piwi family proteins, MIWI2 and MILI. Their associated piRNA profiles reveal differences from *Drosophila* wherein large piRNA clusters act as master regulators of silencing. Instead, in mammals, dispersed transposon copies initiate the pathway, producing primary piRNAs, which predominantly join MILI in the cytoplasm. MIWI2, whose nuclear localization and association with piRNAs depend upon MILI, is enriched for secondary piRNAs antisense to the elements that it controls. The Piwi pathway lies upstream of known mediators of DNA methylation, since piRNAs are still produced in *Dnmt3L* mutants, which fail to methylate transposons. This implicates piRNAs as specificity determinants of DNA methylation in germ cells.

Introduction

It is of paramount importance that germ cell genomes be protected from the uncontrolled propagation of mobile genetic elements. This prevents both long-term reductions in fitness through accumulation of mutations and short-term reductions in fertility due to germ cell loss. For protection to be effective, mobile elements must be distinguished from endogenous genes and selectively silenced. This presents a significant challenge, as few characteristics unambiguously mark the many families of mobile elements that colonize animal genomes (Girard and Hannon, 2008).

© 2008 Elsevier Inc. All rights reserved.

*To whom correspondence should be addressed (hannon@cshl.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

In *Drosophila*, germ cells express a class of small RNAs, piRNAs that are specialized for mobile element repression (Brennecke et al., 2007; Saito et al., 2006; Vagin et al., 2006). These act in a nucleic acid based innate immune system that comprises both genetically encoded (primary piRNAs) and adaptive (secondary piRNAs) resistance mechanisms. Primary piRNAs, are generated from dedicated loci, called piRNA clusters, that contain the highest density of transposon-related sequences in the fly (Brennecke et al., 2007).

Primary piRNAs join Piwi proteins and guide these to the selection of their targets. Two outcomes follow recognition of a transposon mRNA by a piRNA. First, the mRNA is cleaved, resulting in its destruction. The cleavage event also promotes the production of a secondary piRNA derived from the mRNA itself (Brennecke et al., 2007; Gunawardane et al., 2007). Though this is in the sense orientation with respect to the transposon, it can target antisense transposon RNAs that are generated from piRNA clusters. Cleavage of such transcripts reproduces original antisense piRNAs that can again target transposons. As a whole, this creates a cycle, called the “ping-pong cycle” that optimizes the piRNA population to target active elements (Brennecke et al., 2007), reviewed in (Aravin et al., 2007a).

Features of the *Drosophila* piRNA system have been conserved in vertebrates. Signatures of ping-pong amplification cycles have been detected in zebrafish (Houwing et al., 2007) and mammals (Aravin et al., 2007b). Moreover, repeat-enriched mouse piRNA clusters give rise to small RNA species in both male and female germ cells (Aravin et al., 2007b; Tam et al., 2008). Finally, LTR and non-LTR transposons are overexpressed in male germ cells deficient in either of two Piwi family members, *Mili* or *Miwi2* (Aravin et al., 2007b; Carmell et al., 2007; Kuramochi-Miyagawa et al., 2008).

In contrast to *Drosophila*, epigenetically stable repression of transposable elements in mammals requires CpG DNA methylation (Bourc'his and Bestor, 2004; Walsh et al., 1998). Transposon methylation patterns are extensively remodeled during mammalian development. Following fertilization, gametic methylation patterns are mostly lost (Lane et al., 2003). *De novo* methylation patterns are then established around implantation and subsequently maintained in somatic cells throughout the life of the organism. The germ line undergoes a wave of transposon demethylation soon after its emergence during embryogenesis, as it colonizes the gonadal compartment (Hajkova et al., 2002). The subsequent establishment of gametic methylation patterns differs between the two sexes. In males, after migration of primordial germ cells into embryonic gonads and their initial expansion, germ cells arrest their cell cycle around 14.5dpc (days post-coitum) as prospermatogonia, only resuming division 2–3 days after birth (dpp). This is the critical window during which male gametic methylation patterns are established (Kato et al., 2007; Lees-Murdock et al., 2003).

Several members of the DNA methyltransferase family, DNMT3A, DNMT3B and DNMT3L, act in *de novo* methylation of transposable elements. Catalytically active DNMT3A and DNMT3B are important in both germ and somatic cells, where they perform complementary and non-overlapping functions. Constitutive *Dnmt3a* or *Dnmt3b* mutations are embryonically lethal (Okano et al., 1999), and germ cell conditional inactivation of *Dnmt3a* results in sterility (Kaneda et al., 2004). In contrast, DNMT3L operates as a central regulator of *de novo* methylation specifically in the germline. *Dnmt3L*-deficient animals globally fail to establish *de novo* methylation of transposons in their germ cells with no other phenotypic manifestation (Bourc'his and Bestor, 2004; Kato et al., 2007). In males, this results in uncontrolled transposon expression and eventually in spermatogenesis failure and sterility. Recent studies have begun to unravel the biochemistry of *de novo* DNA methylation machinery, showing that DNA methylation might be preceded by specific histone modifications (Jia et al., 2007; Ooi et al., 2007). However, it is still not clear how transposon sequences are specifically recognized to receive such modifications.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

In the male germline, deficiency in either of two Piwi family members, *Mili* or *Miwi2*, results in loss of the DNA methylation marks on transposons, and mutant animals display a phenotype remarkably similar to that of *Dnmt3L*-deficient mice (Aravin et al., 2007b; Carmell et al., 2007; Kuramochi-Miyagawa et al., 2004; Kuramochi-Miyagawa et al., 2008). These data led to hypothesis that Piwi/piRNA complexes might serve as a sequence-specific guides that direct the *de novo* DNA methylation machinery to transposable elements (Aravin et al., 2007b; Kuramochi-Miyagawa et al., 2008). Thus far, we understand little of how the piRNA system operates in the restricted developmental window that is critical to the stable epigenetic silencing of repeat elements. This prompted investigation of both Piwi protein and piRNA expression during the time during which methylation marks are established.

Results

Expression of Piwi family members during germline development

The mouse genome contains three Piwi family members: *Mili*, *Miwi* and *Miwi2*. *Miwi* is expressed from the pachytene stage of meiosis to the haploid round spermatid stage (Deng and Lin, 2002) (Fig. 1A). MILI is also present during meiosis, at which point both MILI and MIWI interact with an extremely abundant class of small RNAs, the pachytene piRNAs (Aravin et al., 2006; Girard et al., 2006). These are derived from specific genomic loci and form a complex population of small RNAs that match only to those sites from which they are derived. The function of this sub-class of piRNAs is elusive.

Mili is also expressed earlier in development. Just before entry into meiosis and the onset of pachytene piRNA expression, MILI binds piRNAs, which are different in character and genomic origin from meiotically expressed piRNAs (Aravin et al., 2007b). These are derived from a set of clustered loci that are repeat-enriched and thus give rise to small RNA populations corresponding to transposons.

Loss of *Miwi2* gives both morphological (e.g., meiotic arrest and progressive germ cell loss) and molecular phenotypes (increased transposon expression) that resemble those resulting from *Mili* deficiency (Aravin et al., 2007b; Carmell et al., 2007; Kuramochi-Miyagawa et al., 2004; Kuramochi-Miyagawa et al., 2008). However, neither the developmental timing of *Miwi2* expression nor its subcellular localization has been reported.

We investigated *Mili* and *Miwi2* expression both by using specific antibodies raised against each family member and using transgenic animals that express GFP- and myc-tagged MILI and MIWI2 proteins under the control of their endogenous promoters. For each protein, examination of several GFP and myc-transgenic lines and comparison to results obtained with antibodies against native proteins revealed essentially identical patterns.

We analyzed the expression of *Mili* and *Miwi2* during germ line development in male and female mice. *Mili* expression could be detected in both sexes as early as 12.5 dpc, a time when migrating primordial germ cells (PGCs) have reached the somatic genital ridge (Fig. 1A, not shown). MILI localized to numerous perinuclear cytoplasmic granules in male PGCs (Fig. 1B). These structures were remarkably similar to nuage, which contain the Piwi proteins, AUB and AGO3, in *Drosophila* female germ cells. MILI was also detected in female PGCs and localized to cytoplasmic granules (Fig. 1B). *Mili* expression continued in both male and female germ cells after birth (Fig. S1A). In ovary it localized to cytoplasmic granules of both arrested and growing oocytes. In adult testis, it was expressed throughout spermatogenesis until the round spermatid stage at which point the majority of MILI localized in a single prominent granule, the chromatoid body, as evident from co-localization with chromatoid body marker MVH (DDX4) (Fig. S1B).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

MIWI2 could be detected in male germ cells beginning around 14.5–15.5 dpc but was absent from female germ cells (Fig. 1B). MIWI2 was present in the nucleus as well as the cytoplasm. In the cytoplasm, MIWI2 occupied granules similar to but fewer in number than those containing MILI. From 15.5 dpc and until birth both proteins were present in male germ cells (Fig 1B). We did not detect MILI or MIWI2 in somatic cells of the embryonic gonad. MIWI2 expression declined soon after birth reaching undetectable levels in 4-day old mice. Remarkably, the very short window of MIWI2 expression during male germ cell development (15.5 dpc-3 dpp) corresponds to the time of cell cycle arrest and *de novo* DNA methylation.

MILI and MIWI2-bound piRNA populations during germline development

MILI and MIWI2 complexes were recovered from male embryonic gonads isolated from 16.5 dpc embryos. Each protein associated with piRNAs of a specific size. MILI bound ~26 nt. RNAs (Aravin et al., 2006; Aravin et al., 2007b), and MIWI2 associated with ~28 nt RNAs (Fig 1C).

We prepared libraries of piRNAs from MILI and MIWI2 complexes at 16.5 dpc and 24–33 nt total RNAs from the same stage. Additionally, we cloned libraries from 24–33 nt RNAs and MILI complexes at days 2 and 10 after birth. The small RNA libraries were sequenced, typically yielding 2–3 million small RNA reads per library. Between 40 and 70% of sequences matched perfectly to the mouse genome, and these were considered for further analysis (Suppl. Table 1). MILI and MIWI2-associated sequences showed a normal distribution with peaks at 26 and 28 nt respectively (Fig. 1D). The profile of total cellular small RNA suggested that MILI-bound piRNAs are slightly more abundant than MIWI2 piRNAs.

At 16.5 dpc 47.5% of all cellular small RNAs were derived from transposon sequences (Fig. 1E and Table S1). The other small RNAs represent fragments of abundant, larger non-coding RNAs (29.6%), sequences derived from un-annotated genomic regions (20.3%) or exons of protein encoding genes (2.5%). The overall fraction of transposon-derived piRNAs remained relatively stable during development from 16.5 dpc to 10 dpp. However, different types of transposons showed distinct patterns. The fraction of LTR and LINE-derived piRNAs decreased while SINE-derived sequences increased during development. The fraction of exon-derived small RNAs also increased substantially (from 2.5% at 16.5 dpc to 18.6% at 10 dpp). In *Mili* knock-out animals, LTR and LINE-derived small RNAs were almost completely eliminated indicating that they represent *bona fide* piRNAs. SINE and exon-derived small RNA were decreased in abundance but not eliminated in knock-outs suggesting that small portion of this class derives from degradation product of SINE-containing and genic transcripts.

Both MILI and MIWI2 associated with repeat-derived piRNAs during prenatal development (Fig. 1E). However, MIWI2 demonstrated greater specificity for transposons. 76% of MIWI2-bound piRNAs mapped to LTR and LINE retrotransposons as compared to 45.7% for MILI-bound species. MILI complexes also contained piRNAs derived from exons of protein-coding genes (4.8%). As observed for total RNA profiles, the fraction of MILI-bound LTR and LINE piRNAs decreased and SINE and exon-derived piRNAs increased from 16.5 dpc to 10 dpp. A large fraction of total cellular small RNAs (48.7%) and MILI-bound piRNAs (60.1%) could be mapped to a unique genomic position (Fig. S2). Both total small RNAs and MILI-associated piRNAs contained a subset (14–17%) that matched highly repetitive sequences (100 or more genomic mappings). These dramatically decreased in *Mili* mutants. In MIWI2 complexes the fraction of uniquely mapped piRNAs was lower (35.6%) and the highly repetitive fraction was substantially higher (30.5%).

Overall, our analysis revealed that both MILI and MIW2 bind transposon-derived piRNAs during embryonic development and that the fraction of piRNAs that match active transposons

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

classes decreases after birth. Interestingly, the partners of the two Piwi proteins are different with MIWI2 complexes being particularly enriched in transposon-derived piRNAs.

Transposon-derived piRNAs

A substantial fraction of piRNAs matched the three major classes of transposable elements present in mammalian genomes: LTR, LINE and SINE retrotransposons. We analyzed piRNAs derived from the representative element of each class that produced the largest number of small RNAs: the IAP LTR-retrotransposon, LINE1 and SINE B1. During development, the abundance of LINE1 piRNAs decreased and SINE increased, while IAP piRNAs showed a double-peaked with higher levels before birth and at 10 dpp (Fig. 2A). LINE- and IAP-derived piRNAs were associated with both MILI and MIWI2 (Fig. 2B). Interestingly, for both elements piRNAs were almost equally distributed among MIWI2 and MILI complexes (Fig. 2B). Sequences derived from exons were almost exclusively bound to MILI (Fig. 2B).

To probe the functions of Piwi-bound small RNAs, we analyzed strand orientation of transposon- and gene-derived piRNAs. Exon-derived piRNAs were highly enriched for sense sequences. All three transposons produced substantial numbers of antisense piRNAs, with different elements showing different characteristics (Fig. 2C). For LINE1, piRNAs were generally enriched in antisense sequences throughout development. In prenatal testis MILI and MIWI2 had opposite strand preferences with MILI binding more sense and MIWI2 binding more antisense piRNAs. Interestingly, strand orientation of LINE1 piRNA in MILI complexes was reversed after birth. IAP and SINE piRNAs were nearly equally divided between sense and antisense sequences in both complexes and in total piRNA populations in prenatal cells. However at 10 dpp, sense IAP and SINE piRNAs became predominant. For all three transposons the fraction of antisense sequences dramatically decreased in MILI-deficient animals.

These results indicate that LINE1 piRNAs are sorted into MILI and MIWI2 complexes according to strand orientation, similar to what is observed for Piwi family members in *Drosophila* (Brennecke et al., 2007). However, this bias was slight for SINE B1 and was absent for IAP. These data also reveal that strand bias even within a given Piwi complex can be dynamic during development.

We next analyzed the distribution of piRNAs along transposon consensus sequences (Fig. 2D). Notably, MILI-bound LINE1 piRNAs were enriched in antisense and biased toward the 5’ end of LINE1 in prenatal germ cells. At 10 dpp, piRNAs were less strand biased and mapped more frequently toward the 3’ ends of the elements, likely reflecting the higher abundance of those sequences in the genome (Kazazian, 2004). The enrichment for piRNAs mapping to LINE1 5’ ends at 16.5 dpc implies that at this developmental stage piRNAs are processed from full-length, potentially active copies.

The ping-pong cycle in prenatal piRNAs

Two different mechanisms, primary processing and ping-pong amplification, have been proposed to generate piRNAs (Brennecke et al., 2007), reviewed in (Aravin et al., 2007a). Primary processing samples single-stranded piRNA precursor transcripts generating a diverse set of piRNA sequences that share a preference for 5’ uridine (1U). Pachytene piRNAs are exclusively produced by the primary processing mechanism. A subset of pre-pachytene piRNAs in mouse (Aravin et al., 2007b) and a substantial fraction of *Drosophila* piRNAs (Brennecke et al., 2007) are generated by a mechanism that depends upon the endonuclease activity of Piwi proteins and that is referred to as the ping-pong amplification cycle.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

The ping-pong cycle requires the presence of transcripts that are complementary to primary piRNAs. Recognition by primary piRNAs guides the endonuclease activity of Piwi proteins, which cleave the transcript 10 nucleotides from the 5' end of the original piRNA (Brennecke et al., 2007; Gunawardane et al., 2007). This event generates 5' end of a new secondary piRNA. These show a strong bias for adenine at position 10 (10A) complementing the 1U bias of primary piRNAs (Fig. 3A, B). Secondary piRNAs can also generate new piRNAs by recognizing and cleaving complementary transcripts to regenerate the piRNA that initiated the cycle. Thus, only secondary piRNAs are enriched for 10A. Although piRNAs with the sequence identical to original primary piRNA can be created during the cycle (following cleavage by a 10A secondary piRNA), these 1U-biased species remain reflective of a primary species that initiated the cycle. Thus, we consider the 1U population to be reflective of primary piRNA biogenesis.

We investigated the existence of the ping-pong cycle in prenatal germ cells and the roles of MILI and MIWI2 in this process. Tracking a specific feature of ping-pong piRNA pairs, the 10nt offset between 5' ends of piRNAs, showed that both MILI and MIWI2 participate in the amplification cycle. The most prevalent signatures indicated MIWI2-MIWI2 and MIWI2-MILI cycles (Fig. 3C).

In *Drosophila*, Piwi proteins participating prominently in the ping-pong cycle show piRNA strand specificity (Brennecke et al., 2007). We tested whether similar characteristics defined the ping-pong cycle in prenatal testis. Ping-pong pairs where the sense strand associated with MILI and antisense with MIWI2 were more abundant than pairs with the opposite character (Fig. 3D). Surprisingly, this held true not only for LINE1 piRNAs that are generally asymmetrically distributed in MILI and MIWI2 complexes, but also for IAP piRNAs, that do not show a protein-dependent strand bias overall.

As was seen for strand orientation, MILI and MIWI2 complexes also discriminated primary and secondary piRNAs. We calculated the preference (Primary/Secondary, P/S, ratio) by taking the number of piRNAs that show 1U but no 10A bias (primary-like) and dividing it by the number that show a 10A but no 1U bias (secondary piRNAs). It should be noted that this approach ignores all sequences with both 1U and 10A, as these cannot be assigned to primary or secondary categories. Prenatal piRNAs are strongly enriched in secondary sequences as compared to pachytene piRNAs, which appear to be generated exclusively by primary processing (P/S ratios of 5.13 and 21.2 respectively). LINE1 and IAP piRNAs showed strong signals for secondary sequences as compared to exon-derived piRNAs, which given a lack of antisense information must be generated by primary processing (Fig. 3E). MIWI2 complexes were ~2 fold enriched in secondary piRNAs as compared to MILI. Overall our data indicate distinct roles for MILI and MIWI2 in the piRNA-processing pathway. MILI is biased toward primary piRNAs and 1U-containing piRNAs generated in the ping-pong cycle. MIWI2 is particularly enriched in secondary sequences.

Finally, we probed the correlation between the strand orientation of transposon piRNAs and their processing category (primary or secondary). In both MILI and MIWI2 complexes, antisense piRNAs were enriched for secondary sequences. MILI-associated sense piRNAs had a P/S ratio of 4.46 versus 2.02 for antisense, and MIWI2-associated sense and antisense piRNAs had P/S ratios of 2.54 and 1.40, respectively (Fig. 3F).

Overall, these data are consistent with a model in which sense transcripts, most likely mRNAs of active transposons, represent the major substrate for primary processing and result in piRNAs associated with MILI. MILI-associated primary sense piRNAs recognize and cleave transcripts that contain transposon sequences in the antisense orientation and generate secondary piRNAs that join MIWI2 complexes. This is precisely opposite to the bias observed

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

in *Drosophila*, where primary piRNAs are mostly derived from piRNA clusters and are enriched for antisense strands while secondary piRNAs are sense.

Ping-pong relationships between MILI and MIWI2 suggest a need for physical proximity. Co-localization of MILI and MIWI2 was investigated by immunofluorescence in 17.5 dpc germ cells. Though MIWI2 was mainly present in the nucleus, MIWI2-containing cytoplasmic granules co-localized with or were in close proximity to MILI-containing granules (Fig. 4A). MILI granules were more abundant as compared to MIWI2 granules, and many MILI granules did not interact with MIWI2 foci. In *Miwi2*-deficient cells, MILI localization in cytoplasmic granules did not change (Fig. 4B, lower panel). In contrast, in *Mili* mutants MIWI2 re-localized from the nucleus to the cytoplasm where it was uniformly distributed rather than concentrating in granules (Fig. 4B, upper panel). Notably, we detected no MIWi2-associated piRNAs in Mili mutants (Fig. 4C) indicating that MIWI2 remains unloaded when MILI is absent. This epistatic relationship between MILI and MIWI2 supports the proposed directionality of their interaction wherein MILI initiates the cycle with primary piRNAs, and the production of secondary piRNAs associated with MIWI2 depends upon the prior existence of these species.

Genomic origins of prenatal piRNAs

Derivation from clustered loci in the genome has been a defining feature of piRNAs, and these loci play an important role in piRNA generation in both *Drosophila* and vertebrates (Aravin et al., 2006; Aravin et al., 2007b; Brennecke et al., 2007; Girard et al., 2006; Lau et al., 2006). To investigate the genomic origin of prenatal piRNAs we searched for unambiguously mapping sequences in close proximity in the genome. Using a threshold value of 10 piRNAs per kilobase, we identified 3399 clusters, which were ranked by their relative contributions to piRNA populations. Though the most prominent cluster gave rise to almost 10% of all uniquely mapped piRNAs, the individual contribution of each subsequent cluster dropped dramatically (Fig. 5A). Combining all clusters yielded about 8 Mb of genomic space, which could accommodate only ~50% of all uniquely mapping piRNAs. This suggests fundamental differences from the pachytene piRNAs (Aravin et al., 2006; Girard et al., 2006) and even those pre-pachytene piRNAs that occupy *Mili* at 10 dpp (Aravin et al., 2007b). In fact, most prenatal clusters actually represent individual transposons and transposon fragments. Expression from all clusters was eliminated in *Mili* mutants (Fig. 5B).

Analyzing the expression of clusters that gave rise to the greatest numbers of piRNAs indicated developmentally regulated expression patterns. Only cluster #2 on chromosome 10 persisted in its expression through day 10 dpp. Interestingly this cluster was also highly expressed in ovary (Fig. 5B).

Unlike piRNA clusters identified in mammals thus far, 6 out of the 8 most prominent prenatal clusters generate piRNAs from both genomic strands (double-strand clusters). This arrangement is very similar to piRNA clusters in *Drosophila* ovary (Brennecke et al., 2007). The two most prominent clusters (#1 chr.7 and #2 chr.10) show profound strand asymmetry with the vast majority of piRNAs being derived from one genomic strand (single-strand clusters). As with *Drosophila flamenco*, these clusters also showed enrichment for similarly oriented transposon fragments and for the generation of piRNAs that are antisense to transposon mRNAs (Fig. 5C). Surprisingly, double-stranded clusters are not particularly enriched in transposon sequences. Both single-strand and double-strand clusters produced piRNAs in MILI and MIWI2 complexes (Fig. 5C and 5E). However, even in double-strand clusters MILI-bound piRNAs were strongly biased for one of the two genomic strands (Fig. 5E). Size profiles of total cellular piRNAs indicated that single-strand clusters produced significantly more MILI-bound piRNAs (Fig. 5D), while double-strand clusters contributed to MILI and MIWI2 complexes equally (Fig. 5F).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

piRNAs derived from clusters participate in the ping-pong cycle (data not shown). For cluster-derived piRNAs, MIWI2 remained biased towards secondary piRNAs as compared to MILI (Fig. 5G). Single-strand clusters produced more primary piRNAs in both MILI and MIWI2 complexes than did double-strand clusters. Indeed, the P/S ratio of MILI-bound piRNAs derived from the single-strand cluster on chr. 7 (16.21) was close to that of pachytene piRNAs (21.2).

If ping-pong amplification operates by interaction of piRNAs from single-strand clusters with transposon mRNAs in trans, MIWI2, which is enriched in secondary piRNAs, should contain more transposon-derived piRNAs than MILI. Indeed, for the two single-strand transposon-rich clusters (#1, chr. 7 and #2, chr.10), MIWI2 was particularly enriched in transposon-derived piRNAs (Fig. 5H, left panel). The difference between MILI and MIWI2 complexes was even more dramatic when strand orientation was taken into account. For both clusters, MIWI2 was enriched in piRNAs that match transposons in the antisense orientation (Fig. 5H, right panel). These data indicate that the ping-pong cycle strongly shapes piRNA populations and enhances the production of piRNAs that match transposons in the antisense orientation.

DNA methylation and Piwi/piRNA pathway

DNA methylation is critical to stable, epigenetically inherited silencing of transposons, and this is lost upon mutation of either MILI or MIWI2 (Aravin et al., 2007b; Carmell et al., 2007; Kuramochi-Miyagawa et al., 2008). DNMT3L, a catalytically defective member of the DNA methyltransferase family, is essential for both proper transposon methylation and transposon repression (Bourc'his and Bestor, 2004). DNMT3L acts together with the catalytically active *de novo* methyltransferases, DNMT3A and DNMT3B, to establish methylation patterns (Chen et al., 2005; Gowher et al., 2005; Suetake et al., 2004). We sought to order MILI and MIWI2 with respect to other pathway components. Since DNMT3L acts as a coordinator of methylation activities in the male germ line, we examined the integrity of the piRNA pathway in *Dnmt3L*-deficient animals.

We immunoprecipitated MILI from testes of 10 dpp *Dnmt3L*-deficient animals and their wild-type littermates and found that piRNAs were still present in mutant animals (Fig. 6A). Cloning and analysis of small RNA libraries showed that the fraction of LTR and LINE retrotransposon piRNAs increased, while the fraction of SINE piRNAs decreased in *Dnmt3L* mutants (Fig. 6B, left panel). This was consistent with Northern blotting for an abundant IAP-derived piRNA, which also increased in abundance in the mutant (Fig. 6C).

Several lines of evidence indicated that increases in LTR and LINE piRNAs might be linked to de-repression of these elements and the increased capacity of transposon mRNAs for entry into the piRNA pathway. First, LINE and LTR elements lost methylation and silencing, showed increased expression and contributed a greater number of sense piRNAs in *Dnmt3L* mutants. In contrast SINEs (B1) were neither affected at the level of DNA methylation and expression nor contributed increasingly to piRNA populations in mutant animals (unpublished data). Moreover, LINE and LTR piRNAs that increased in *Dnmt3L* mutants corresponded to elements that are close to consensus and thus potentially expressed under circumstances where methylation was lost (Fig. 6B, right panel).

As an example, piRNAs derived from IAP elements showed the greatest change in *Dnmt3L* mutants, and this was mainly due to a dramatic increase in sense small RNAs (Fig. 6C, D). This pattern was also obvious when the distribution of piRNAs along the IAP consensus was displayed (Fig. 6E). Normally, the ratio of primary to secondary sequences in MILI complexes decreases after birth and approaches that of MIWI2 in prenatal germ cells (Fig. 6F, compare to Fig. 3E). In *Dnmt3L*-deficient animals the ratio of primary to secondary IAP piRNA increased ~ 6-fold as compared to wild-type. This strongly supports the model that mRNAs

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

from IAP elements, which show increased expression in *Dnmt3L* mutants, flow into the piRNA pathway as a source of primary piRNAs.

Considered, these data support a model in which the piRNA pathway acts upstream of DNMT3L, and consequently DNMT3A and 3B, to help establish patterns of DNA methylation on repeat elements. This interpretation rests on the observation that while mutations in methyltransferase family members impact cytosine methylation, they leave the piRNA pathway largely intact, affecting only the composition of small RNA populations in manner that can be rationalized by the impact of loss of methylation on transposon expression. In contrast, loss of the piRNA pathway prevents the recognition and silencing of potentially active transposons by the DNMT3L pathway.

Discussion

Dynamic expression of Piwis and piRNAs

The mouse genome encodes three Piwi proteins. MIWI is expressed beginning at the pachytene stage of meiosis (Deng and Lin, 2002) and together with MILI binds to non-repetitive pachytene piRNAs whose function remains unknown (Aravin et al., 2006; Girard et al., 2006). MIILI is expressed very early in germ cell development and persists through the completion of meiosis. Deficiency in either MILI or MIWI2 causes increased transposon expression and similar defects in meiosis and germ cell survival (Aravin et al., 2007b; Carmell et al., 2007; Kuramochi-Miyagawa et al., 2008), suggesting their cooperation in the same pathway. Results reported herein support an intimate connection between MILI and MIWI2 in transposon control.

MIWI2 protein had not previously been detected, leaving its developmental expression profile and subcellular localization unknown. We found that both MILI and MIWI2 are expressed in prospermatogonia during embryogenesis. MILI can be detected in both male and female germ cells starting at 12.5 dpc, soon after PGCs arrive at the embryonic gonad and continues to be expressed after birth throughout both oogenesis and spermatogenesis. In contrast, *Miwi2* expression can be detected exclusively in male germ cells in a window from 15 dpc to soon after birth. The timing of *Miwi2* expression precisely corresponds to the stage of cell cycle arrest when prospermatogonia establish *de novo* methylation patterns (Lees-Murdock et al., 2003; Trasler, 2006; Walsh et al., 1998). Furthermore, it coincides with the time when expression of LINE1 elements (Trelogan and Martin, 1995) and an IAP transgene (Dupressoir and Heidmann, 1996) can be detected in wild-type testes. Overall, our data show that two Piwi members are co-expressed in primordial male germ cells at the moment that is critical for establishing epigenetically stable silencing of mobile elements.

Developmental profiling revealed unexpectedly dynamic piRNA populations. Previously we reported that the expression of abundant, repeat-poor piRNAs, pachytene piRNAs, during meiosis is preceded by the appearance of another transposon enriched piRNA population, pre-pachytene piRNAs (Aravin et al., 2007b). Here we showed that the composition of the pre-pachytene piRNA pool changes throughout development, even though at least one binding partner, MILI, is expressed throughout the entire window examined. Changes are observed in the types of transposons targeted, the strand bias of populations, the distribution of piRNAs along transposon consensus sequences and the degree to which primary and secondary piRNAs populate the pathway.

It is likely that the dynamics of piRNA populations reflect a complex interplay between the expression of Piwi proteins, transposable elements, and piRNA clusters. Particularly, we note a switch from the ‘MILI/MIWI2’ ping-pong cycle that operates before birth to one containing MILI alone after birth. Since amplification of piRNAs in the ping-pong loop depends on the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

presence of transcripts from active elements, piRNA populations should be greatly influenced by the expression level of transposons. Accordingly, we found the highest fraction of LINE and IAP-derived piRNAs at 16.5 dpc, the time at which both elements are actively expressed (Dupressoir and Heidmann, 1996; Trelogan and Martin, 1995). Finally, the expression of piRNA clusters changes during development. The majority of clusters expressed in pre-natal testes no longer produce piRNAs post-natally, at which point a different set of clusters predominates. Overall, our data highlight the highly dynamic nature of piRNA populations that are shaped to provide an adequate response to active transposable elements.

The piRNA ping-pong cycle

Analysis of piRNA sequences in both flies and mouse indicates the existence of two classes of piRNAs, primary and secondary (reviewed in (Aravin et al., 2007a). Most primary piRNAs in *Drosophila* are antisense to transposons and are derived from few large piRNA clusters that contain a diversity of transposon fragments (Brennecke et al., 2007). In mice, during prenatal development, the situation appears to be quite different. We find that primary piRNAs in mouse are predominantly sense oriented with respect to transposon transcripts and appear to derive mainly from transposon mRNAs instead of from piRNA clusters. The processing of primary piRNAs from transposon mRNAs is supported by several observations. Primary LINE1 piRNAs are biased toward the 5' end of this element suggesting that they are generated from transcripts of active full-length LINE1 elements; the majority of LINE1 fragments in the genome are inactive, 5' truncated copies (Kazazian, 2004). Furthermore, increases in sense, primary IAP piRNAs are observed in *Dnmt3L* mutants that de-repress these elements. These observations indicate that transposon mRNAs can be efficiently used as substrates for primary piRNA processing. Though long piRNA clusters do exist in mouse prenatal germ cells, dispersed transposon copies outweigh their contribution to the system.

Currently, it is unclear how transposon transcripts are distinguished from cellular mRNAs as an input into the piRNA system. An alternative hypothesis is that they are indeed not distinguished, but that the transcriptome is instead essentially sampled at random, with only those sequences that encounter complimentary antisense transcripts truly engaging the pathway and increasing their relative abundance through a ping-pong cycle. This hypothesis is supported to some degree by the prevalence of piRNAs from exons of normal protein coding genes. These genic piRNAs are almost exclusively primary and might be considered as by-products of a random sampling of transcriptome for piRNA generation. In contrast, piRNAs from transposon-poor double-stranded piRNA clusters might represent by-products of the ping-pong amplification acting on loci in the genome that happen to be transcribed bi-directionally. In this case, complementary transcripts are readily available in cis allowing robust operation of the ping-pong cycle in the absence of transcripts in trans.

Overall, our data indicate that in mouse the initiating input for the ping-pong cycle is different than in *Drosophila*. In mouse primary processing of mRNAs of active elements generates sense piRNAs that start the cycle, while in *Drosophila* primary piRNAs are often antisense to transposons and are processed from few specialized piRNA clusters (Fig. 7). Recognition of transposon mRNAs as a source for primary processing and an independence from piRNA clusters might allow greater flexibility in response to expansion of transposable elements in mammals. However it also poses the problem of how transposons are distinguished from host genes.

Specialized functions of MILI and MIWI2 proteins

Both the content of their associated piRNA populations and their intracellular location suggest that MILI and MIWI2 play quite specialized roles in the piRNA pathway and in transposon silencing. Several lines of evidence suggest that MILI might be the principal recipient of

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

piRNAs generated by the primary processing mechanism. MILI complexes are enriched for 1U-containing species, and exonic piRNAs for which no ping-pong partners are found populate MILI complexes almost exclusively. Single-strand clusters, which appear to be processed mainly by the primary mechanism, contribute disproportionately to MILI. Finally, MILI seems to act upstream of MIWI2, consistent with its being a recipient of primary piRNAs.

Subcellular compartmentalization of these proteins also supports functional specialization. MIWI2 is found predominantly in the nucleus, while MILI is cytoplasmic. The impacts of mutations in this cytoplasmic family member on DNA methylation can be explained by downstream effects on MIWI2. Indeed we found that MIWI2 lacks piRNA partners and is lost from the nucleus in MILI mutants. A subset of MIWI2 is cytoplasmic and is found in or near MILI-containing granules. This co-localization is consistent with MILI and MIWI2 acting as partners in the ping-pong cycle.

Interactions between Piwi/piRNA and chromatin/DNA methylation pathways

During the development of primordial germ cells into prospermatogonia, methylation on mobile elements is erased and re-established (Trasler, 2006). We and others previously showed that mutation of either MIWI2 or MILI results in substantial de-methylation and de-repression of IAP and LINE1 elements (Aravin et al., 2007b; Carmell et al., 2007; Kuramochi-Miyagawa et al., 2008). Results presented here implicate these proteins in establishment rather than in maintenance of methylation patterns and support the idea that piRNAs serve as guides that direct the DNA methylation machinery to transposon sequences. This conclusion is supported not only by the nuclear localization of MIWI2 at the critical time when methylation patterns are established but also by indications from small RNA sequencing that both MILI and MIWI2 act upstream of DNMT3L, which in turn acts upstream of *de novo* methyltransferases, DNMT3A and 3B.

Overall, our data can be synthesized into a model (Fig. 7) in which loading of MIWI2 complexes with piRNAs depends on MILI function. This occurs through a ping-pong cycle that functions in cytoplasmic granules reminiscent of nuage and results in MIWI2 becoming enriched in antisense piRNAs targeting active elements. MIWI2 likely shuttles between the cytoplasm, where it acts in the ping-pong cycle, and the nucleus, where the majority of protein appears. Here it may recognize chromatin-bound nascent primary transcripts from active elements. Using co-immunoprecipitation from embryonic testes we did not detect interaction between MIWI2 and *de novo* methyltransferases, Dnmt3a and Dnmt3b, (not shown) therefore it is unlikely that piRNA-containing complexes directly bring DNA methyltransferases to their targets. With reference to pathways that are becoming increasingly well understood, particularly in *S. pombe* (Verdel and Moazed, 2005), piRNA complexes, bound to nascent transcripts, may simultaneously cleave those RNAs and recruit chromatin modifying enzymes to establish epigenetically stable repression marks that eventually lead to *de novo* DNA methylation. However, the precise pathway through which small RNAs act to induce the deposition of methylation marks remains a mystery at present.

Materials and methods

Animals and transgenic constructs

CD-1 wild-type mice were purchased from Charles River Lab. The *Miwi2* and *Dnmt3L* knock-out strains are described in (Carmell et al., 2007) and (Bourc'his and Bestor, 2004), respectively. The *Mili* knock-out strain was obtained from Haifan Lin (Yale University) and is described in (Kuramochi-Miyagawa et al., 2004). Transgenic animals expressing tagged MILI and MIWI2 were produced by BAC recombineering (see Supplementary methods).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Antibodies

Rabbit polyclonal antibodies against MILI were described in (Aravin et al., 2007b). Synthetic peptides (N-MSGRARVRARGIC-C: and N-GITTGHSAREVGRSSRDC-C) were used to raise rabbit polyclonal antibodies against MIWI2.

MILI and MIWI2 detection

To detect GFP-fused MILI and MIWI2, dissected testes were fixed for 10–15 min in 4% paraformaldehyde in PBS, washed in PBS, and mounted in 30% glycerol in PBS. For immunofluoresce, testis were fixed in either Bouin's solution or 4% paraformaldehyde (for details, see Supplementary Methods).

Simultaneous detection of MILI and MIWI2 in germ cells was performed using slides prepared from 3xMyc-MIWI2 transgenic animals using anti-myc (clone 4A6, Upstate) and anti-MILI antibodies.

Small RNA libraries

MILI and MIWI2 ribonucleoprotein complex immunopurification, RNA isolation, library construction and annotation were performed as described previously (Aravin et al., 2007b). For ovarian libraries 24–33 nt total RNA was isolated from ovaries dissected from CD1 females. Sequences reported in this manuscript are available in the Gene Expression Omnibus under accession number GSE12757.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Sang Yong Kim (CSHL) for generating transgenic animals. We thank Maria Mosquera, Lisa Bianco, Jodi Coblentz, and Gula Nourjanova (CSHL) for animal assistance and histology. We thank Stephen Hearn (Cold Spring Harbor Laboratory) for microscopy assistance and Emily Hodges, Michelle Rooks, Dick Mccombie, Danea Rabbolini and Laura Cardone for help with Illumina sequencing. We thank Catherine Schlingheyde for help with experiments and members of the Hannon laboratory, especially Antoine Molaro, for comments on the manuscript. G.J.H. is an investigator of the Howard Hughes Medical Institute. This work was supported by grants from the National Institutes of Health to G.J.H. and an NIH Pathway to Independence Award K99HD057233 to A.A.A.

Literature Cited

Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T, et al. A novel class of small RNAs bind to MILI protein in mouse testes. Nature 2006;442:203–207. [PubMed: 16751777]

Aravin AA, Hannon GJ, Brennecke J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science 2007a;318:761–764. [PubMed: 17975059]

Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K, Hannon GJ. Developmentally regulated piRNA clusters implicate MILI in transposon control. Science 2007b;316:744–747. [PubMed: 17446352]

Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature 2004;431:96–99. [PubMed: 15318244]

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 2007;128:1089–1103. [PubMed: 17346786]

Carmell MA, Girard A, van de Kant HJ, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell 2007;12:503–514. [PubMed: 17395546]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Chen ZX, Mann JR, Hsieh CL, Riggs AD, Chedin F. Physical and functional interactions between the human DNMT3L protein and members of the de novo methyltransferase family. J Cell Biochem 2005;95:902–917. [PubMed: 15861382]

Deng W, Lin H. miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev Cell 2002;2:819–830. [PubMed: 12062093]

Dupressoir A, Heidmann T. Germ line-specific expression of intracisternal A-particle retrotransposons in transgenic mice. Mol Cell Biol 1996;16:4495–4503. [PubMed: 8754850]

Girard A, Hannon GJ. Conserved themes in small-RNA-mediated transposon control. Trends Cell Biol 2008;18:136–148. [PubMed: 18282709]

Girard A, Sachidanandam R, Hannon GJ, Carmell MA. A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature 2006;442:199–202. [PubMed: 16751776]

Gowher H, Liebert K, Hermann A, Xu G, Jeltsch A. Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. J Biol Chem 2005;280:13341–13348. [PubMed: 15671018]

Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. Science 2007;315:1587–1590. [PubMed: 17322028]

Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J, Surani MA. Epigenetic reprogramming in mouse primordial germ cells. Mech Dev 2002;117:15–23. [PubMed: 12204247]

Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB, et al. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell 2007;129:69–82. [PubMed: 17418787]

Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. Nature 2007;449:248–251. [PubMed: 17713477]

Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, Sasaki H. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature 2004;429:900–903. [PubMed: 15215868]

Kato Y, Kaneda M, Hata K, Kumaki K, Hisano M, Kohara Y, Okano M, Li E, Nozaki M, Sasaki H. Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. Hum Mol Genet 2007;16:2272–2280. [PubMed: 17616512]

Kazazian HH Jr. Mobile elements: drivers of genome evolution. Science 2004;303:1626–1632. [PubMed: 15016989]

Kuramochi-Miyagawa S, Kimura T, Ijiri TW, Isobe T, Asada N, Fujita Y, Ikawa M, Iwai N, Okabe M, Deng W, et al. Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development 2004;131:839–849. [PubMed: 14736746]

Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Totoki Y, Toyoda A, Ikawa M, Asada N, Kojima K, Yamaguchi Y, Ijiri TW, et al. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev 2008;22:908–917. [PubMed: 18381894]

Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter J, Reik W. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis 2003;35:88–93. [PubMed: 12533790]

Lau NC, Seto AG, Kim J, Kuramochi-Miyagawa S, Nakano T, Bartel DP, Kingston RE. Characterization of the piRNA complex from rat testes. Science 2006;313:363–367. [PubMed: 16778019]

Lees-Murdock DJ, De Felici M, Walsh CP. Methylation dynamics of repetitive DNA elements in the mouse germ cell lineage. Genomics 2003;82:230–237. [PubMed: 12837272]

Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 1999;99:247–257. [PubMed: 10555141]

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 2007;448:714–717. [PubMed: 17687327]

Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev 2006;20:2214–2222. [PubMed: 16882972]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Suetake I, Shinozaki F, Miyagawa J, Takeshima H, Tajima S. DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. J Biol Chem 2004;279:27816–27823. [PubMed: 15105426]

Tam OH, Aravin AA, Stein P, Girard A, Murchison EP, Cheloufi S, Hodges E, Anger M, Sachidanandam R, Schultz RM, Hannon GJ. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature 2008;453:534–538. [PubMed: 18404147]

Trasler JM. Gamete imprinting: setting epigenetic patterns for the next generation. Reprod Fertil Dev 2006;18:63–69. [PubMed: 16478603]

Trelogan SA, Martin SL. Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. Proc Natl Acad Sci U S A 1995;92:1520–1524. [PubMed: 7878012]

Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. Science 2006;313:320–324. [PubMed: 16809489]

Verdel A, Moazed D. RNAi-directed assembly of heterochromatin in fission yeast. FEBS Lett 2005;579:5872–5878. [PubMed: 16223485]

Walsh CP, Chaillet JR, Bestor TH. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet 1998;20:116–117. [PubMed: 9771701]

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

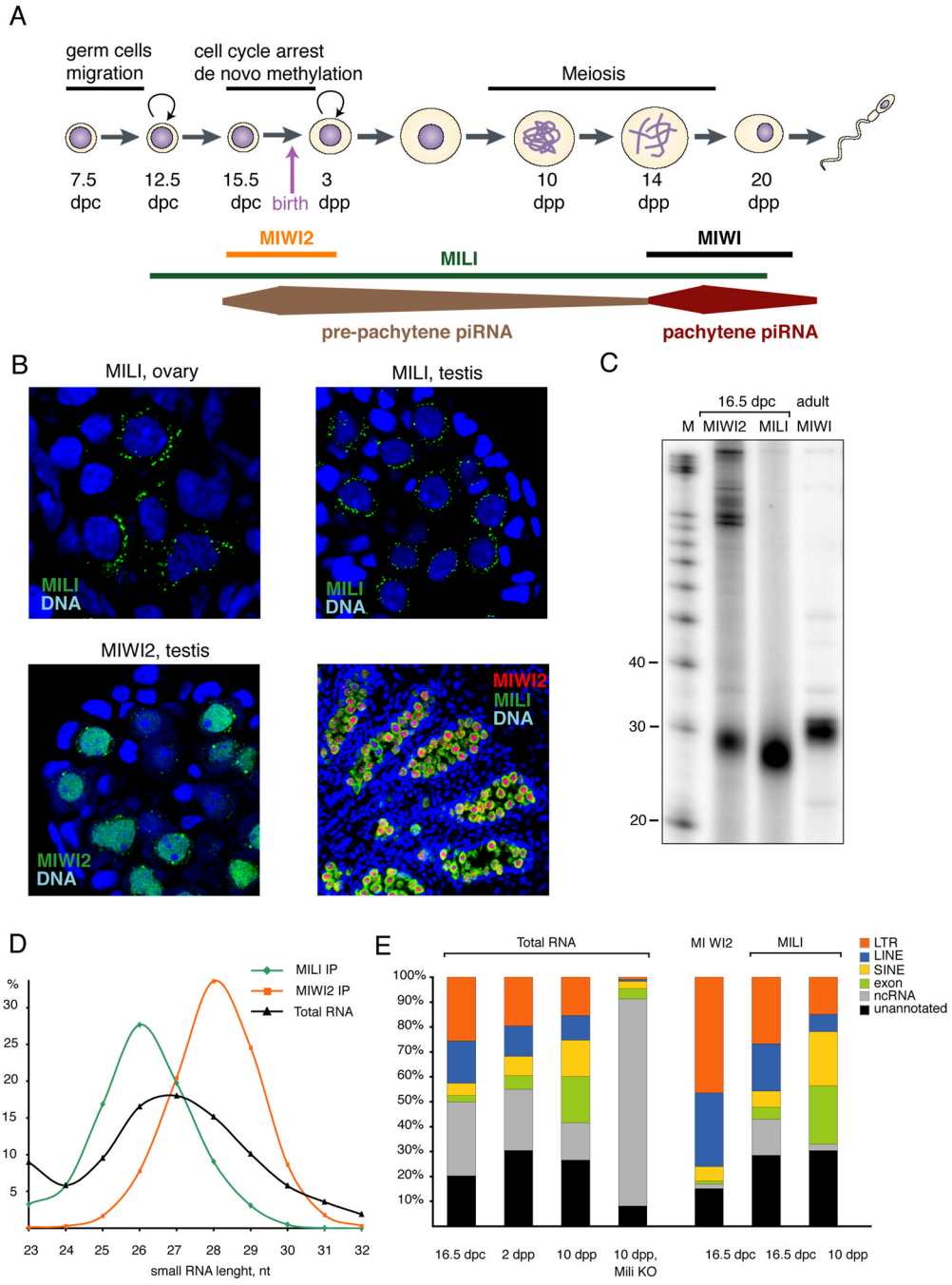

Figure 1. Expression of Piwi proteins though germ cell development

(A) A scheme of spermatogenesis is shown with the timing of expression of *Mili*, *Miwi* and *Miwi2* indicated. After migration, primordial germ cells (PGCs) arrive at the gonad around 11.5 dpc and expand prior to undergoing cell cycle arrest at 15.5 dpc. The timing of cell cycle arrest coincides with establishment of *de novo* DNA methylation patterns on transposable elements and imprinted genes. Germ cells resume division after birth at around 3 dpp and initiate meiotic division at 10 dpp. The first cells at pachytene and haploid round spermatid stages appear at days ~14 and ~20, respectively. **(B)** Expression of GFP-MILI and GFP-MIWI2 transgenes in pre-natal (17.5 dpc) germ cells is shown. MILI exclusively expresses in germ cells and concentrates in perinuclear cytoplasmic granules in both developing testis and ovary.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Miwi2 is absent in female germ cells and localizes in the nucleus as well as in cytoplasmic granules in male germ cells. *Mili* and *Miwi2* are co-expressed in male germ cells at 17.5 dpc and are absent from somatic cells. **(C)** MILI, MIWI2 and MIWI complexes were immunoprecipiated from embryonic gonad at 16.5dpc (MILI and MIWI2) or adult (MIWI) testes. Associated piRNAs were 5' labeled using sequential phosphatase/kinase reactions. **(D)** Size profiles are shown for MILI and MIWI2-bound piRNAs cloned from prenatal testis. **(E)** Small RNA libraries were prepared from size-selected (24–33 nt) total RNA and MILI and MIWI2-immunopurified complexes from the indicated developmental stages and tissue sources. Their small RNA content was classified as indicated.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

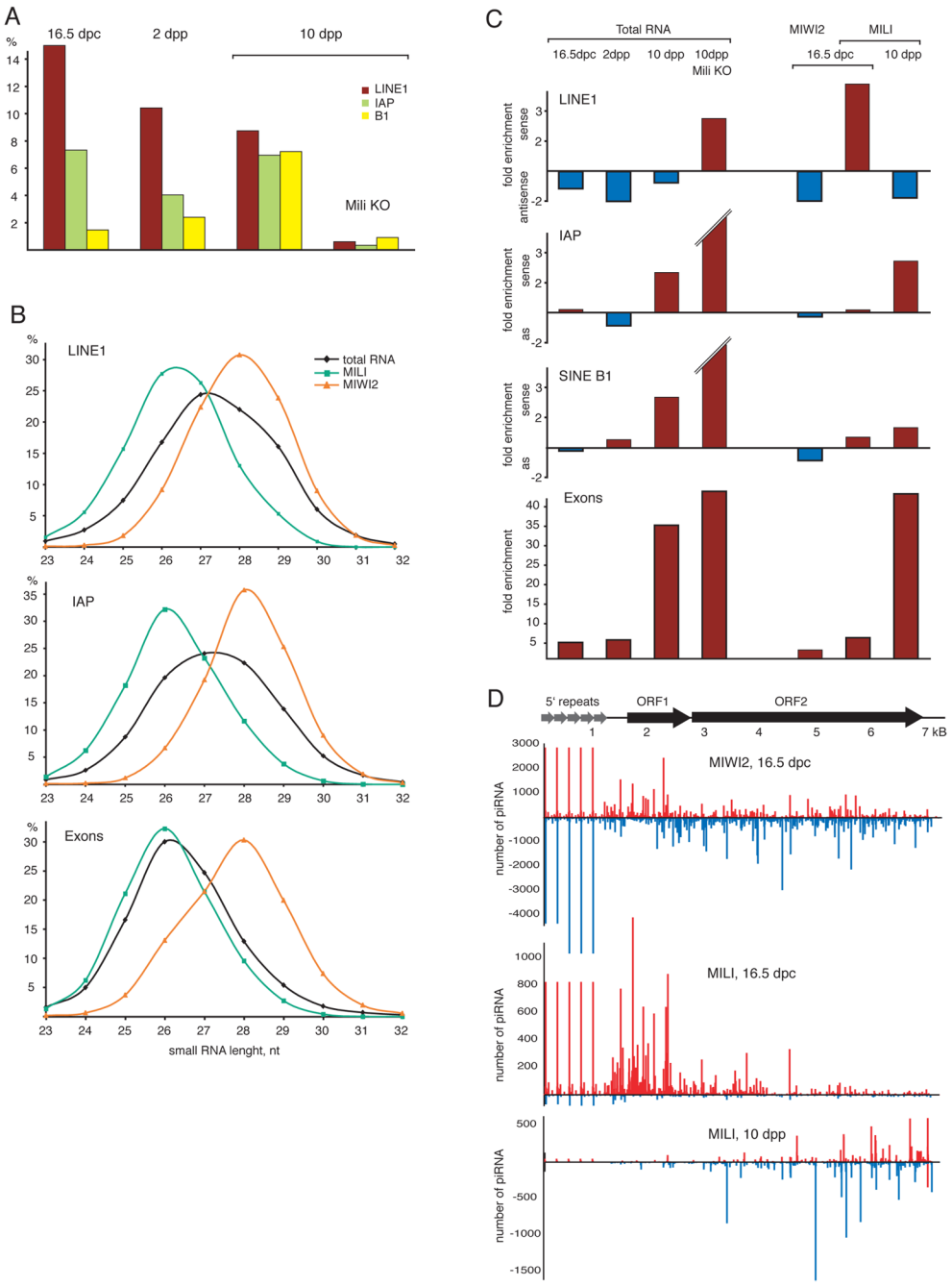

Figure 2. Repeat and gene-derived piRNAs

(A) Shown is the fraction of LINE1, LTR IAP and SINE B1 piRNA in total small RNA libraries at the indicated developmental time points. **(B)** The size distribution of LINE1, IAP and exon-derived piRNAs in total RNA, MILI and MIWI2 complexes is plotted. **(C)** The strand orientation of piRNAs derived from transposable elements and exons of protein-coding genes is shown as fold enrichment for sense (red) or antisense (blue) piRNAs. **(D)** The distribution of piRNAs on LINE1 retrotransposon consensus sequences is shown for the indicated piRNA libraries. piRNAs were mapped to LINE1-A consensus with up to three mismatches. The 5' end of LINE1 consist of several ~200bp repetitive units.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

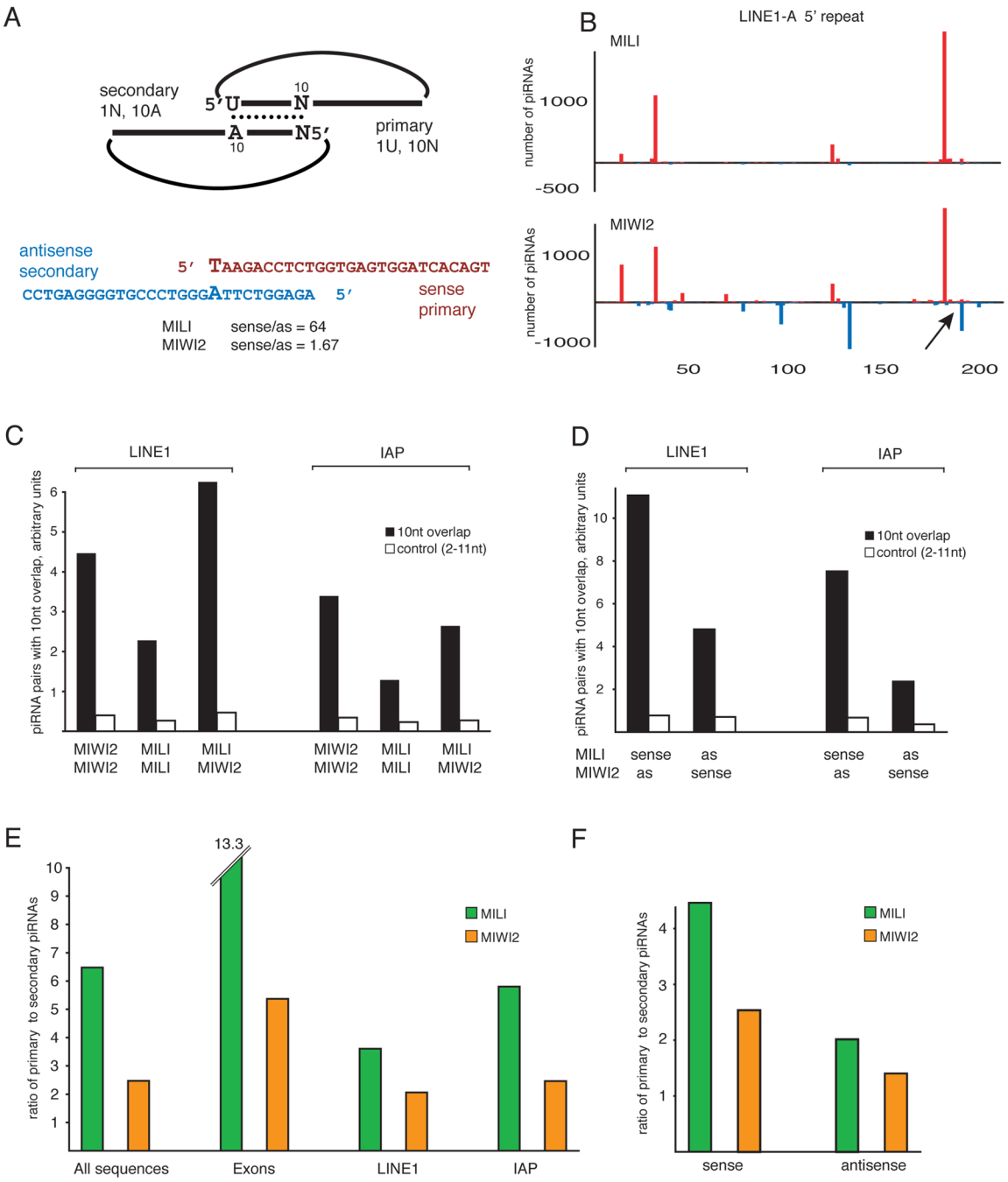

Figure 3. Ping-pong amplification in prenatal piRNAs

(A) A schematic ping-pong pair is shown. piRNAs generated in the ping-pong cycle are complementary to each other and have a 10 nt offset between their 5' ends. Primary piRNAs have a bias for uridine at position 1 and do not have nucleotide bias at position 10. Secondary piRNAs, generated by primary piRNA guided cleavage have a bias for adenine at position 10 and do not have a bias at position 1. Below is an example of an actual ping-pong pair derived from the 5' repeats of LINE1-A as shown in **(B)**. **(B)** The distribution of MILI and MIWI2-associated piRNAs is shown across the 5' repeats of LINE1-A. MILI has a preference for sense piRNAs, while MIWI2 is bound to both sense and antisense sequences. The most prominent ping-pong pair is indicated by arrow and its sequences are shown in **(A)**. **(C)** The ping-pong

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

interaction between piRNAs associated with each Piwi family member was measured as the number of piRNA pairs that have a 10 nt overlap between their 5' ends normalized to the total number of sequences (arbitrary units). Comparison to a control set (overlap at position 2–11) shows that both Piwi proteins are involved in the ping-pong cycle. **(D)** The extent of the ping-pong interaction measured as in **(C)** was calculated for sense and antisense transposon piRNAs separately. **(E)** The ratio of primary (1U, no-10A) to secondary (no-1U, 10A) piRNA was calculated for MILI and MIWI2-bound piRNAs. **(F)** The correlation between strand orientation of transposon-derived piRNA and their processing category is displayed.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

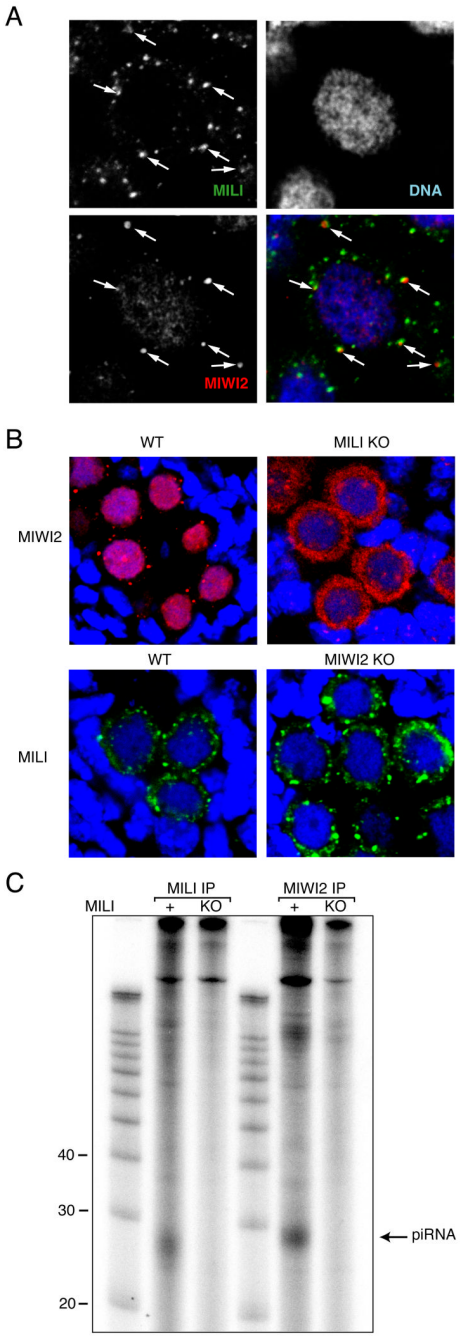

Figure 4. Interaction between MILI and MIWI2 complexes in germ cells

(A) Colocalization of MILI and MIWI2 granules is shown. Detection of MILI and MIWI2 in 17.5 dpc germ cells was performed in Myc-MIWI2 transgenic animals using anti-myc and anti-MILI antibodies. Note that MILI granules are smaller but more numerous as compared to MIWI2 granules. **(B)** MIWI2 localization depends on MILI. MIWI2 was detected in 17.5 dpc prenatal testes of heterozygous and homozygous MILI embryos (upper panels). MIWI2 is present in the germ cells of MILI mutants, but almost completely delocalizes from the nucleus to the cytoplasm where it is uniformly distributed. MILI localization in cytoplasmic granules and does not change in *Miwi2*-deficient animals (bottom panels).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

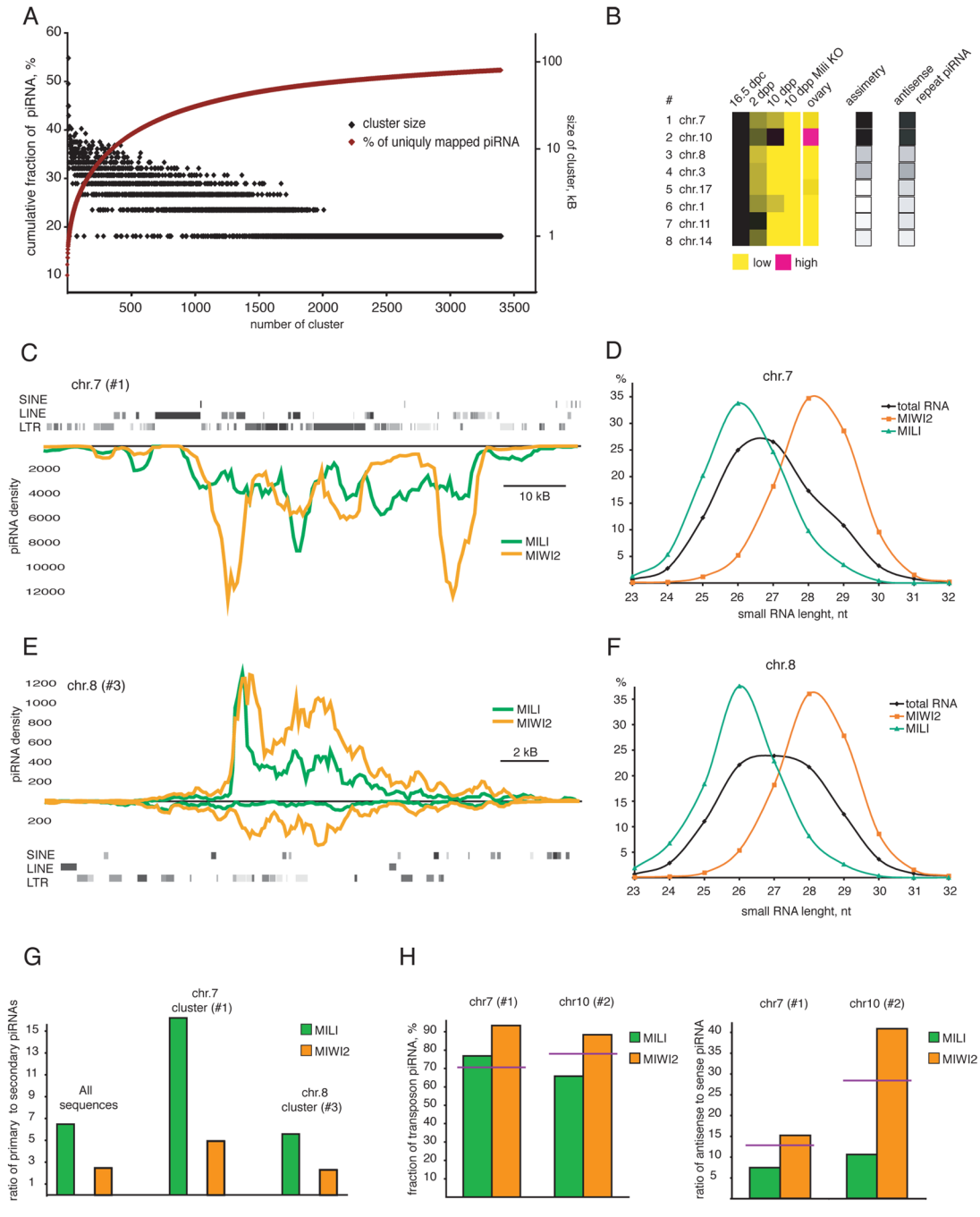

Figure 5. Genomic origins of prenatal piRNAs

(A) Prenatal MIWI2 piRNA clusters were identified by scanning the genome using a 1kB sliding window to find loci that produce at least 10 uniquely-mapped piRNAs per kB. More than 3000 piRNA clusters were thus identified and arranged by the number of uniquely-mapped piRNAs that they produce from left to right. Shown is the genomic size of each cluster (black diamonds) and the cumulative fraction of piRNAs contributed by clusters (red curve). **(B)** Expression patterns and features of the 8 most prominent piRNA clusters. Expression is calculated as a fraction of cluster-derived piRNAs in total RNA populations normalized for expression level at 16.5 dpc. Also shown is the genomic strand orientation of piRNAs produced by each cluster (strand asymmetry) and the fraction of cluster-derived piRNAs that matches

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

the antisense strand of transposons. piRNA density and size profiles of piRNAs for clusters #1 and #3 are shown in **(C)**, **(D)** and **(E)**, **(F)**, respectively. **(C)** piRNA density is shown for the most prominent piRNA cluster (chr. 7: 6526000-6612000). The cluster spans ~ 70 kB and is enriched in transposon sequences (LINE and LTR). The majority of transposons are located on the plus genomic strand and piRNAs are exclusively derived from the minus strand. Therefore, the majority of cluster-derived piRNAs are antisense to transposons. **(D)** The size profile of piRNAs derived from cluster #1 is shown. **(E)** piRNA density is plotted for cluster #3 (chr. 8: 48701000-48723000). This cluster is not significantly enriched in transposon sequences. **(F)** The size profile of piRNAs derived from cluster #3 is shown. **(G)** Processing features are displayed for piRNAs derived from clusters. The ratio of primary (1U, no-10A) to secondary (no-1U, 10A) piRNAs is shown for total MILI and MIWI2-bound populations and for piRNAs uniquely mapped to two piRNA clusters (#1, single-strand, **(C)** and #3, double-strand, **(E)**). For the double strand cluster #3, the ratio of primary to secondary piRNAs in both complexes is similar to that of total population. For the single-strand cluster #1, both MILI and MIWI2 are enriched in primary piRNAs. **(H)** Shown are fractions of repeat-derived piRNAs (left panel) and their sense/antisense ratio (right panel) for the two most prominent piRNA clusters. The genomic sequences of both clusters are enriched in transposable elements. The expected level of repetitive piRNAs and their sense/antisense ratios are shown (red line) if cluster sequences are randomly sampled.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

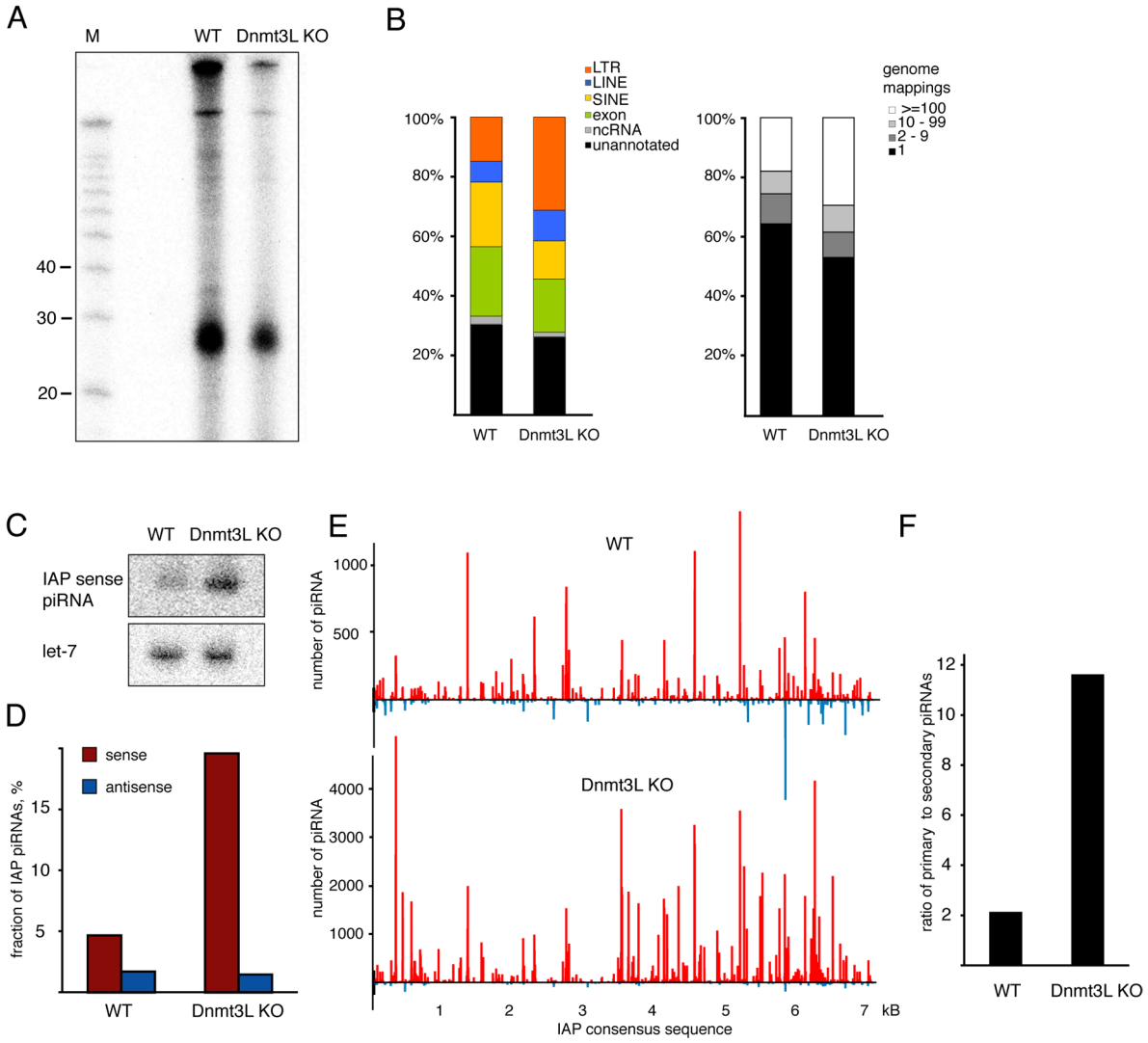

Figure 6. Links between the DNA methylation and piRNA pathways

(A) MILI-piRNA complexes were immunoprecipitated from testes of 10 dpp wild type and *Dnmt3L* knock-out animals, and isolated RNAs were 5' labeled. **(B)** piRNA populations were analyzed in *Dnmt3L* mutants. piRNAs isolated from MILI complexes shown in **(A)** were cloned, sequenced and annotated. Shown are the annotation (left panel) and genomic mapping (right panel) for RNAs from wild-type and *Dnmt3L* mutants. **(C)** Northern hybridization with LNA probe to detect an abundant IAP sense piRNA in 10 dpp testes of wild type and *Dnmt3L* KO mice. Hybridization with let-7 miRNA was used as a loading control. **(D)** The fraction of sense and antisense IAP piRNAs is shown for wild-type and *Dnmt3L* mutant libraries. **(E)** The distribution of piRNAs on the IAP retrotransposon consensus is shown for MILI complexes from wild-type and *Dnmt3L* mutant animals. **(F)** Primary and secondary IAP piRNA in *Dnmt3L* mutants. Shown is the ratio of primary (1U, no-10A) to secondary (no-1U, 10A) piRNAs in the indicated libraries.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

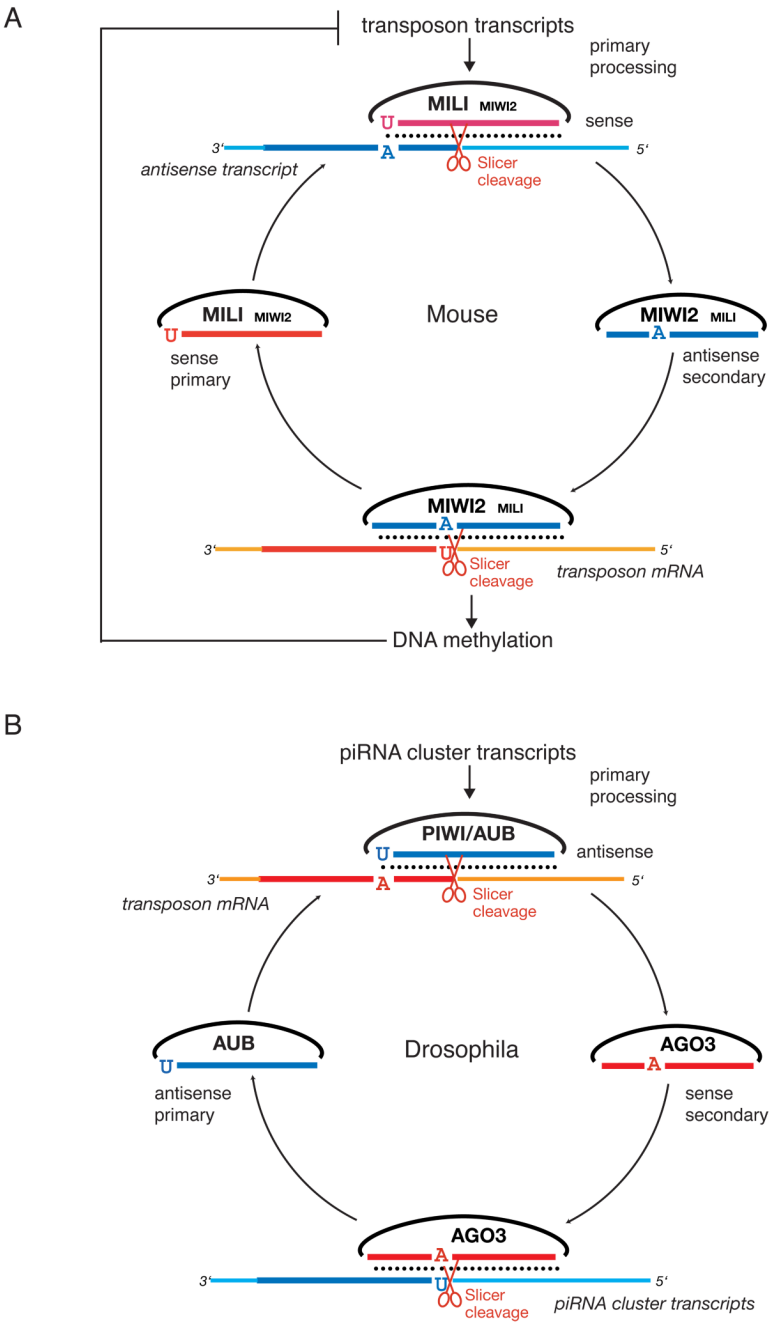

Figure 7. A comparison of ping-pong amplification in mouse and fly

The mouse **(A)** ping-pong cycle is reversed as compared to *Drosophila* **(B)**. In *Drosophila* primary processing of long transcripts derived from piRNA clusters produces both sense and antisense piRNAs that enter a ping-pong cycle that involves AGO3 and AUB. AGO3 primarily binds sense secondary piRNAs and AUB binds primary antisense piRNAs. In mouse piRNA clusters are not the major source of primary piRNAs (Fig. 5A). mRNAs of active transposable elements likely represent the substrate for primary processing resulting in sense piRNAs that preferentially associate with MILI. In prenatal testis both MILI and MIWI2 participate in the amplification cycle. MIWI2 is specifically enriched in secondary antisense piRNAs as compared to MILI. Antisense piRNAs guide DNA methylation of transposable elements

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

sequences in the nucleus probably through recognition of nascent transposon transcripts. After birth, when MIWI2 is no longer expressed, MILI continues to operate the cycle alone. If DNA methylation of transposon sequences is impaired due to downstream mutations in methyltransferase proteins, overexpression of transposon transcripts boosts primary processing and increases the fraction of primary sense piRNAs.