HHS Public Access

Author manuscript

Nat Struct Mol Biol. Author manuscript; available in PMC 2013 January 29.

Published in final edited form as:

Nat Struct Mol Biol. 2012 August ; 19(8): 773–781. doi:10.1038/nsmb.2347.

Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis

Anastassios Vourekas[1], **Qi Zheng**[2,3], **Panagiotis Alexiou**[1,6], **Manolis Maragkakis**[1,6], **Yohei Kirino**[4], **Brian D. Gregory**[2,3,5], and **Zissimos Mourelatos**[1,3]

[1]Department of Pathology and Laboratory Medicine, Division of Neuropathology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[2]Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[3]PENN Genome Frontiers Institute, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[4]Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, California, USA

[5]Genomics and Computational Biology Graduate Program, University of Pennsylvania, Philadelphia, Pennsylvania USA

Abstract

Germ cells implement elaborate mechanisms to protect their genetic material and to regulate gene expression during differentiation. Piwi proteins bind piRNAs, a class of small germline RNAs whose biogenesis and functions are still largely elusive. We employed high throughput sequencing after crosslinking and immunoprecipitation (HITS-CLIP) coupled with RNA-Seq to characterize the genome-wide target RNA repertoire of Mili (Piwil2) and Miwi (Piwil1), two Piwi proteins expressed in mouse postnatal testis. We report the *in vivo* pathway of primary piRNA biogenesis and implicate distinct nucleolytic activities that process Piwi-bound precursor transcripts. Our studies indicate that pachytene piRNAs are the end products of RNA processing. HITS-CLIP demonstrates that Miwi binds spermiogenic mRNAs directly, without utilizing piRNAs as guides, and independent biochemical analyses of testis mRNA-ribonucleoproteins (mRNPs) establishes that Miwi functions in the formation of mRNP complexes that stabilize mRNAs essential for spermiogenesis.

Correspondance should be addressed to Z.M. (mourelaz@uphs.upenn.edu), TEL: +1-215-746-0014, FAX: +1-215-898-9969.

[6]These authors contributed equally to this work.

Supplementary information is available on the Nature Structural & Molecular Biology website.

Accession Codes

All HITS-CLIP and RNA-seq library reads as well as genomic locations deposited in NCBI GEO under the accession numbers GSE27622 and GSE27609 respectively.

Author Contributions

A.V. and Z.M. conceived and directed experiments. A.V. performed HITS-CLIP, RNA-Seq and Nycodenz experiments; Y.K. performed overexpression of Mili and Miwi and *in vitro* crosslinking experiments. Q.Z., M.M., P.A., A.V., B.D.G., Z.M. analyzed data and A.V. wrote the manuscript with input and editing from Q.Z., B.D.G. and Z.M.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Keywords

Argonaute; Mili; Miwi; Piwi; piRNA; miRNA; HITS-CLIP; piRNP; germline; mouse testis; male fertility; mRNA translational regulation

Introduction

Germline development and function engages genes that are conserved in all sexually reproducing metazoans including the Piwi family [1–3] and genes encoding Tudor domain-containing proteins (Tdrds) [4, 5], arranged in functional networks at a genetic and biochemical level [5–9]. Piwi proteins and Tdrds co-localize in germ-plasm/granules, which are cytoplasmic, electron-dense structures such as the intermitochondrial cement (IMC) of spermatocytes and chromatoid bodies of spermatids in mammals [5,10]. Piwi proteins bind piRNAs to form pi-ribonucleoproteins (piRNPs) [11–14]. piRNAs have been implicated in transposon control and a mechanism for amplification of secondary, retrotransposon-derived piRNAs known as “ping-pong” was described (reviewed in Ref. [15]). However, pachytene piRNAs, the most abundant class of mammalian piRNAs, are mainly processed from intergenic precursor transcripts devoid of transposons and repetitive elements [11–13]. The *in vitro* characterization of a 3′-5′ trimming activity for the formation of the 3′ ends of piRNAs was reported [16], and although a model for primary piRNA biogenesis has been hypothesized (reviewed in Ref. [17]), comprehensive, *in vivo* evidence is still lacking. Whether pachytene piRNAs target RNAs *in trans* is also unknown.

The discovery of piRNAs raised the possibility of a germ cell-specific mechanism that utilizes piRNAs in miRNA-like targeting of RNAs for repression and/or degradation. This is an appealing model, given that regulation of mRNA stability and translation is vital for germline development in many organisms[18]. For instance, during mammalian spermiogenesis, early haploid spermatids transcribe many spermatid-enriched/specific genes, whose mRNAs are stored in the cytoplasm in repressed mRNPs until later stages when mRNA translation resumes. Such regulation is necessitated by the transcriptional inactivation of the compacting haploid nuclear chromatin at the onset of spermatid elongation [19,20].

Miwi, Mili and Miwi2 (Piwil4) are the three mouse (*Mus musculus*) Piwi proteins[2,3,21]. In postnatal testis, Mili is expressed until the pachytene stage of meiosis, while Miwi is expressed from pachytene to haploid round spermatid stage; Mili binds to piRNAs that are ~26–28 nucleotides (nt) long, while Miwi binds to piRNAs that are ~29–31 nt and both proteins localize to the cytoplasm [15]. Genetic disruption of Mili and Miwi2 and other genes that are implicated in piRNA biogenesis, such as Mov10l1, Ddx4 (a.k.a. MVH, mouse Vasa homologue), Tdrd1 and Tdrd9 results in male but not female sterility, as a consequence of developmental arrest of spermatogenesis at the pachytene stage; this is attributed to genetic damage caused by rogue retrotransposon activity in the absence of a piRNA silencing mechanism[22–27]. In contrast, deletion of Miwi, Tdrd6 and Tdrd7 causes an arrest at or right after the round spermatid stage[2,28,29] without significant disturbance of piRNA levels or

retrotransposon upregulation [28,29]. In the absence of Miwi, spermiogenic mRNAs are destabilized, contributing to the spermiogenic arrest of the developing germ cells [2].

We report here RNAs targeted *in vivo* by Mili and Miwi in the adult mouse testis, using HITS-CLIP [30]. piRNAs and longer piRNA precursor RNAs bound by both proteins were sequenced, allowing us to propose a coherent model for *in vivo* piRNA biogenesis. Moreover, using HITS-CLIP, biochemical and genetic approaches, we show that Miwi binds spermiogenic mRNAs without piRNA guides, thus participating in the formation of the repressive mRNPs that store spermiogenic messages in postmeiotic spermatids. Combined with RNA-Seq data from five time points of testis development, we provide a comprehensive view of the piRNP molecular pathway, from piRNA processing to the novel function of a Piwi protein as a key regulator of gene expression during spermatogenesis.

Results

A comprehensive view of Piwi RNA targets

To test whether a CLIP approach would be able to identify the RNA targets of Mili and Miwi we first established, using recombinant proteins and synthetic RNAs, that piRNP–RNA target complexes are specific, and that UV can crosslink both piRNAs and larger RNAs, including RNA targets, to Mili and Miwi (Supplementary Fig. 1). CLIP assays (Supplementary Fig. 2a) of Mili and Miwi using testicular tissue resulted in the formation of specific complexes of both proteins with RNA (Fig. 1a, b). cDNA libraries prepared from RNA that was extracted from the membrane segments containing the main radioactive signal are enriched in piRNAs (Fig. 1a, b, blue bars; Supplementary Table 1), while larger complexes extracted from higher-molecular-weight positions are enriched in larger RNAs (Fig. 1a, b, red bars; Supplementary Table 1). Immunoblotting for the presence of known interacting proteins (Tdrd1, Tdrd6, MVH, Mili, Miwi) in IPs of Miwi and Mili verified that CLIP conditions abolished co-immunoprecipitation of protein partners (data not shown). RNA extracted from three sets of control experiments performed in the same stringent conditions (Mili/Miwi IP from non-UV treated testis, Supplementary Fig. 2b; testis CLIP using non-immune serum, Fig. 1a, b; and kidney CLIP using antibodies against Mili and Miwi, Fig. 1a, b) was in non-existent amounts, and attempts to generate cDNA libraries repeatedly failed. Overall, these results demonstrate that our Miwi and Mili CLIP libraries (Fig. 1c, d) are highly specific.

We sequenced 8 libraries, 3 enriched in piRNAs (2 replicates for Mili and 1 for Miwi), and 5 enriched in larger RNAs (large CLIP tags) (2 replicates for Mili and 3 for Miwi) for a total of 58,857,315 mapped reads (Supplementary Table 1). Based on the size distribution of mapped reads and distinct population peaks, we designated the size range of Mili and Miwi piRNAs as 23–31 (peak at 26–27) and 25–33 (peak at 29–30) nt, respectively (Fig. 2a, b). Tags smaller than piRNAs were not analyzed further because it was impossible to determine whether they were derived from piRNAs or piRNA precursors (see below). piRNAs isolated from both proteins show a strong bias for uridine at their 5′ position (Fig. 2c), while larger tags often start with adenosine (Fig. 2d). No significant bias at nucleotide 10 of piRNAs was observed (Fig. 2c) attesting to the absence of “ping-pong”, secondary piRNA amplification. The majority of the piRNAs (60–70%) cluster in repeat-devoid large intergenic hotspots

harboring potential bidirectional transcription start sites described previously [11–13] (Figure 2e). Therefore, CLIP-identified piRNAs bear all the hallmarks of previously identified piRNAs by conventional methods. Furthermore, we analyzed the overlap of CLIP identified Mili and Miwi piRNAs, with Mili and Y12 immunopurified piRNA libraries [6] prepared by IPs under standard conditions (Supplementary Table 2). The overwhelming majority of the CLIP identified piRNAs (65–80%) overlap with standard IP piRNAs. piRNA populations mapping within intergenic piRNA hotspots (see below) are almost identical (>94% overlap) (Supplementary Table 2). Evidently, the conditions and experimental procedures of CLIP do not alter the piRNA profile of Mili and Miwi, and therefore CLIP can reliably identify the *in vivo* RNA cargo of these proteins. Based on the distinct size and sequence characteristics, we segregated mapped reads into piRNAs and larger CLIP tags (lgClips) (Fig. 2a, b, **see also** Supplementary Fig. 3). This sorting is essential for the understanding and accurate description of Piwi protein function. The genomic classification of CLIP tags shows a notable consistency across replicates for both proteins (Fig. 2e, f), indicating the high reliability of our HITS-CLIP approach. The overwhelming majority of non-repeat piRNAs and lgClips located within RefSeq mRNAs have the same strand polarity as the corresponding annotated transcripts in that region (Mili: 85.2% of piRNAs and 92.2% of lgClips; Miwi: 78.8% of piRNAs and 94.5% of lgClips).

To provide a reference for the transcriptome at key time points of testis development, we developed a novel method for directional RNA-Seq (termed Solid Support Directional RNA-Seq, Supplementary Fig. 4a, b) of total RNA purged of ribosomal RNA in order to provide an unbiased account of all transcript classes. We generated two replicate libraries from adult testis (Supplementary Fig. 4c), and one library from each of the following time points in postnatal testis development: 6 days post partum (dpp) (enriched in spermatogonia), 14 dpp (enriched in pachytene spermatocytes), 21 dpp (enriched in haploid round spermatids) and 30 dpp (enriched in elongating spermatids) yielding 43,543,879 mapped reads in total (Supplementary Table 3). Our RNA-Seq data can determine whole transcriptome abundance of testis during key time points of spermatogenesis (Supplementary Fig. 4d).

A model for primary piRNA biogenesis

Intergenic piRNA hotspots (IPHs) lack any defined boundaries, therefore we used a comprehensive statistical model to determine them (see Supplementary Note). We used the designated IPHs and the corresponding intergenic piRNAs, which constitute the most abundant of all CLIP tag classes (Fig. 2e, f), to accurately estimate the reproducibility between replicate HITS-CLIP libraries (R=0.83 – 0.95, Supplementary Fig. 4e, f). IPHs and a few RefSeq mRNAs (throughout their lengths, or 3′-UTRs only [31]) were highly enriched in both small and large tags for both Mili and Miwi (Fig. 3a and Supplementary Fig. 5a, b). Accordingly, piRNA and lgClip densities within IPHs are highly correlated, for both Mili and Miwi (Supplementary Fig. 5c). These observations indicate that both proteins bind the precursor transcripts during primary piRNA biogenesis. Closer examination of Mili bound IPH large tags revealed that they share common 5′ ends with piRNAs (Fig. 3b), indicating that mature 5′ ends of piRNAs are formed before their 3′ ends. The density-plot of all 5′-5′ and 3′-3′ distances between piRNAs and lgClips (piLg-dist) uniquely mapped within IPHs

reveals an extremely high peak at 0 only for 5′-5′ but not 3′-3′ distances (Fig. 3c). Additionally, the 3′ ends of lgClips have a significant preference for extending downstream but not upstream of piRNAs, which is evidenced by the significantly higher positive compared to negative values for 3′-3′ piLg-dist ($p < 2.58\times10\text{-}15$, t-test) but not for 5′-5′ piLg-dist ($p>0.069$, t-test) (Fig. 3c). Moreover, genome browser inspection (Fig. 3b) and the slope of the 3′-3′ distance curve from negative values toward zero (Fig. 3c), indicate progressive shortening of lgClip 3′ ends, suggestive of a 3′ to 5′ exonucleolytic processing for maturation of piRNA 3′ ends. The same analysis performed only for RefSeq exons that contain abundant piRNAs (>30) and lgClips reveals similar patterns for both proteins, supporting a common mechanism for piRNA processing of transcripts with different genomic origins (Supplementary Fig. 5d).

Our results suggest an ordered succession of molecular events, and predict the participation of two distinct ribonucleolytic activities during piRNA biogenesis. Piwi proteins bind the precursor transcript and/or intermediate fragments processed by an unknown endonucleolytic activity. Next, the 5′-U of the intermediate fragments are preferentially recognized and accommodated in the MID domain of Piwi proteins. Subsequently, the 3′ end is processed, most likely by an exonucleolytic activity, before it is 2′-O methylated by Hen1[32–34], which halts further trimming and promotes binding of the methylated piRNA 3′- to the PAZ domain of Piwi proteins [35, 36]. Both Mili and Miwi lgClips originating from IPHs show enrichment in U at the 5′ end position when compared to total lgClips, as expected from our primary piRNA biogenesis model (Fig. 3d), suggesting a common mechanism for the processing of primary piRNAs for these two proteins. Finaly, the bulk of IPH piRNAs (~93%) are processed from large clusters (85 clusters of average size 16,465 nt. Supplementary Fig. 5e–h; Supplementary Table 4) which, according to our RNA-Seq data, are present at the pachytene stage, and largely absent from spermatogonia.

Absence of pachytene piRNA complementary targets

An outstanding question about postnatal mammalian piRNAs is whether they can direct posttranscriptional regulation of complementary target RNAs by piRNPs through direct base pairing, similar to miRNAs loaded in miRNPs. Our data show that, as is the case for target RNAs crosslinked to miRNPs [37, 30], a tripartite piRNP (Piwi-piRNA-complementary target RNA) can be crosslinked efficiently (Supplementary Fig. 1). We reasoned that if piRNAs have complementary targets *in vivo*, fragments of such targets should be enriched in our CLIP libraries containing lgClips. We performed exhaustive searches for full or partial complementarity (Fig. 4a) between lgClips and piRNAs for both Miwi and Mili using combined lgClip-enriched datasets (Supplementary Table 1). We observed that only 5–8% of intergenic and uniquely mapped RefSeq-exon piRNAs have any complementary lgClip sequences (Fig. 4b), and a similar percentage of intergenic and uniquely mapped RefSeq mRNA derived lgClips were paired with any piRNA (Fig. 4b). We examined the complementarity pattern of the piRNAs with potential complementary targets, by plotting the mean percentage of pairing events per nucleotide position, normalized by the number of binding events for each specific piRNA. As a control, we used shuffled piRNA sequences against the same lgClip datasets. This analysis can reveal whether the 5′ or 3′ ends of piRNAs have increased potential to recognize complementary sequences, in other words the

existence of seed regions. Remarkably, the complementarity potential is identical to the control and there is no preference for pairing of the 5′ (or even 3′) end as would be expected if this pairing followed a miRNA-like seed rule (Fig. 4c). This suggests that the observed pattern is due to random complementarity. Repeat-derived piRNAs are more oftenly paired (Fig. 4b), and they exhibit slightly elevated complementarity score compared to the control and other piRNA classes (Fig. 4c), on par with the recently reported role of a piRNA-mediated mechanism of L1 retrotransposon silencing in postnatal testis [27, 38]. Complementary nucleotides of RefSeq exon derived lgClips, display the same degree of conservation for real and shuffled piRNAs, both for Mili and Miwi (Fig. 4d), thus excluding the remote possibility that the few potentially targeted mRNA sequences identified by this analysis are somehow conserved, although individual piRNA sequences are not. Overall, our results strongly suggest that in the postnatal male germ line, Mili and Miwi proteins do not use piRNAs as guides to target non-repeat mRNAs.

Miwi targets spermiogenic mRNAs

In addition to RNAs that are processed into piRNAs, we observed specific mRNAs harboring abundant Miwi lgClips but very few piRNAs (Supplementary Fig. 6a–d), suggesting that these transcripts are not bound by Miwi for piRNA processing. To verify this observation, we plotted the density of piRNAs and lgClips positions side-by-side relative to all RefSeq-mRNAs/exons using the combined lgClip enriched libraries for both proteins separately (Fig. 5a). The density of Mili piRNAs correlates significantly (R = 0.8) with that of lgClips throughout the length of the bound mRNAs, indicating that the binding of Mili to its mRNA targets (lgClips) is functionally coupled with their processing into piRNAs (Fig. 5a). Conversely, Miwi-bound exonic piRNAs are fewer and their density shows a far less significant correlation (R = 0.43), even with a much higher density of lgClips along RefSeq-mRNAs (Fig. 5a). These results suggest that Mili binds mRNA targets specifically for piRNA processing, while Miwi binds to a subset of target transcripts outside a piRNA biogenesis context, potentially to protect or stabilize these mRNAs against degradation and/or piRNA processing. We will refer to such mRNAs as “Miwi-covered” mRNAs. We generated an MA-plot showing generalized-log-odds (glog-odds) for lgClips compared to piRNA abundance (M value), versus total CLIP-tag abundance (lgClip + piRNA) (A value) for all RefSeq exons (Fig. 5b). To reliably identify Miwi-covered mRNAs, we focused our analysis to RefSeq mRNAs that have a Miwi lgClip/piRNA ratio more than 8-fold (glog-odds ≥ 3) across all replicates ($p < 0.05$, t-test). We identified 575 Miwi-covered mRNAs representing 460 unique RefSeq genes that have significantly more Miwi lgClips than piRNAs (Supplementary Table 5). Contrary to 3′ UTR binding that is coupled with piRNA processing, there is a rather uniform distribution of Miwi lgClips throughout the mRNA length, indicating a “beads-on-a-string” type of binding of Miwi on these transcripts (Supplementary Fig. 6a–e). The Miwi-covered mRNAs are highly enriched or are only expressed in the post-meiotic stages of spermatogenesis evidenced by our RNA-Seq datasets (Fig. 5c). This further distinguishes these RNAs from piRNA precursor transcripts, which are expressed through meiosis (Supplementary Fig. 6a–d). Gene Ontology (GO) analysis verified that processes involved in spermatogenesis and male gamete generation are the most enriched ones in Miwi-covered genes (Supplementary Table 6). Importantly, MA plots for IPHs (Supplementary Fig. 5f, g) show similar enrichment of lgClips and piRNAs for

both Mili and Miwi (M values ~0) that is consistent with the role of Piwi in processing of these precursor transcripts, thus emphasizing the contrasting result of the RefSeq exon MA plot revealing mRNAs that are bound but not processed, as described above.

We validated the specificity of the RefSeq mRNA lgClips of Miwi by performing the recently described crosslinking-induced mutation site (CIMS) analysis[39] (Supplementary Fig. 6f, g). We sought to identify any potential sequence (by n-mer analysis) or structural biases (by UNAfold analysis) on the Miwi-covered mRNA areas, which could play a role Miwi binding (Supplementary Note). Our analyses did not reveal sequence motifs for Miwi binding but showed that mRNAs targeted by Miwi have the potential to adopt more energetically favorable secondary structures than random background; however the significance of this finding is not clear at the present time (Supplementary Note).

piRNA-free Miwi engages in spermiogenic, repressed mRNPs

To confirm the direct binding of spermiogenic mRNAs by Miwi, we pursued an independent, biochemical approach. Adult and 28 dpp mouse testis lysates were resolved by isopycnic ultracentrifugation in a Nycodenz density gradient (Fig. 6a and Supplementary Fig. 7a), a system that has been used extensively to define and study repressed spermiogenic mRNPs [40–42]. The distinct distribution of RNA and proteins within the density gradient demonstrates the successful separation of various macromolecular complexes[40–42]: actively translated mRNAs (Gapdh, Ldhc) forming poly-ribosomes (Fig. 6a–c, lanes 1–3), mRNAs packed in repressed mRNPs (Prm1, Tnp2, Smcp, Odf1) (Fig. 6a–c, lanes 4–5), and proteins complexed with smaller RNA species or in free state (Fig. 6a–c, e, lanes 6–10). It is worth noting the presence of Miwi in all fractions, and in particular in fractions 4 and 5 that are marked by the presence of the bulk of repressed spermiogenic mRNAs with Msy2 (Fig. 6d). The bulk of Mili, Miwi and piRNAs are mainly found in fractions 6 and 7 (Fig. 6d and e). A prolonged exposure still failed to identify piRNAs in fractions 4 and 5 (Supplementary Fig. 7b). To verify that Miwi present in fractions 4 and 5 participates in the formation of the repressive mRNPs, we performed oligo-dT pull-downs (Fig. 6f). Miwi was detected by western blot in the protein extract of the oligo-dT pull-down using fractions 4 and 5 (but not 8), and therefore it forms complexes with polyadenylated mRNA in these fractions (Fig. 6f). The total protein content of four independent oligo-dT pull-down experiments using Nycodenz gradient fractions 4 and 5 from adult testis and late round spermatid stage (28 dpp, see below) lysates was analyzed by mass spectrometry. Strikingly, Miwi was among the top protein hits identified in all four samples (Supplementary Table 7) along with Pabpc1,2,4, Ybx1, Msy2,4, which are all proteins that are known to participate in the formation of the repressed spermiogenic mRNPs [43–46].

To verify the binding of spermiogenic mRNAs to Miwi, we performed a reciprocal immunoprecipitation of Miwi (Fig. 6g). Significantly more mRNA from spermiogenic genes is extracted from Miwi immunoprecipitates using fraction 4 (containing the repressed spermiogenic mRNPs) than with any of the fractions 2, 7 or 8 (4–8-fold) (Fig. 6c and g). This represents a two-fold enrichment of such mRNAs in the Miwi immunoprecipitates from fraction 4, compared to their levels in Nycodenz fractions 4 versus 2, supporting the notion that the binding detected by this assay is specific. Mili immunoprecipitates from the

same fractions are not enriched in spermiogenic mRNAs (Supplementary Fig. 7c), further supporting the specificity of the above findings. Remarkably, piRNAs are absent from these repressed mRNPs (Fig. 6g, sample 4) and spermiogenic mRNAs are absent from fractions containing Miwi–piRNA complexes (Fig. 6g, samples 7 and 8). This finding corroborates the results of the bioinformatic analysis and substantiates that piRNAs are not the mediators of Miwi's mRNA binding activity. Thus, both the binding of the piRNA precursors and the binding of the spermiogenic mRNAs by Miwi is a manifestation of the same ability, namely to bind RNA in a manner that is not sequence specific.

In the absence of Miwi, spermiogenesis is arrested at the round spermatid stage, before the onset of spermatid elongation (stage IX) [2]. We performed a careful characterization of postmeiotic spermatids (see also Supplementary Note) and determined that at 28 dpp, elongation has not yet ensued (Fig. 6h, Supplementary Fig. 7a, d) and that the number of round spermatids at that stage is similar in mice heterozygous (HET) or homozygous null (KO) for Miwi (Supplementary Fig. 7d). We found that the levels of intergenic piRNA precursors and LINE1 retrotransposons in 28 dpp Miwi KO testis do not show significant change compared to those from HET mice (Supplementary Fig. 7e). We noticed that Mili levels are not reduced in Miwi KO (Fig. 6h) and that while the total levels of piRNAs are similar in KO and HET testis, the piRNA size is smaller (~26 to 28 nt) and corresponds to the size of piRNAs that are bound by Mili, indicating that Mili can compensate, at least partially, for the loss of Miwi during piRNA precursor processing (Supplementary Fig. 7f). We examined the levels of a number of mRNAs that are bound by Miwi as determined by HITS-CLIP (Ldhc, Odf1, Smcp, Tnp2, Prm1), and exhibit distinct distribution among the polysomes and repressive mRNPs (Fig. 6c, i and Supplementary Fig. 7a). mRNAs highly repressed in post meiotic spermatids (Odf1, Smcp, Tnp2, Prm1) display a dramatic (20-fold to 100-fold) decrease in Miwi KO testis, while mRNAs with predominant or exclusive presence in the polysome fractions show small or no change in abundance (Ldhc and Gapdh, respectively) (Fig. 6i and Supplementary Fig. 7g).

Discussion

Mili and Miwi are integral parts of the piRNA biogenesis machinery, binding the precursor transcripts during nucleolytic processing into mature piRNAs. A model for primary piRNA biogenesis emerges, which predicts the involvement of two distinct nucleolytic activities that process Piwi-bound precursors: an endonuclease, which creates intermediate piRNA precursor fragments, and a 3′-5′ exonuclease that trims the 3′ ends to give rise to mature piRNAs (Fig. 7). Our data provide an *in vivo*, high-throughput demonstration of the *in vitro* recapitulation of the maturation of the piRNA 3′ ends using BmN4 cell extracts reported by the Tomari lab [16]. We predict that at least one of the nucleolytic activities or associated cofactors that generate piRNAs is expressed at much lower levels or ceases to be active or is sequestered in post-meiotic spermatids. Further support for our hypothesis that piRNA biogenesis is markedly reduced in haploid spermatids, comes from the recent discovery that piRNA biogenesis requires the activity of mitoPLD (Pld6, homolog of *Drosophila zucchini*), a mitochondrial phospholipase that catalyzes conversion of mitochondrial cardiolipin to phosphatidic acid [47,48]. mitoPLD is essential for the organization of the IMC, the site where all known factors required for piRNA biogenesis are recruited[47,48]. Pachytene

spermatocytes that generate the bulk of piRNAs in adult mice contain prominent IMC [47] (Supplementary Fig. 7h) contrary to haploid spermatids, which do not contain IMC [49] (Supplementary Fig. 7h).

Most pachytene piRNAs do not act as sequence-specific mediators of Piwi-target RNA interactions. Although we cannot exclude the possibility that piRNAs utilize a “tiny seed” sequence for complementary target recognition that was missed by our analysis, given the extreme sequence diversity of pachytene piRNAs, it is difficult to imagine how such a mechanism could recognize specific RNA targets for regulation. Our results indicate that pachytene piRNAs are the end products of a germ cell degradation mechanism that targets mostly non-coding transcripts with possible meiotic functions. Such functions may include roles in homologous chromosome synapsis and chromosomal crossover or in marking the chromatin for subsequent silencing during spermiogenesis. Alternatively, the appearance of these transcripts may be a byproduct of transcriptional-mediated chromatin remodeling during meiosis. In either case the long non-coding RNAs are cleared in the cytoplasm as piRNAs.

Why do pachytene piRNAs persist if they do not target RNAs? An intriguing possibility is that piRNAs have a “passive” role as part of Piwi-Tdrd containing piRNPs. In such a context, they may be required for the dramatic cytoplasmic remodeling that takes place in spermiogenesis. Part of this process is the formation of the chromatoid body that is best characterized as an aggresome that contains Miwi-piRNPs, Tdrd6, Tdrd7 and many other proteins from different cellular compartments [50, 28]. During the transformation of the round spermatid to sperm, much of the cytoplasm of the early spermatid is shed into the seminiferous tubules (as residual body) including much of the chromatoid body, and is phagocytosed by Sertoli cells [51, 49]. Mutations in genes whose protein products organize the chromatoid body, such as Tdrd6 or Tdrd7 lead to chromatoid body defects and arrest of spermiogenesis [29, 28]. In addition, in later stages of spermiogenesis, transcription ceases and the proteins that are required for sperm formation are translated from pre-existing spermiogenic mRNAs that were previously made and stored. Thus, without new RNA synthesis there is no need for degrading the piRNAs further in order to recycle nucleotides in sperm. Further work is required to address these fascinating possibilities.

Miwi stabilizes mRNAs within translationally repressed mRNPs, by binding directly to spermiogenic mRNAs, without using piRNAs as guides. We propose that direct binding of Miwi onto spermiogenic mRNAs protects them from ribonucleases until the mRNAs are ready for translation in later spermiogenic steps. In that aspect this stabilizing function is similar to the protection that Miwi and other Piwi proteins afford on piRNAs, which are stabilized when bound to Piwi proteins. We envision that as spermiogenesis progresses, the stored mRNPs are remodeled and mRNA translation ensues with displacement of Miwi from coding sequences by the ribosomes but with likely persistence of Miwi in 3′-UTRs (and thus persistence of Miwi in polysomal fractions; Fig. 6d) until Miwi’s ultimate disappearance in later elongating spermatids. Such ribosomal clearing may be similar to the removal of the exon-junction complex from mRNAs during the pioneer round of translation [52]. Alternatively, degradation of Miwi may accompany the remodeling and translation of spermiogenic mRNPs in elongating spermatids.

Are there mechanisms that select RNAs destined to bind to Piwi proteins? Apart from the first nucleotide, Ago proteins bind to the backbone of RNAs, and do not use any base information for selecting their guide RNAs [53]. In the case of miRNAs, the miRNA loading complex and the structure of double-stranded miRNA duplexes are the major mechanisms that select the miRNAs that will bind to Ago proteins under physiological conditions [54]. Our analyses indicate a possible bias of Miwi towards associating with “more structured” areas of mRNAs, although the exact role of elements of RNA secondary structure in the formation of Miwi mRNPs warrants further investigation. Other mechanistic aspects of the loading process such as the roles of helicases and numerous Tdrds are unknown. We favor a stochastic binding of RNAs to Piwi proteins that is largely patterned by the temporal expression of various transcripts and the Piwi-Tdrd germ granule complexes and their cellular compartmentalization. Consequently, Mili and Miwi adopt distinct spatiotemporal expression patterns [3] that directly reflect their functions. Jointly, they process piRNA precursors during their overlapping presence in the pachytene stage of meiosis. Next, Mili levels decline with the appearance of haploid round spermatids, and Miwi actuates the formation of mRNPs that store translationaly inactive spermiogenic mRNAs. Interestingly, cold-shock domain containing proteins are deposited on spermiogenic mRNAs cotrancriptionally[45]. Such mechanism may mark nascent transcripts in the nucleus, and may contribute to their export and sorting in cytoplasmic Piwi-Tdrd RNPs, promoting distinct fates: degradation for piRNA precursors, or storage for developmentally regulated mRNAs.

In conclusion, our study reveals an elegant interplay between Piwi proteins, the germ cell transcriptome and piRNA processing enzymes that is choreographed around the temporal expression patterns of the Piwi proteins themselves, the piRNA generating nucleases and the germline transcriptome. Piwi proteins and associated protein factors and nucleases emerge as the central players in mammalian piRNPs and set a new ground for the understanding of the biological roles and significance of piRNAs and piRNPs in mammalian germline development.

Online Methods

Antibodies

Affinity purified, rabbit polyclonal anti-peptide antibodies (Abs) were generated: Tdrd1, EVGGSKGDRKPPTC; Tdrd6, CETKTSKFYERSTRS (Genscript). Other antibodies used: anti-Mili (17.8) [6], rabbit anti-Miwi [9], rabbit anti-mouse IgG (Jackson Immunoresearch: 315-005-008), E7 anti beta-tubulin (Developmental Studies Hybridoma Bank), anti-Msy2 (C-13; Santa Cruz sc-21321), anti-Tnp2 (a generous gift by W. Stephen Kistler, University of South Carolina) [19,56].

Histology and EM

Miwi mice, developed by the Lin lab, [2] were purchased from MMRRC – University of Missouri (mouse strain 029995). Testes from adult and various dpp (±18 hours) from Miwi heterozygotes and homozygotes were fixed in Bouin’s, embedded in paraffin and stained with Hematoxylin and Eosin. Electron Microscopy of testes from Miwi HET ad Miwi KO mice was performed at the Electron Microscopy Resource Laboratory at PENN.

Preparation of recombinant Miwi and Mili, and *in vitro* crosslinking with piRNAs and complementary targets

Flag-Miwi and Flag-Mili were expressed in baculovirus-infected, Sf9 cells and purified with anti-Flag M2 agarose (Sigma) as previously described[6]. For crosslinking experiments between recombinant Piwi proteins and piRNA, we followed a similar strategy to that previously described [37]. Crosslinking was performed on ice by irradiation for 30 min with a 365 nm hand-held lamp (EL series UV lamp, UVP). Cross-linked proteins were separated by NuPAGE (4–12 % Bis-Tris, Invitrogen) and detected by storage-phosphor autoradiography. For crosslinking between Piwi/piRNA complex and its target RNA, we first loaded Piwi protein with excess cold piRNA followed by washings to remove unbound piRNA. The Piwi/piRNA complex was incubated with a 3′-end labeled target RNA containing s^4U at 28 °C for 60 min in lysis buffer, followed by crosslinking.

Mili and Miwi HITS-CLIP

Adapted from the Argonaute HITS-CLIP [30] with stated modifications. Testes from C57BL/6J mice were harvested, detunicated and kept in ice cold HBSS (GIBCO 14175). A cell suspension in HBSS was created in 10 cm plates (always kept on ice) by mild pipetting followed by immediate UV irradiation (3x) at 254 nm (400mJ/cm^2) in Stratalinker 1800 (Stratagene). The cells were pelleted, washed with PBS, and the final cell pellet was flash-frozen in liquid nitrogen and kept at −80 °C.

For each immunoprecipitation, 150 μL of protein A Dynabeads® slurry (Invitrogen 100-02D) was used. For Mili antibody, 8 μL rabbit anti-mouse IgG was incubated with the beads, in 350 μL Ab binding buffer for 45 min at RT; beads were incubated in 350 μL Ab binding buffer (0.1 M Na-phosphate pH 8, 0.1% NP-40) with 10 μL of 17-8 ascites fluid at 4 °C for ~4 hrs. For Miwi antibody, 15μg were added to the beads at 4 °C for ~2–3 hours; beads were washed with 1xPXL (1X PBS (no Mg^{2+}/Ca^{2+}), 0.1% SDS, 0.5% deoxycholate, 0.5% NP-40) or 1xPMPG (1X PBS (no Mg^{2+}/Ca^{2+}) 2% Empigen). Mili CLIP was performed with 1xPXL (1xPMPG is also suitable), while Miwi CLIP with 1xPMPG.

For each IP, UV treated cells from two testis were lysed in 300 μL of 1X PXL or 1xPMPG with protease inhibitors and RNasin (2u/μL); lysates were treated with DNase (Promega) for 10 min at 37 °C. For optimal retrieval of long crosslinked RNAs, lysates for Mili CLIP were treated with 5U of RNase T1 (Roche 10109193001) per 250 μL, and lysates for Miwi CLIP were not treated with exogenous RNases.

Cell lysates were centrifuged at 90,000 × g for 30 min at 4 °C. Antibody beads were incubated with lysates for 3 h at 4 °C. Low and high salt washes of IP beads were performed with 1x and 5xPXL or 1x and 5xPMPG (5x is 5xPBS, detergents same as 1x) respectively.

RNA linkers (RL3 and RL5), DNA primers (DP3 and DP5, DSFP3 and DSFP5), and Illumina sequencing primer were as described previously [30], as well as 3′ adapter labeling and ligation to CIP treated RNA CLIP tags.

IP beads were eluted at 70 °C for 12 min using 30 μL of 2x Novex reducing loading buffer. Samples were analyzed by NuPAGE. Crosslinked RNA-protein complexes were transferred

on nitrocellulose (Invitrogen LC2001) and the membrane was exposed to film for one to two hours. The main radioactive signal of the immunoprecipitated crosslinked RNA-protein complex appears ≥7 kDa above the apparent molecular weight of the protein alone as expected (Fig. 1a, b). Membrane fragments containing the main radioactive signal and fragments up to ≥15 kDa higher were cut. RNA extraction and 5′ linker ligation were performed as described previously [30].

RT-PCR and re-amplification were performed as described previously [30]. Reaction products from both reactions were analyzed on 3% Metaphor 1xTAE/EtBr gels. piRNAs form a distinct band at ~70 bps and ~120 bps after first and second PCR respectively, while cDNA from larger RNAs forms a smear extending to larger sizes. DNA was extracted with QIAquick Gel Extraction Kit and submitted for Illumina deep sequencing.

Solid Support Directional RNA-Seq (SSD RNA-Seq)

Total RNA from mouse adult testis (C57BL/6J) was isolated using TRIZOL and treated with DNase. To remove the ribosomal RNA from 10 μg of total RNA, the RiboMinus Eukaryote kit (Invitrogen, A10837) was used following manufacturer's instructions. 2 μg of ribosomal RNA-depleted total RNA (RiboMinus-RNA) were treated with fragmentation buffer (Ambion, AM8740) at 70 °C for 10 min. The reaction was stopped by adding 1 μL 10x stop solution, and was desalted using G25 Sephadex microspin column (GE 27-5325-01). Fragmented RiboMinus-RNA was treated with Antarctic phosphatase at 37 °C for 30 min, extracted with phenol and precipitated with ethanol. 50 pmoles RL3-3′ biotin adapter (IDT) were ligated to the fragmented RiboMinus-RNA using T4 RNA ligase in 80 μL of reaction mixture at 16 °C overnight.

100 μL of M280Streptavidin Dynabeads® (Invitrogen 112.05D) slurry were used. The ligation mixture was incubated with the beads for 15 min at RT in 0.5x BW buffer (1x: 10mM Tris pH 7.5, 2M NaCl, 1 mM EDTA). After washing twice with BW buffer and twice with PNW buffer (50 mM Tris pH 7.4, 5 mM $MgCl_2$, 0.1% NP-40), the beads were resuspended in a T4 polynucleotide kinase reaction mixture (80 μL) for the addition of 5′ phosphate to the captured RNA. The beads were washed, resuspended in the 5′ adaptor (RL5, 150 pmoles in reaction) ligation mixture (80 μL) at 16 °C for 6 hours, then washed twice with BW buffer and twice with H_2O. For reverse transcription (RT), the beads were split in two aliquots, and each aliquot was mixed with Titan One Tube RT-PCR System (11888382001 Roche); RT was performed at 50 °C for 45 min, and at 55 °C for 15 min. PCR, RE-PCR and agarose gel analysis (with care to exclude adapter-dimers, see Supplementary Fig. 4b) were performed as in HITS-CLIP.

Nycodenz density gradient ultracentrifugation and subsequent analyses

Nycodenz density gradient separation of RNPs was performed as previously described [40–42] with modifications. A 20%–60% (top to bottom) Nycodenz gradient (4.8 mL) in 1xKMH (200 mM KCl, 2 mM $MgCl_2$, 20 mM HEPES pH 7.4, 0.5% NP-40, 0.1U/μL rRNAsin, and protease inhibitors) was prepared as a step gradient by overlaying 5 equal parts of Nycodenz solutions and was let to diffuse overnight at 4 °C. 0.2 mL of post nuclear testis lysate in 1xKMH was laid over the gradient and centrifuged at 150,000 g for 20 hours. The gradient

was collected in 10 equal fractions. Samples from each fraction were used for protein determination by Bradford and RNA extraction with Trizol LS. Right before RNA extraction, 500 ng of *in vitro* transcript of Renilla Luciferase (rLuc) mRNA was added to each fraction for normalization purposes in subsequent steps. There was no discernible interference by the presence of Nycodenz with any of the analyses carried out using gradient fractions.

Oligo-dT pull down experiments

100 μL of Dynabeads® Oligo $(dT)_{25}$ suspension (Invitrogen 610-11) were washed twice with 1xKMH, and were mixed with 100 μL of Nycodenz fraction, diluted with equal volume of 1xKMH. The suspension was rotated gently for 20 min at 4°C, and was then washed three times with 1xKMH buffer. Two thirds of the beads were eluted with 2x SDS reducing loading buffer (Invitrogen NP0007), and the rest were used for RNA extraction with TRIZOL.

qRT-PCR

Equal volume of RNA extracted from each fraction was reverse transcribed by Supersript III (Invitrogen 18080-051) in the presence of random hexamers. Equal volume of the cDNA was mixed with primers (Qiagen QuantiTect Assay, and: rLuc F: 5′-CGCTGAAAGTGTAGTAGATGTG, R: 5′-TCCACGAAGAAGTTATTCTCCA; Line1 (L1Md_T) pair A: F: 5′-AGATTCCATAGAGAACATCGG, R: 5′-CTTGTTGAAGATGTTTGCTGG; pair B: F: 5′-ACCCAAACCTATGGGACACA, R: 5′-CTGCCGTCTACTCCTCTTGG; IPH #494 F: 5′-TGTTCACTTACATATCAGGGTC, R: 5′-GTAAAGCCCAAGAGCAAGAG; IPH #204 F: 5′-AGTCTGTGTAGTAGTTTCCTGAG, R: 5′-TGTCCACTTCCATGTTACCT) and Power SYBR Green reaction mix (Applied Biosystems 4367659). The reactions were run on StepOnePlus™ (Applied Biosystems).

Mass spectrometry

Oligo-dT eluates from pull-down experiments from Nycodenz gradient fractions from testes from adult and 28 dpp mice were analyzed by mass spectrometry at Taplin Mass Spectrometry Facility, Harvard as entire eluates and also from individual bands cut out from silver-stained gels. Individual bands from silver stained-gels from Tdrd6 and Tdrd1 immunoprecipitates were subjected to mass spectrometry at the Proteomics Core Facility, PENN.

Bioinformatic analysis

Detailed bioinformatic methods can be found in the Supplementary Information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Apologies to colleagues whose papers could not be cited due to space limitations. We are grateful to Jonathan Schug for assistance with Illumina analysis approaches in the early phase of this project and to Svetlana Fayngerts, Stavros Rafail, and members of the Mourelatos lab for technical help and lively discussions. The antibody against Tnp2 was a generous gift from W. S. Kistler (Department of Chemistry and Biochemistry, University of South Carolina, Columbia, South Carolina, USA). Supported by NIH grant GM0720777 and in part by IRM-PENN and PSOM-PENN grants to ZM. Supported in the Gregory lab in part by Grant #IRG-78-002-30 from the American Cancer Society, as well as the Penn Genome Frontiers Institute and a grant with the Pennsylvania Department of Health. The Department of Health specifically disclaims responsibility for any analyses, interpretations, or conclusions.

References

1. Cox DN, et al. A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev. 1998; 12:3715–27. [PubMed: 9851978]
2. Deng W, Lin H. miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev Cell. 2002; 2:819–30. [PubMed: 12062093]
3. Kuramochi-Miyagawa S, et al. Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development. 2004; 131:839–49. [PubMed: 14736746]
4. Boswell RE, Mahowald AP. tudor, a gene required for assembly of the germ plasm in Drosophila melanogaster. Cell. 1985; 43:97–104. [PubMed: 3935320]
5. Chen C, Nott TJ, Jin J, Pawson T. Deciphering arginine methylation: Tudor tells the tale. Nat Rev Mol Cell Biol. 2011
6. Kirino Y, et al. Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. Nat Cell Biol. 2009; 11:652–8. [PubMed: 19377467]
7. Nishida KM, et al. Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines. EMBO J. 2009; 28:3820–31. [PubMed: 19959991]
8. Vagin VV, et al. Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. Genes Dev. 2009; 23:1749–62. [PubMed: 19584108]
9. Kirino Y, et al. Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization. RNA. 2010; 16:70–8. [PubMed: 19926723]
10. Yokota S. Historical survey on chromatoid body research. Acta histochemica et cytochemica. 2008; 41:65–82. [PubMed: 18787638]
11. Lau NC, et al. Characterization of the piRNA complex from rat testes. Science. 2006; 313:363–7. [PubMed: 16778019]
12. Aravin A, et al. A novel class of small RNAs bind to MILI protein in mouse testes. Nature. 2006; 442:203–7. [PubMed: 16751777]
13. 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]
14. Grivna ST, Beyret E, Wang Z, Lin H. A novel class of small RNAs in mouse spermatogenic cells. Genes Dev. 2006; 20:1709–14. [PubMed: 16766680]
15. Siomi MC, Sato K, Pezic D, Aravin AA. PIWI-interacting small RNAs: the vanguard of genome defence. Nat Rev Mol Cell Biol. 2011; 12:246–58. [PubMed: 21427766]
16. Kawaoka S, Izumi N, Katsuma S, Tomari Y. 3′ End Formation of PIWI-Interacting RNAs In Vitro. Molecular cell. 2011; 43:1015–22. [PubMed: 21925389]
17. Pillai RS, Chuma S. piRNAs and their involvement in male germline development in mice. Dev Growth Differ. 2012
18. Besse F, Ephrussi A. Translational control of localized mRNAs: restricting protein synthesis in space and time. Nat Rev Mol Cell Biol. 2008; 9:971–80. [PubMed: 19023284]
19. Heidaran MA, Showman RM, Kistler WS. A cytochemical study of the transcriptional and translational regulation of nuclear transition protein 1 (TP1), a major chromosomal protein of mammalian spermatids. The Journal of cell biology. 1988; 106:1427–33. [PubMed: 3372585]

20. Kleene KC. Patterns of translational regulation in the mammalian testis. Molecular reproduction and development. 1996; 43:268–81. [PubMed: 8824926]
21. Carmell MA, et al. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell. 2007; 12:503–14. [PubMed: 17395546]
22. Reuter M, et al. Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. Nat Struct Mol Biol. 2009; 16:639–46. [PubMed: 19465913]
23. Shoji M, et al. The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. Developmental cell. 2009; 17:775–87. [PubMed: 20059948]
24. Kuramochi-Miyagawa S, et al. MVH in piRNA processing and gene silencing of retrotransposons. Genes Dev. 2010; 24:887–92. [PubMed: 20439430]
25. Frost RJ, et al. MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. Proc Natl Acad Sci U S A. 2010; 107:11847–52. [PubMed: 20547853]
26. Zheng K, et al. Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway. Proc Natl Acad Sci U S A. 2010; 107:11841–6. [PubMed: 20534472]
27. De Fazio S, et al. The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. Nature. 2011; 480:259–263. [PubMed: 22020280]
28. Tanaka T, et al. Tudor domain containing 7 (Tdrd7) is essential for dynamic ribonucleoprotein (RNP) remodeling of chromatoid bodies during spermatogenesis. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108:10579–84. [PubMed: 21670278]
29. Vasileva A, Tiedau D, Firooznia A, Muller-Reichert T, Jessberger R. Tdrd6 is required for spermiogenesis, chromatoid body architecture, and regulation of miRNA expression. Current biology: CB. 2009; 19:630–9. [PubMed: 19345099]
30. Chi SW, Zang JB, Mele A, Darnell RB. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature. 2009; 460:479–86. [PubMed: 19536157]
31. Robine N, et al. A broadly conserved pathway generates 3′UTR-directed primary piRNAs. Curr Biol. 2009; 19:2066–76. [PubMed: 20022248]
32. Gunawardane LS, et al. A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in Drosophila. Science. 2007; 315:1587–90. [PubMed: 17322028]
33. Horwich MD, et al. The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr Biol. 2007; 17:1265–72. [PubMed: 17604629]
34. Kirino Y, Mourelatos Z. The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs. RNA. 2007; 13:1397–401. [PubMed: 17652135]
35. Simon B, et al. Recognition of 2′-O-methylated 3′-end of piRNA by the PAZ domain of a Piwi protein. Structure. 2011; 19:172–80. [PubMed: 21237665]
36. Tian Y, Simanshu DK, Ma JB, Patel DJ. Structural basis for piRNA 2′-O-methylated 3′-end recognition by Piwi PAZ (Piwi/Argonaute/Zwille) domains. Proc Natl Acad Sci U S A. 2011; 108:903–10. [PubMed: 21193640]
37. Kirino Y, Mourelatos Z. Site-specific crosslinking of human microRNPs to RNA targets. RNA. 2008; 14:2254–9. [PubMed: 18719246]
38. Reuter M, et al. Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. Nature. 2011
39. Zhang C, Darnell RB. Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. Nat Biotechnol. 2011; 29:607–14. [PubMed: 21633356]
40. Herbert TP, Hecht NB. The mouse Y-box protein, MSY2, is associated with a kinase on non-polysomal mouse testicular mRNAs. Nucleic acids research. 1999; 27:1747–53. [PubMed: 10076007]
41. Bagarova J, Chowdhury TA, Kimura M, Kleene KC. Identification of elements in the Smcp 5′ and 3′ UTR that repress translation and promote the formation of heavy inactive mRNPs in spermatids by analysis of mutations in transgenic mice. Reproduction. 2010; 140:853–64. [PubMed: 20876225]

42. Kleene KC, Bagarova J, Hawthorne SK, Catado LM. Quantitative analysis of mRNA translation in mammalian spermatogenic cells with sucrose and Nycodenz gradients. Reproductive biology and endocrinology: RB&E. 2010; 8:155. [PubMed: 21184686]
43. Giorgini F, Davies HG, Braun RE. Translational repression by MSY4 inhibits spermatid differentiation in mice. Development. 2002; 129:3669–79. [PubMed: 12117816]
44. Kimura M, Ishida K, Kashiwabara S, Baba T. Characterization of two cytoplasmic poly(A)-binding proteins, PABPC1 and PABPC2, in mouse spermatogenic cells. Biology of reproduction. 2009; 80:545–54. [PubMed: 19020299]
45. Yang J, Medvedev S, Reddi PP, Schultz RM, Hecht NB. The DNA/RNA-binding protein MSY2 marks specific transcripts for cytoplasmic storage in mouse male germ cells. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:1513–8. [PubMed: 15665108]
46. Yang J, et al. Absence of the DNA-/RNA-binding protein MSY2 results in male and female infertility. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:5755–60. [PubMed: 15824319]
47. Huang H, et al. piRNA-associated germline nuage formation and spermatogenesis require MitoPLD profusogenic mitochondrial-surface lipid signaling. Dev Cell. 2011; 20:376–87. [PubMed: 21397848]
48. Watanabe T, et al. MITOPLD is a mitochondrial protein essential for nuage formation and piRNA biogenesis in the mouse germline. Dev Cell. 2011; 20:364–75. [PubMed: 21397847]
49. Hermo L, Pelletier RM, Cyr DG, Smith CE. Surfing the wave, cycle, life history, and genes/proteins expressed by testicular germ cells. Part 2: changes in spermatid organelles associated with development of spermatozoa. Microsc Res Tech. 2010; 73:279–319. [PubMed: 19941292]
50. Haraguchi CM, et al. Chromatoid bodies: aggresome-like characteristics and degradation sites for organelles of spermiogenic cells. J Histochem Cytochem. 2005; 53:455–65. [PubMed: 15805420]
51. Austin CR, Sapsford CS. The development of the rat spermatid. J R Microsc Soc. 1951; 71:397–406. [PubMed: 14939269]
52. Dostie J, Dreyfuss G. Translation is required to remove Y14 from mRNAs in the cytoplasm. Curr Biol. 2002; 12:1060–7. [PubMed: 12121612]
53. Faehnle CR, Joshua-Tor L. Argonaute MID domain takes centre stage. EMBO reports. 2010; 11:564–5. [PubMed: 20634804]
54. Liu X, Jin DY, McManus MT, Mourelatos Z. Precursor MicroRNA-Programmed Silencing Complex Assembly Pathways in Mammals. Mol Cell. 2012; 46:507–17. [PubMed: 22503104]
55. Oko RJ, Jando V, Wagner CL, Kistler WS, Hermo LS. Chromatin reorganization in rat spermatids during the disappearance of testis-specific histone, H1t, and the appearance of transition proteins TP1 and TP2. Biology of reproduction. 1996; 54:1141–57. [PubMed: 8722637]
56. Alfonso PJ, Kistler WS. Immunohistochemical localization of spermatid nuclear transition protein 2 in the testes of rats and mice. Biology of reproduction. 1993; 48:522–9. [PubMed: 8452928]

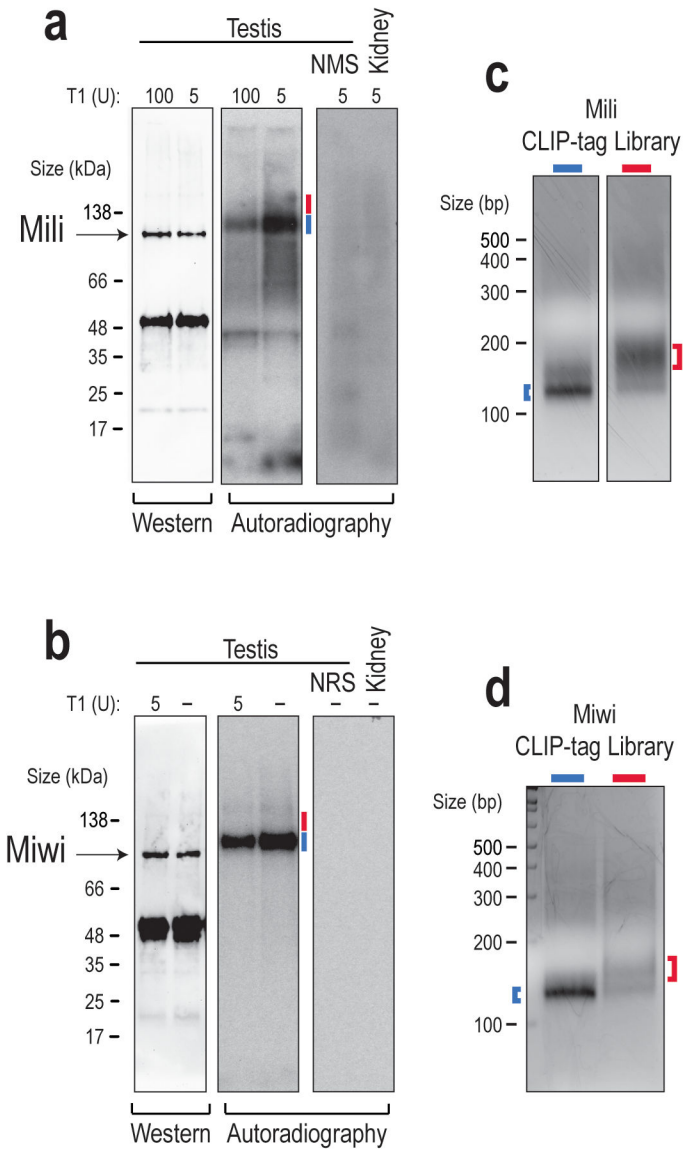

Figure 1. Mili and Miwi HITS-CLIP

(a, **b)** CLIPs were performed using highly stringent conditions with buffers containing 2% Empigen, a potent zwitterionic detergent. Autoradiograms and Western blots of immunoprecipitated, UV-crosslinked, RNA-protein complexes ligated to radiolabeled 3′ adapter are shown for Mili **(a)** and Miwi **(b)**. Negative control CLIPs were performed using non-immune mouse (NMS) or rabbit serum (NRS) with crosslinked testis, and anti-Mili or anti-Miwi antibodies with crosslinked kidney. Cells were lysed after crosslinking and the lysates were incubated in the absence or presence of RNase T1 prior to immunoprecipitation, as indicated. RNA (CLIP tags) was extracted from the membranes after cutting at indicated areas: a blue line marks the major radioactive signal containing mainly piRNA-protein complexes; a red line marks larger RNA-protein complexes which appear as a smear extending to higher molecular weights.

(c, **d)** cDNA libraries from Mili **(c)** and Miwi **(d)** CLIP tags were prepared by RT-PCR, gel purified and sequenced with Illumina. Blue and red brackets denote piRNA and large tag enriched cDNA samples, respectively.

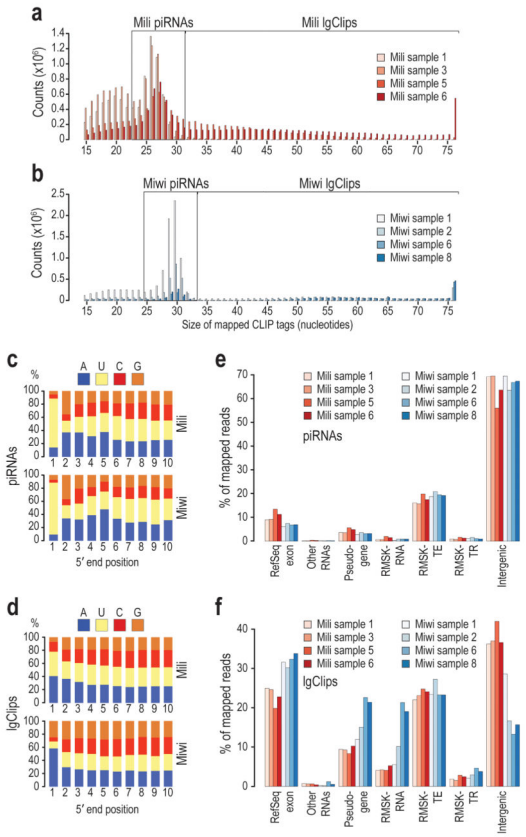

Figure 2. Size distribution, nucleotide preference and genomic classification of mapped CLIP reads

(a, **b)** Size distribution of mapped HITS-CLIP reads for Mili (**a**) and Miwi (**b**)

(c, **d)** Nucleotide preference of the first ten positions for piRNAs (**c**) and lgClips (**d**).

(e, **f)** Genomic classification of piRNAs (**e)** and lgClips (**f**). RMSK: Repeat Masker, TR: terminal repeats, TE: transposable elements.

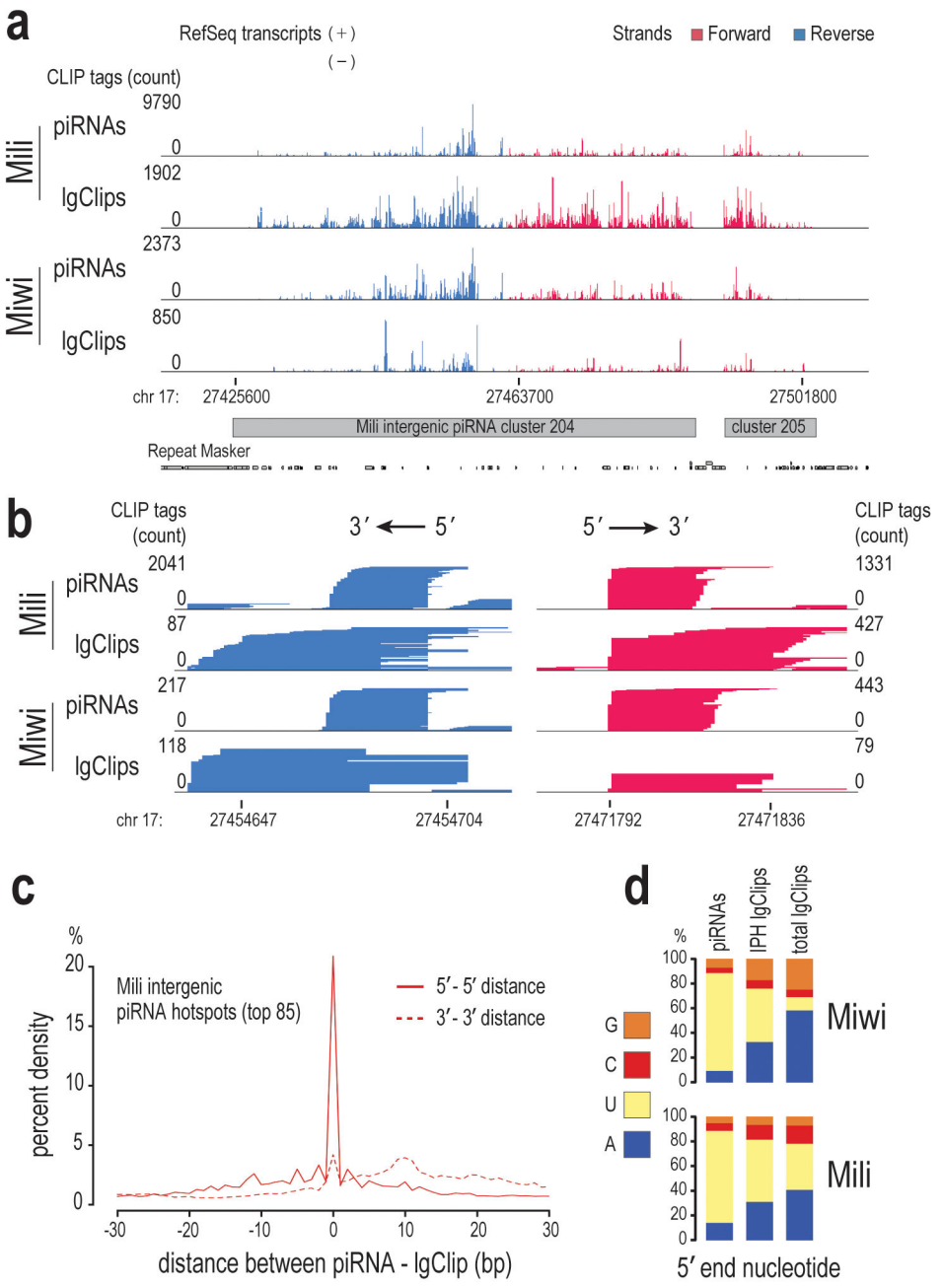

Figure 3. Processing of intergenic piRNA precursors

(a) Genome browser snapshot showing mapped CLIP-tags on an intergenic piRNA hotspot (IPH). Magenta = forward, blue = reverse tags. Numbers on the side of tracks denote number of stacked tags. Genomic coordinates (mm9) are marked below the tracks.

(b) Zoom-in view of **(a)**, revealing piRNA processing. Note the concurrence of 5′ ends between piRNAs and lgClips (intermediate piRNA precursors), and progressive shortening of lgClip 3′ ends.

(c) Density plot of piRNA-lgClip distances (piLg-dist) for CLIP tags that map within IPHs. The distance is defined as the genomic position of lgClip end (5′ or 3′), minus the genomic

position of the respective piRNA end (5′ or 3′) for the forward strand, and vice versa for the reverse strand. On the Y-axis, the percent of piLg-dist over distances for all pairwise combinations between lgClips and piRNAs is plotted. Only piRNAs that lie within a distance of 30 nucleotides upstream or downstream of any given lgClip were considered. Distance measuring requires unique genomic coordinates; therefore only uniquely mapped tags were used in this analysis.

(d) Nucleotide preference at the 5′ end position of piRNAs and lgClips originating from IPHs, and total lgClips, for the indicated proteins. IPH lgClips are enriched for 5′ U in comparison with total lgClips.

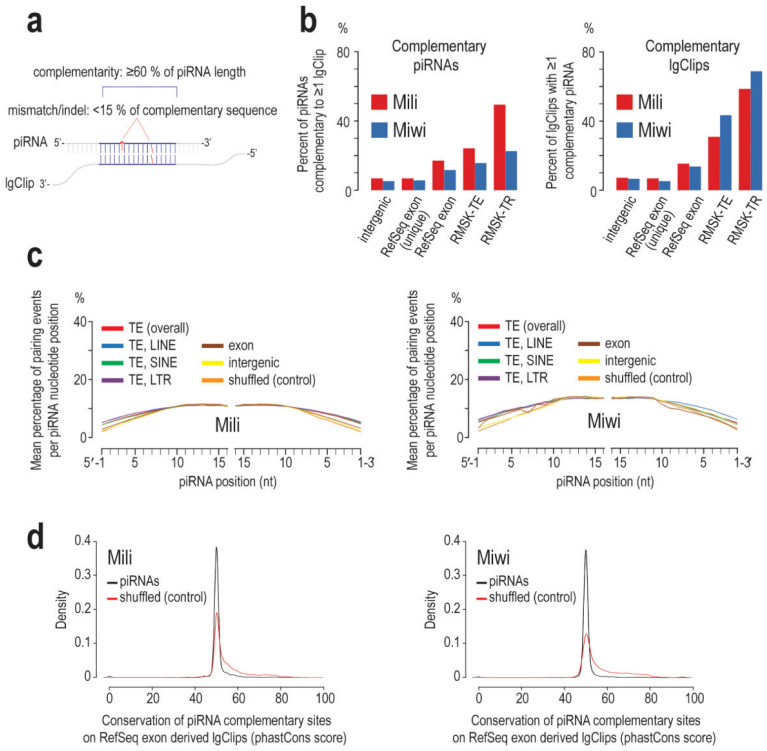

Figure 4. Non-repeat piRNAs lack complementary RNA targets

(a) BLASTN was used to examine complementarity potential for all piRNAs versus all lgClips for each protein. Complementary hits between piRNAs and lgClips were reported as identity ≥85% and a match-ratio of ≥60% for the piRNA sequence.

(b) Left panel: Percent of piRNAs complementary to one or more lgClips. Less than 10% of unique RefSeq exon or intergenic piRNAs match to a complementary lgClip. Right panel: Percent of lgClips with one or more complementary piRNAs. Less than 10% of unique RefSeq exon or intergenic lgClips can be targeted by a complementary piRNA.

(c) The mean percentage of pairing events per nucleotide position, divided by the number of events for each piRNA is plotted, for Mili (left panel), and Miwi (right panel). All piRNA classes show a random pattern of complementarity, with no preference for pairing the 5′ or 3′ ends, almost identical to the control (shuffled piRNA sequences, yellow line).

(d) Conservation of piRNA complementary sites within RefSeq exon derived lgClips (black line). As a control, the complementary sites of shuffled piRNA sequences are analyzed (red line); no differences are observed.

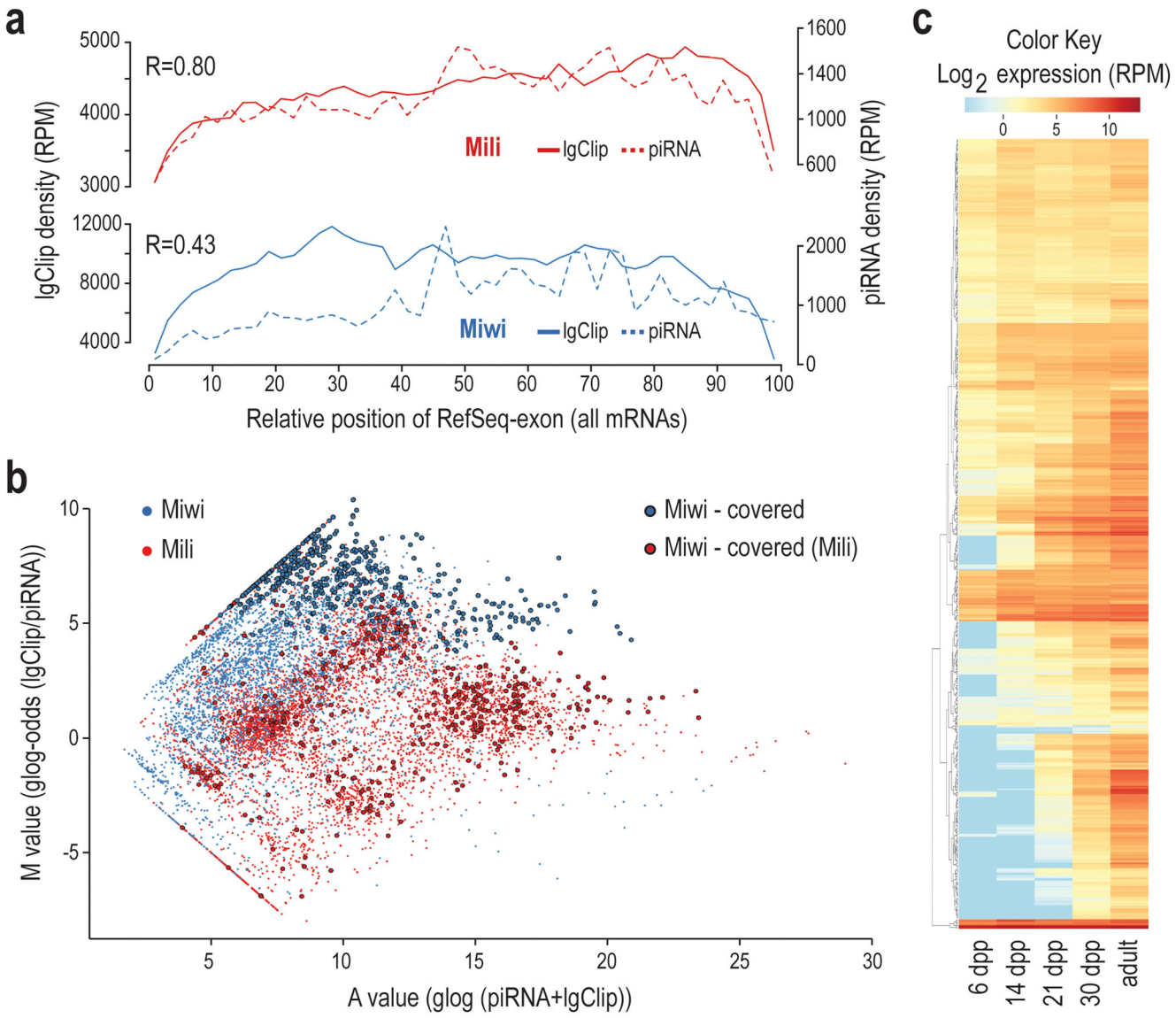

Figure 5. Absence of piRNA processing in Miwi-bound mRNAs essential for spermiogenesis

(a) Density plots of piRNAs (dashed lines) and lgClip (solid lines) relative positions on RefSeq mRNAs, for Mili (red) and Miwi (blue). In contrast to Miwi, Mili piRNA and lgClip densities are highly correlated.

(b) MA plot of generalized-log-odds (glog-odds) for lgClips over piRNA abundance (M value), versus total CLIP-tag glog abundance (lgClip + piRNA) (A value) for all RefSeq exons. mRNAs that have more lgClips than piRNAs tend to appear on the top of the graph. RefSeq mRNAs that had a Miwi lgClip/piRNA ratio more than 8-fold (glog-odds ≥ 3) across all three Miwi replicates ($p<0.05$, t-test) were marked as covered (outlined circles). As a result, we identified 575 Miwi-covered mRNAs representing 460 unique RefSeq genes that have significantly more Miwi lgClips than piRNAs (see Supplementary Table 5).

(c) Heatmap of log2 abundances of Miwi-covered mRNAs in five time points in testis development. Most of Miwi-covered mRNAs are highly enriched or are only expressed in the post meiotic stages of spermatogenesis (21 dpp and after).

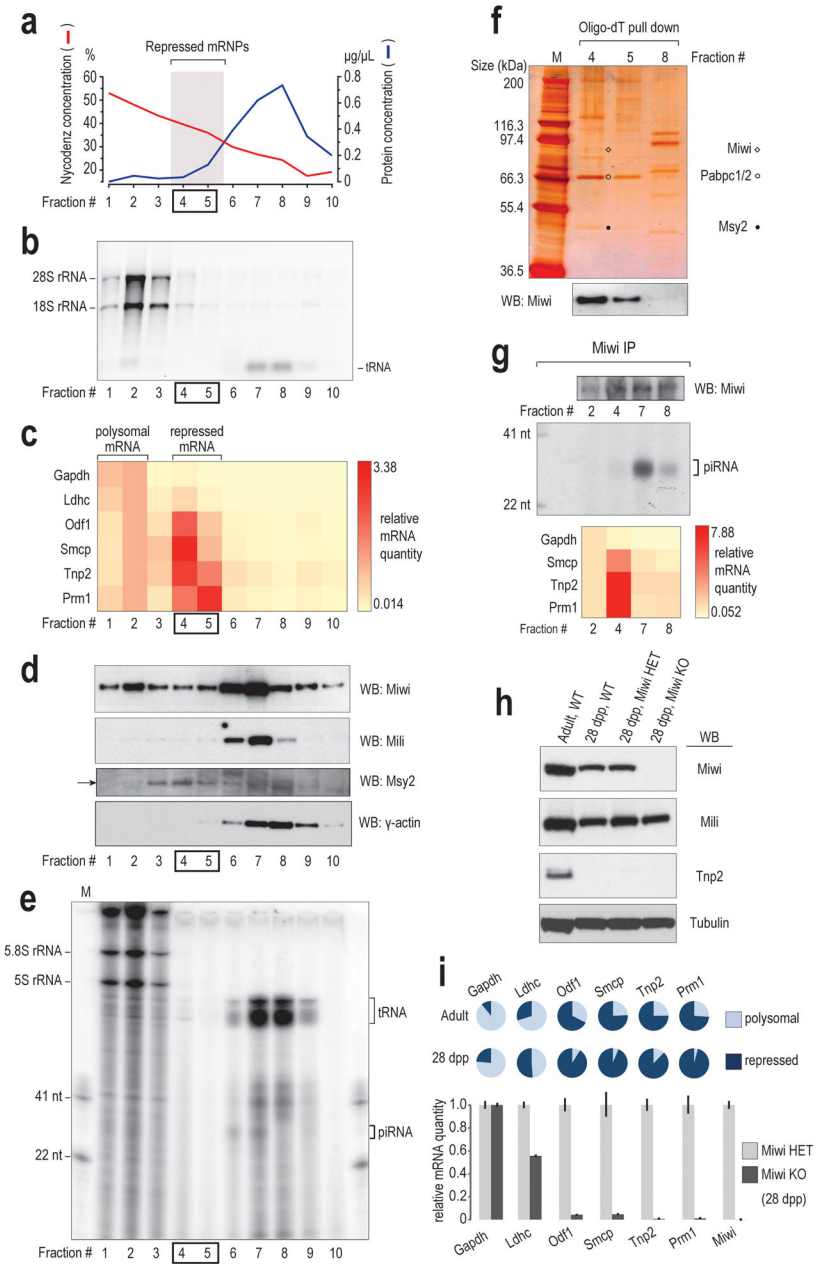

Figure 6. Miwi, devoid of piRNAs, binds and stabilizes spermiogenic mRNAs in repressed mRNPs

(a) Isopyknic density gradients of post-nuclear testis lysate.

(b) Total RNA content.

(c) Heat map of mRNA levels determined by qRT-PCR, normalized to spiked luciferase RNA. Repressed spermiogenic mRNAs are predominantly in fractions 4 and 5 (bracketed).

(d) Western blots.

(e) Small RNA content analyzed by 5′-end labeling and UREA-PAGE. piRNAs are in fractions 6 and 7.

(f) Silver stain of Oligo-dT pull downs and Western for Miwi. Proteins were identified by mass spectrometry from excised bands and from entire pulldown eluates (see Supplementary Table 7).

(g) Miwi protein (top panel), piRNAs (middle) and mRNAs (bottom) from Miwi immunoprecipitates. Spermiogenic mRNAs and piRNAs are bound to Miwi in a non-overlapping manner.

(h) Western blots from adult (wild type, WT) and 28 dpp (WT, Miwi HET and KO) testis lysates. Tnp2 (expressed in elongating spermatids [55,56]), is not detected in 28 dpp testis, verifying the absence of elongating spermatids at that age.

(i) qRT-PCR from 28dpp Miwi HET and KO. Gapdh was used as endogenous control (average Ct, HET=14.94; KO=14.65), and the Miwi HET 28dpp sample as reference. Upper panel: distribution of six mRNAs from adult and 28dpp WT mice across the polysomal (fractions 1 and 2) and repressed mRNPs (fractions 4 and 5; see Supplementary Fig. 7). Bottom panel: Relative levels of the same mRNAs from total RNA isolated from 28 dpp Miwi HET and KO testes. Spermiogenic mRNAs that are highly enriched in fractions 4 and 5 (Odf1, Smcp, Prm1, Tnp2), show a dramatic decrease in the absence of Miwi. Ldhc, which is equally distributed between polysomes and repressed mRNPs at 28dpp, shows a smaller reduction in Miwi KO.

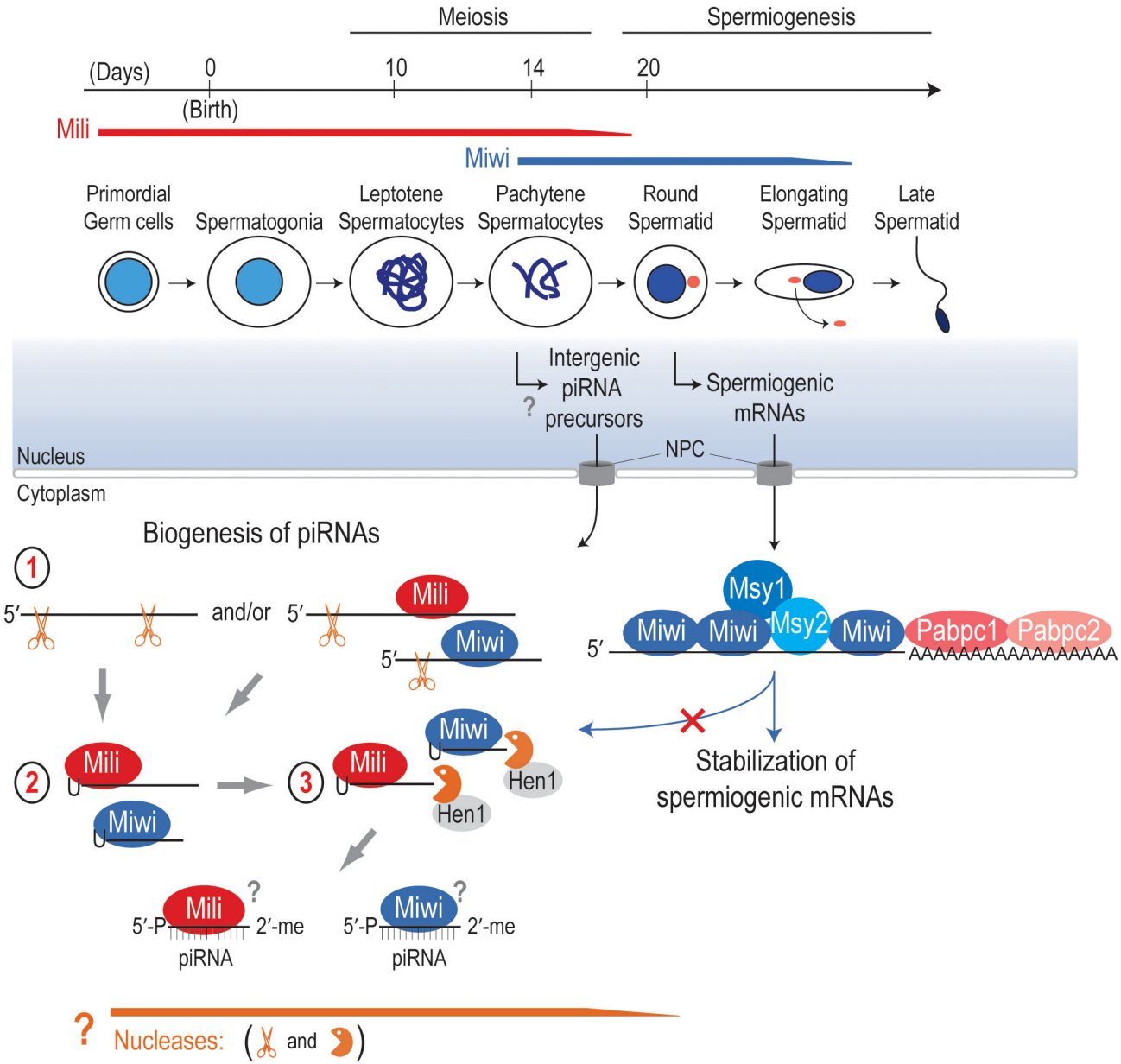

Figure 7. Model for primary piRNA biogenesis, and functions of Mili and Miwi piRNPs during mouse male germ cell differentiation

Two distinct biological functions of Piwi proteins are outlined in this model: piRNA biogenesis and stabilization of spermiogenic mRNAs by Miwi. piRNA processing of Piwi bound transcripts occurs in distinct steps (circled numbers). Two nucleolytic activities, an endonuclease (scissors) that generates the 5′ ends, and a 3′-5′ exonuclease (pacman) that generates the 3′-ends are implicated; their identities are unknown. Mili and Miwi bind to long intergenic transcripts that appear during meiosis, before and/or after endonucleolytic cuts of the long precursors generate intermediate fragments. The function of these intergenic transcripts in meiosis is unknown but may include roles in meiotic chromatin remodeling. 3′-ends of piRNAs are methylated by Hen1 and protected from further trimming. The

piRNA-generating nucleases, or associated cofactors, are not expressed, cease to be active or are sequestered in haploid spermatids. Pachytene Piwi/piRNA complexes are thus end products of RNA clearance but they may also have as yet undetermined non-RNA targeting functions (i.e. chromatoid body organization). Another pool of Miwi binds spermiogenic mRNAs throughout their lengths. This binding is an integral part of the cooperative formation of the mRNPs that maintain these messages prior to spermatid elongation. Red dot indicates the chromatoid body in spermatids that is enriched in Miwi/piRNAs, MVH and Tdrds and which is ultimately eliminated from the maturing spermatid as part of the residual body. NPC: Nuclear Pore Complex.