nature structural & molecular biology

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] & Zissimos Mourelatos[1,3]

Germ cells implement elaborate mechanisms to protect their genetic material and to regulate gene expression during differentiation. Piwi proteins bind Piwi-interacting RNAs (piRNAs), small germline RNAs whose biogenesis and functions are still largely elusive. We used high-throughput sequencing after cross-linking and immunoprecipitation (HITS-CLIP) coupled with RNA–sequencing (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 demonstrated that Miwi binds spermiogenic mRNAs directly, without using piRNAs as guides, and independent biochemical analyses of testis mRNA ribonucleoproteins (mRNPs) established that Miwi functions in the formation of mRNP complexes that stabilize mRNAs essential for spermiogenesis.

Germline development and function engages genes that are conserved in all sexually reproducing metazoans including the Piwi family genes[1–3] and genes encoding Tudor domain–containing proteins (Tdrds)[4,5] that are arranged in functional networks at a genetic and biochemical level[5–9]. Piwi proteins and Tdrds colocalize in germ 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' has been described[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′ to 5′ trimming activity for the formation of the 3′ ends of piRNAs has been reported[16], and although a model for primary piRNA biogenesis has been hypothesized[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 uses 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 and spermatid-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, whereas Miwi is expressed from pachytene to haploid round spermatid stage; Mili binds to piRNAs that are ~26–28 nucleotides (nt), whereas Miwi binds to piRNAs that are ~29–31 nt, and both proteins localize to the cytoplasm[15]. Disruption of genes that encode Mili and Miwi2 and of other genes that are implicated in piRNA biogenesis, such as those that encode Mov10l1, Ddx4 (also known as MVH, mouse Vasa homolog), 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 genes that encode Miwi, Tdrd6 and Tdrd7 causes an arrest at or right after the round spermatid stage[2,28,29] without substantial disturbance of piRNA levels or retrotransposon upregulation[28,29]. In the absence of Miwi, spermiogenic mRNAs are destabilized, contributing to the spermiogenic arrest of developing germ cells[2].

We report here RNAs targeted *in vivo* by Mili and Miwi in the adult mouse testis, using HITS-CLIP[30]. We sequenced piRNAs and longer piRNA precursor RNAs bound by both proteins, which allowed us to propose a coherent model for *in vivo* piRNA biogenesis.

[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. [6]These authors contributed equally to this work. Correspondence should be addressed to Z.M. (mourelaz@uphs.upenn.edu).

Received 17 January; accepted 25 June; published online 29 July 2012; doi:10.1038/nsmb.2347

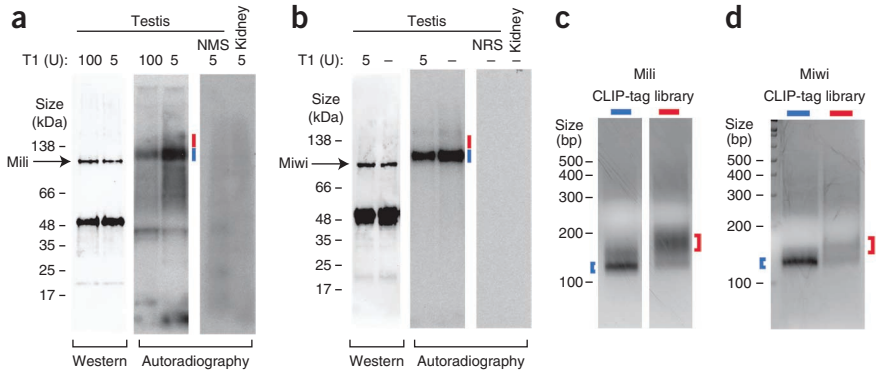

Figure 1 Mili and Miwi HITS-CLIP. (**a**,**b**) CLIPs were performed using highly stringent conditions with buffers containing 2% Empigen. Autoradiograms and western blots of immunoprecipitated, UV-light cross-linked, RNA-protein complexes ligated to radiolabeled 3′ adaptor are shown for Mili (**a**) and Miwi (**b**). Negative control CLIPs were performed using non-immune mouse serum (NMS) or non-immune rabbit serum (NRS) with cross-linked testis samples, and anti-Mili or anti-Miwi antibodies with cross-linked kidney samples. Cells were lysed after cross-linking, and the lysates were incubated in the absence or presence of RNase T1 before immunoprecipitation, as indicated. RNA (CLIP tags) was extracted from the membranes after cutting at indicated areas: blue lines mark the major radioactive signal containing mainly piRNA–protein complexes; red lines mark larger RNA–protein complexes, which appear as smears extending to higher molecular weights. (**c**,**d**) cDNA libraries from Mili (**c**) and Miwi (**d**) CLIP tags were prepared by reverse-transcriptase PCR, gel-purified and sequenced with Illumina GAIIx. Blue and red brackets denote piRNA and large tag–enriched cDNA samples, respectively.

Moreover, using HITS-CLIP and 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. In addition to 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 could be used 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 light can cross-link 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 were enriched in piRNAs, whereas larger complexes extracted from higher-molecular-weight positions were enriched in larger RNAs (**Fig. 1a**,**b** and **Supplementary Table 1**). Immunoblotting for the presence of known interacting proteins (Tdrd1, Tdrd6, MVH, Mili and Miwi) in immunoprecipitates of Miwi and Mili verified that CLIP conditions abolished co-immunoprecipitation of protein partners (data not shown). RNA extracted in three control experiments performed in the same stringent conditions (Mili and Miwi immunoprecipitation from non–UV light–treated testis, **Supplementary Fig. 2b** and testis CLIP using non-immune serum and kidney CLIP using antibodies to Mili and Miwi, **Fig. 1a**,**b**) was nonexistent, and attempts to generate cDNA libraries repeatedly failed. These results demonstrate that our Miwi and Mili CLIP libraries (**Fig. 1c**,**d**) are highly specific.

We sequenced eight libraries, three enriched in piRNAs (two replicates for Mili and one for Miwi) and five enriched in larger RNAs

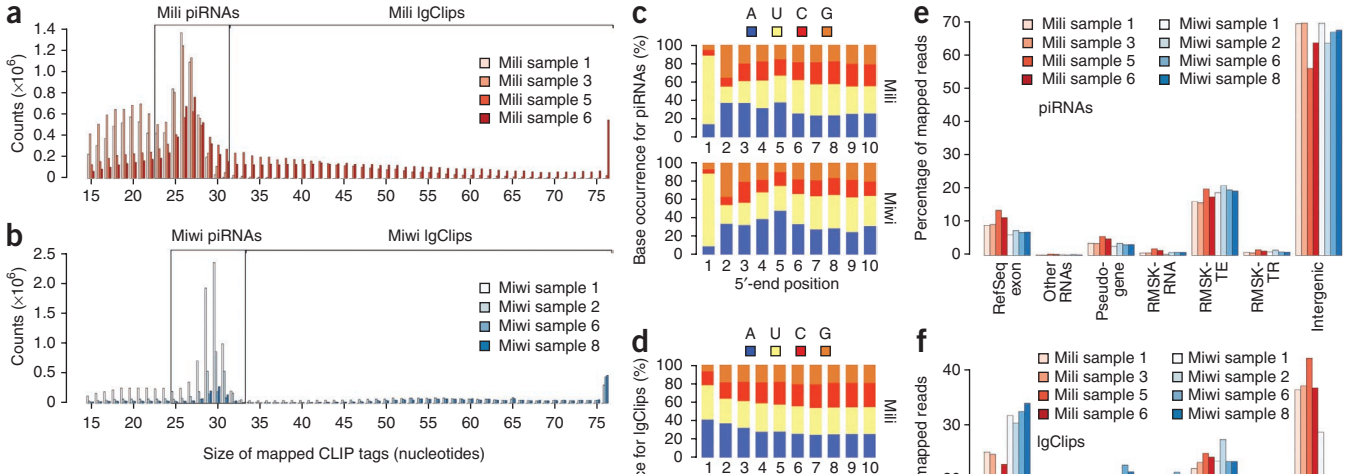

Figure 2 Size distribution, nucleotide preference and genomic classification of mapped CLIP reads. (**a**,**b**) Size distribution of mapped HITS-CLIP reads for Mili (**a**) and Miwi (**b**). (**c**,**d**) Nucleotide preference of the first ten positions for piRNAs (**c**) and IgClips (**d**). Sample numbers refer to identifiers of individual sequenced libraries; the numbers correspond to the order in which they were placed in the Illumina analyzer. (**e**,**f**) Genomic classification of piRNAs (**e**) and IgClips (**f**). RMSK, Repeat Masker; TR, terminal repeats; TE, transposable elements.

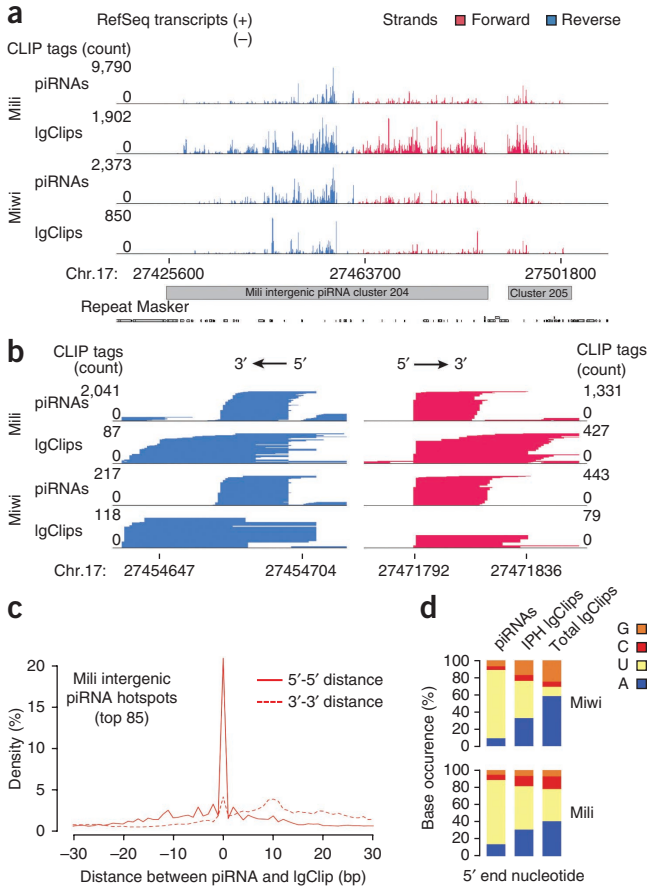

Figure 3 Processing of intergenic piRNA precursors. (**a**) Genome browser snapshot showing mapped CLIP tags on an IPH. Magenta, forward tags; blue, reverse tags. Numbers on the side of tracks denote number of stacked tags. Genomic coordinates (mm9) are marked below the tracks. (**b**) Magnification of the image in **a**, revealing piRNA processing. Note the concurrence of 5′ ends between piRNAs and IgClips (intermediate piRNA precursors) and progressive shortening of IgClip 3′ ends. (**c**) Density plot of piRNA-IgClip distances (piLg-dist) for CLIP tags that map in IPHs. The distance is defined as the genomic position of IgClip 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, percentage of piLg-dist over distances for all pairwise combinations between IgClips and piRNAs is plotted. Only piRNAs within 30 nucleotides upstream or downstream of any given IgClip 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 IgClips originating from IPHs, and total IgClips, for the indicated proteins. IPH IgClips are enriched for 5′ uridine in comparison with total IgClips.

(large CLIP tags) (two replicates for Mili and three for Miwi), 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 nt (peak at 26–27 nt) and 25–33 nt (peak at 29–30 nt), respectively (**Fig. 2a**,**b**). We did not analyze tags smaller than piRNAs because it was impossible to determine whether they were derived from piRNAs or piRNA precursors (see below). piRNAs isolated from both proteins showed a strong bias for uridine at their 5′ position (**Fig. 2c**), whereas larger tags often started with adenosine (**Fig. 2d**). We observed no substantial bias at nucleotide 10 of piRNAs (**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 (**Fig. 2e**), which contain potential bidirectional transcription start sites that have been described previously[11–13]. 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-bound piRNAs, with Mili- and Y12-immunopurified piRNA libraries[6] prepared by immunoprecipitations under standard conditions (**Supplementary Table 2**). The overwhelming majority of the CLIP-identified piRNAs (65–80%) overlapped with standard immunoprecipitation–derived piRNAs. piRNA populations mapping in intergenic piRNA hotspots (IPHs; see below) were almost identical (>94% overlap) (**Supplementary Table 2**). Evidently, the conditions and experimental procedures of CLIP did not alter the piRNA profile of Mili and Miwi, and therefore CLIP can be used to 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** and **Supplementary Fig. 3**). This sorting was essential to understand and accurately describe Piwi protein function. The genomic classification of CLIP tags showed a notable consistency across replicates for both Milli and Miwi (**Fig. 2e**,**f**), indicating the high reliability of our HITS-CLIP approach. The overwhelming majority of nonrepeat piRNAs and lgClips located within RefSeq mRNAs had 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 method for directional RNA-seq (termed solid support directional RNA-seq; **Supplementary Fig. 4a**,**b**) of total RNA purged of ribosomal RNA to provide an unbiased account of all transcript classes. We generated two replicate libraries from adult testis (**Supplementary Fig. 4c**) and one library at each of the following time points in postnatal testis development: 6 d postpartum (d.p.p.) (enriched in spermatogonia), 14 d.p.p. (enriched in pachytene spermatocytes), 21 d.p.p. (enriched in haploid round spermatids) and 30 d.p.p. (enriched in elongating spermatids), yielding 43,543,879 mapped reads in total (**Supplementary Table 3**). Our RNA-seq data can be used to determine whole-transcriptome abundance in the testis during key time points of spermatogenesis (**Supplementary Fig. 4d**).

A model for primary piRNA biogenesis

IPHs lack any defined boundaries; therefore, we used a comprehensive statistical model to determine the genomic coordinates of these elements (**Supplementary Note**). We used the designated IPHs and the corresponding intergenic piRNAs, which constituted 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′ untranslated regions (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 in IPHs were 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 have 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 in IPHs revealed an extremely high peak at 0 only for 5′–5′ but not 3′–3′ distances (**Fig. 3c**). Additionally, the

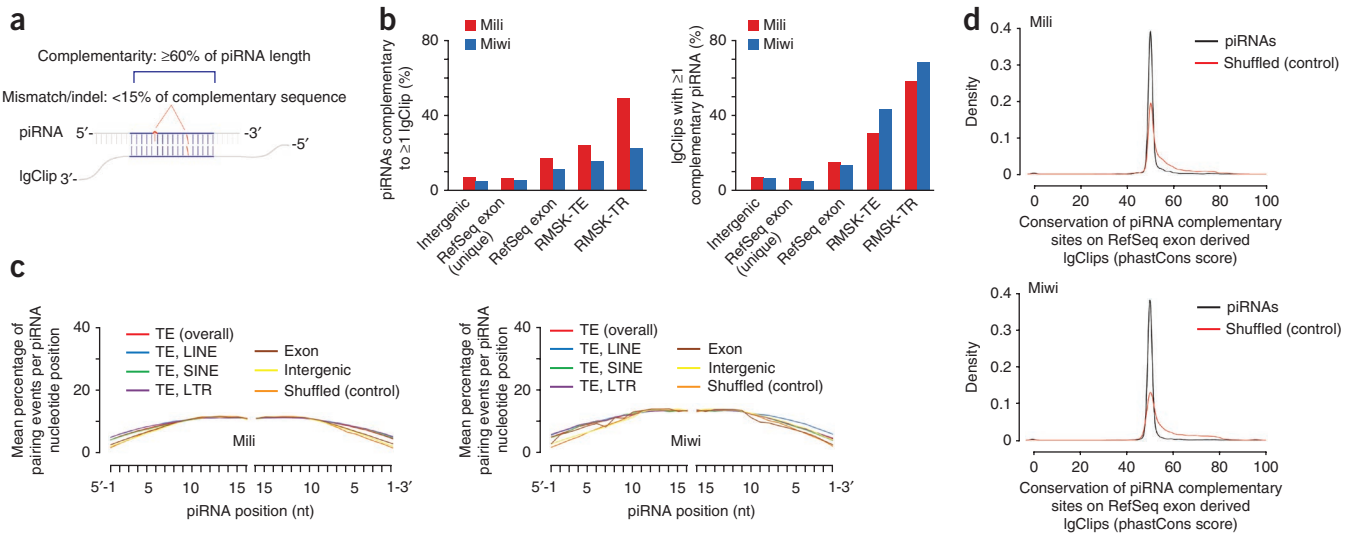

Figure 4 Nonrepeat piRNAs lack complementary RNA targets. (**a**) BLASTN was used to examine complementarity potential for all piRNAs versus all lgClips for each protein. Complementary hits between piRNAs and lgClips were reported as identity ≥85% and a match-ratio of ≥60% for the piRNA sequence. (**b**) Complementarity among lgClips and piRNAs. (**c**) Pairing events per nucleotide position, divided by the number of events for each piRNA, for Mili (left) and Miwi (right). TE, transposable element; LINE, long interspersed element; SINE, short interspersed element; LTR, long terminal repeat. All piRNA classes showed a random pattern of complementarity, with no preference for pairing the 5′ or 3′ ends, almost identical to the control (shuffled piRNA sequences). (**d**) Conservation of piRNA complementary sites within RefSeq exon derived lgClips. As a control, the complementary sites of shuffled piRNA sequences were analyzed.

3′ ends of lgClips had 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 ($P < 2.58 \times 10\text{-}15$, *t*-test) but not for 5′–5′ piLg-dist ($P > 0.069$, *t*-test; **Fig. 3c**). Moreover, Genome Browser inspection and the slope of the 3′–3′ distance curve from negative values toward zero (**Fig. 3b**,**c**) 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 piRNAs) and lgClips revealed 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′ uridine 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 (refs. 32–34), which halts additional 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 uridine 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. Finally, the bulk of IPH piRNAs (~93%) are processed from large clusters (85 clusters with average size of 16,465 nt; **Supplementary Fig. 5e–h** and **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 post-transcriptional 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 cross-linked to miRNPs[30,37], a tripartite piRNP (Piwi–piRNA–complementary target RNA) can be cross-linked 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 exhaustively searched for full or partial complementarity (**Fig. 4a**) between lgClips and piRNAs for both Miwi and Mili using combined lgClip-enriched data sets (**Supplementary Table 1**). Only 5–8% of intergenic and uniquely mapped RefSeq-exon piRNAs had any complementary lgClip sequences, 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 data sets. 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. The complementarity potential was identical to that of the control, and there was 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 were more often paired and exhibited slightly elevated complementarity score compared to the control and other piRNA classes (**Fig. 4b**,**c**), 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 exhibited 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

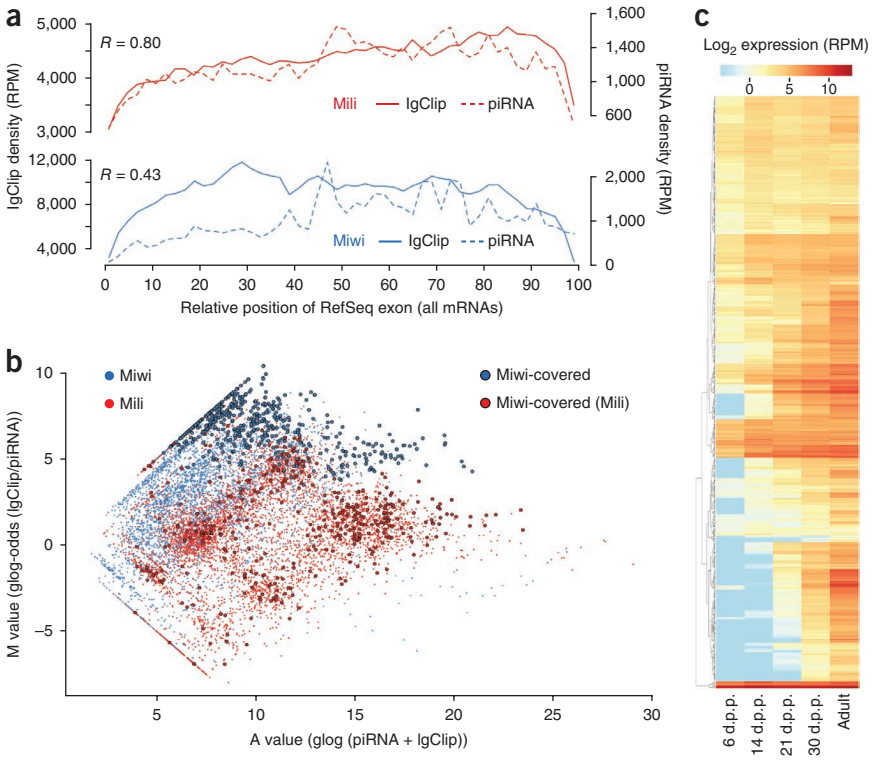

Figure 5 Absence of piRNA processing in Miwi-bound mRNAs essential for spermiogenesis. (**a**) Density plots of piRNAs and IgClip relative positions on RefSeq mRNAs for Mili and Miwi. RPM, reads per million mapped reads. (**b**) MA plot of generalized log-odds (glog-odds) for IgClips over piRNA abundance (M value), versus total CLIP-tag glog abundance (piRNA plus IgClip) (A value) for all RefSeq exons. Outlined circles correspond to mRNAs that had a Miwi (blue) and Mili (red) IgClip/piRNA ratio more than eightfold (glog-odds ≥ 3) across all three replicates ($P < 0.05$, t test). (**c**) Heat map of $\log_2$ abundances of Miwi-covered mRNAs in five time points during testis development. Most of Miwi-covered mRNAs are highly enriched or are only expressed in the post meiotic stages of spermatogenesis (21 d.p.p. and after).

postnatal male germ line, Mili and Miwi proteins do not use piRNAs as guides to target nonrepeat mRNAs.

Miwi targets spermiogenic mRNAs

In addition to RNAs that are processed into piRNAs, we observed specific mRNAs harboring abundant Miwi IgClips 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 positions of piRNAs and IgClips relative to all RefSeq exons using the combined IgClip-enriched libraries for both proteins separately (**Fig. 5a**). The density of Mili piRNAs correlated significantly ($R = 0.8$) with that of IgClips throughout the length of the bound mRNAs, indicating that binding of Mili to its mRNA targets (IgClips) was functionally coupled with their processing into piRNAs (**Fig. 5a**). Conversely, Miwi-bound exonic piRNAs were fewer in number, and their density exhibited a far lower correlation ($R = 0.43$), even with a much higher density of IgClips along RefSeq mRNAs (**Fig. 5a**). These results suggest that Mili binds mRNA targets specifically for piRNA processing, whereas 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 IgClips compared to piRNA abundance (M value), versus total CLIP-tag abundance (IgClip plus piRNA) (A value) for all RefSeq exons (**Fig. 5b**). We focused our analysis to RefSeq mRNAs that have a Miwi IgClip/piRNA ratio more than eight-fold (glog-odds ≥ 3) across all replicates, and we identified 575 Miwi-covered mRNAs representing 460 unique RefSeq genes that had significantly more Miwi IgClips than piRNAs ($P < 0.05$, t test) (**Supplementary Table 5**). Contrary to 3′ UTR binding that is coupled with piRNA processing, there was a rather uniform distribution of Miwi IgClips 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 were highly enriched or were only expressed in the postmeiotic stages of spermatogenesis evidenced in our RNA-seq data sets (**Fig. 5c**). This 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 for Miwi-covered genes (**Supplementary Table 6**). MA plots for IPHs (**Supplementary Fig. 5f,g**) showed similar enrichment of IgClips and piRNAs for both Mili and Miwi (M values of ~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, which revealed mRNAs that are bound but not processed, as described above.

We validated the specificity of the RefSeq mRNA IgClips of Miwi by performing the recently described cross-linking–induced mutation site 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 have 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 importance of this finding is not clear yet (**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. We resolved adult and 28-d.p.p. mouse testis lysates 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 in the density gradient demonstrated the successful separation of various macromolecular complexes[40–42]: actively translated mRNAs (*Gapdh* and *Ldhc*) forming polyribosomes, mRNAs packed in repressed mRNPs (*Prm1*, *Tnp2*, *Smcp* and *Odf1*) and proteins complexed with smaller RNA species or in free state (**Fig. 6a–e**). Miwi was present 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 were mainly found in fractions 6 and 7 (**Fig. 6d,e**). A prolonged exposure of ^{32}P-labeled RNAs still did not identify piRNAs in fractions 4 and 5 (**Supplementary Fig. 7b**). To verify that Miwi present in fractions 4

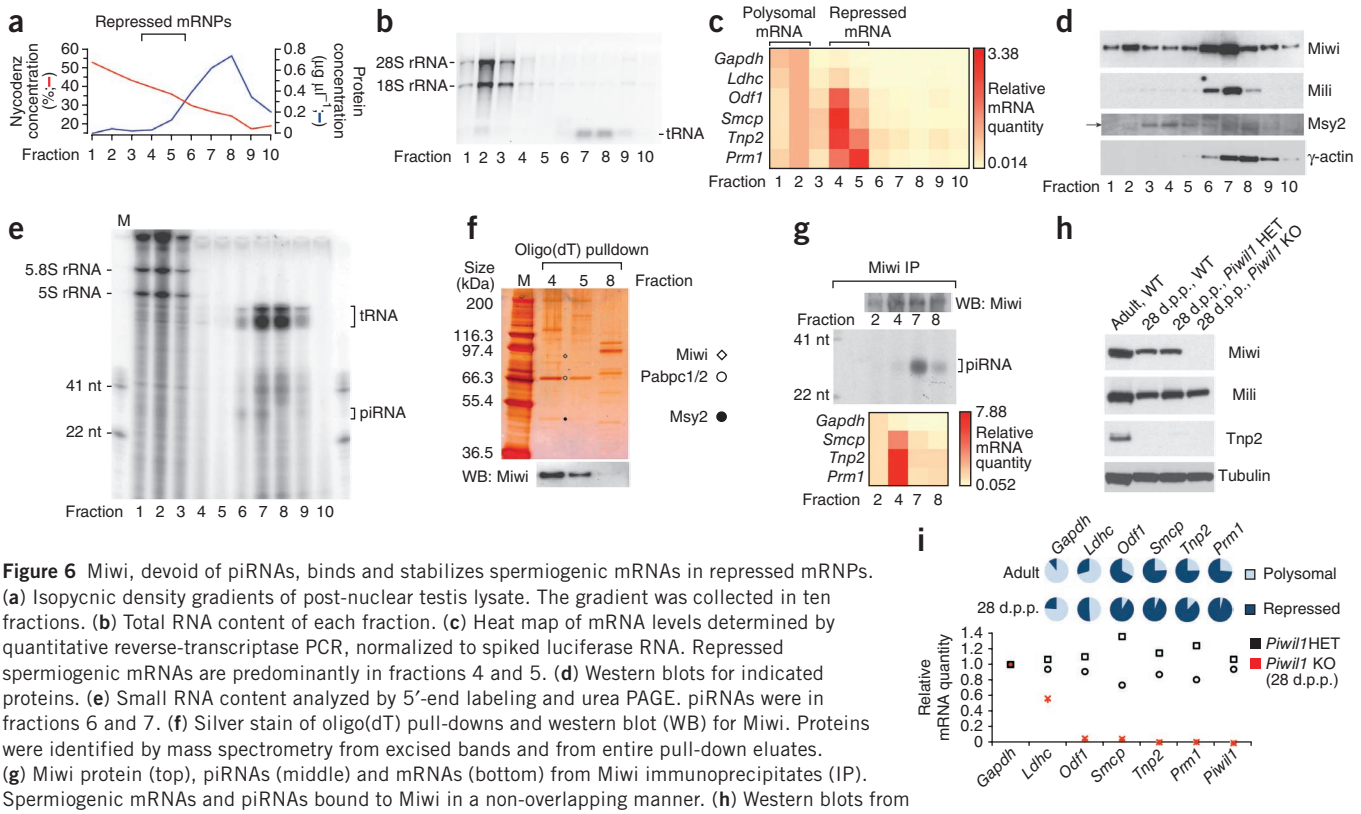

Figure 6 Miwi, devoid of piRNAs, binds and stabilizes spermiogenic mRNAs in repressed mRNPs. (**a**) Isopycnic density gradients of post-nuclear testis lysate. The gradient was collected in ten fractions. (**b**) Total RNA content of each fraction. (**c**) Heat map of mRNA levels determined by quantitative reverse-transcriptase PCR, normalized to spiked luciferase RNA. Repressed spermiogenic mRNAs are predominantly in fractions 4 and 5. (**d**) Western blots for indicated proteins. (**e**) Small RNA content analyzed by 5′-end labeling and urea PAGE. piRNAs were in fractions 6 and 7. (**f**) Silver stain of oligo(dT) pull-downs and western blot (WB) for Miwi. Proteins were identified by mass spectrometry from excised bands and from entire pull-down eluates. (**g**) Miwi protein (top), piRNAs (middle) and mRNAs (bottom) from Miwi immunoprecipitates (IP). Spermiogenic mRNAs and piRNAs bound to Miwi in a non-overlapping manner. (**h**) Western blots from adult wild type (WT) and 28-d.p.p. WT, *Piwil1* heterozygous (HET) and knockout (KO) testis lysates. *Tnp2* (expressed in elongating spermatids[55,56]) was not detected in 28-d.p.p. testis, verifying the absence of elongating spermatids at that age. (**i**) Quantitative reverse-transcriptase PCR ($n = 2$) from 28-d.p.p. *Piwil1* HET and KO. *Gapdh* was used as endogenous control (average threshold cycle (Ct), HET = 14.94; KO = 14.65), and the *Piwil1* HET 28-d.p.p. sample as reference. Distribution of six mRNAs from adult and 28-d.p.p. WT mice across the polysomal (fractions 1 and 2) and repressed mRNPs (fractions 4 and 5) (top). Relative levels of the same mRNAs from total RNA isolated from 28-d.p.p. *Piwil1* HET and KO testes (bottom).

and 5 participates in the formation of the repressive mRNPs, we performed oligo(dT) pull-downs (**Fig. 6f**). We detected Miwi 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 d.p.p., see below) lysates was analyzed by mass spectrometry. Miwi was among the top protein hits identified in all four samples (**Supplementary Table 7**) along with Pabpc1, Pabpc2, Pabpc4, Ybx1, Msy2 and Msy4, 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**). Considerably more mRNA from spermiogenic genes was extracted from Miwi immunoprecipitates using fraction 4 (containing the repressed spermiogenic mRNPs) than with any of the fractions 2, 7 or 8 (fourfold to eightfold) (**Fig. 6c**,**g**). This represents a twofold 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**), supporting the specificity of the above findings. Notably, piRNAs were absent from these repressed mRNPs and spermiogenic mRNAs were absent from fractions containing Miwi–piRNA complexes (**Fig. 6g**). This finding corroborated the results of the bioinformatic analysis and substantiates that piRNAs were 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 carefully characterized postmeiotic spermatids (**Supplementary Note**) and determined that at 28 d.p.p. elongation has not yet ensued (**Fig. 6h** and **Supplementary Fig. 7a**,**d**) and that the number of round spermatids at that stage was similar in mice heterozygous for *Piwil1*, the gene that encodes Miwi, or homozygous null (knockout) for *Piwil1* (**Supplementary Fig. 7d**). We found that the levels of intergenic piRNA precursors and LINE1 retrotransposons in 28-d.p.p. *Piwil1* knockout testis did not change significantly ($P > 0.05$, *t* test) compared to those from *Piwil1* heterozygous mice (**Supplementary Fig. 7e**). We noticed that Mili levels were not reduced in a *Piwil1* knockout (**Fig. 6h**) and that although total levels of piRNAs were similar in *Piwil1* knockout and heterozygous testis, the piRNA size was smaller (~26–28 nt) and corresponded 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 several mRNAs that are bound by Miwi as determined by HITS-CLIP (*Ldhc*, *Odf1*, *Smcp*, *Tnp2* and *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* and *Prm1*) exhibited a dramatic (20-fold to 100-fold) decrease

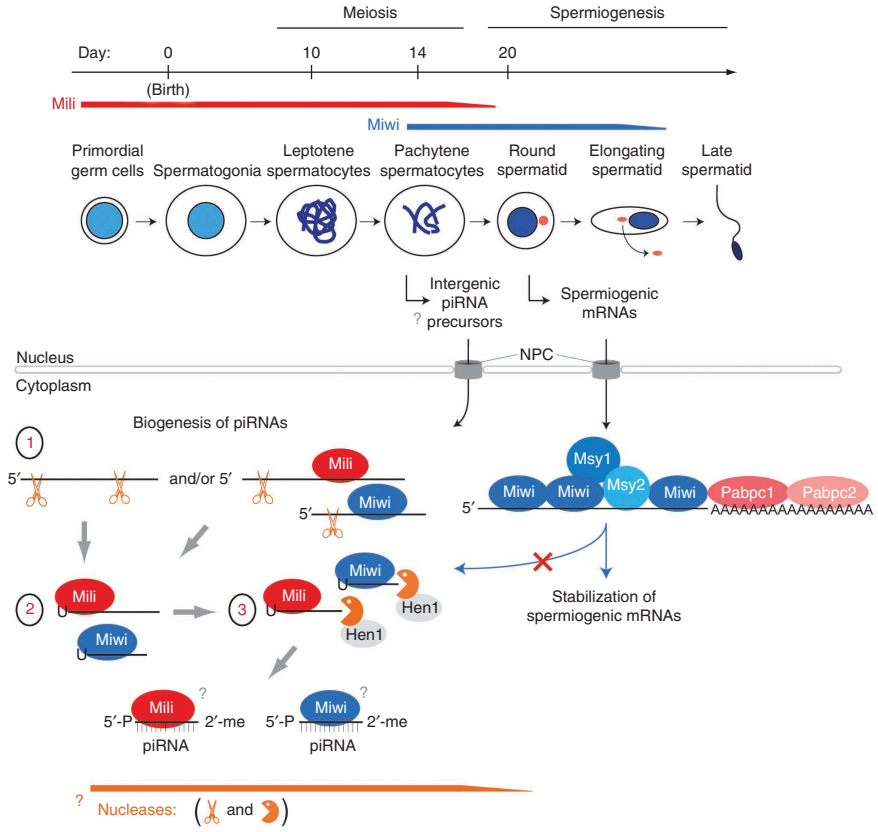

Figure 7 Model for primary piRNA biogenesis, and functions of Mili and Miwi piRNPs during mouse male germ cell differentiation. Two distinct biological functions of Piwi proteins are outlined in this model: piRNA biogenesis and stabilization of spermiogenic mRNAs by Miwi. piRNA processing of Piwi bound transcripts occurs in distinct steps. 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. The 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 (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 before 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.

in abundance in *Piwil1* knockout testis, whereas mRNAs with predominant or exclusive presence in the polysome fractions had a 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 that has been reported previously[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. 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 melanogaster* 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] 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 use a 'tiny seed' sequence for complementary target recognition that we missed in 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 noncoding 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 transcription-mediated chromatin remodeling during meiosis. In either case, the long noncoding 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[28,50]. 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[49,51]. Mutations in genes whose protein products organize the chromatoid body, such as Tdrd6 or Tdrd7, lead to chromatoid body defects and arrest of spermiogenesis[28,29]. 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 to degrade the piRNAs further to recycle nucleotides in sperm. Additional work is required to address these possibilities.

Miwi stabilizes mRNAs in 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 RNases 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 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 toward associating with 'more structured' areas of mRNAs, although the exact role of elements of RNA secondary structure in the formation of Miwi mRNPs warrants additional investigation. Other mechanistic aspects of the loading process, such as the roles of helicases and many 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 throughout 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. Cold-shock domain–containing proteins are deposited on spermiogenic mRNAs co-trancriptionally[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 revealed 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 importance of piRNAs and piRNPs in mammalian germline development.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Gene Expression Omnibus: GSE27622 (HITS-CLIP reads) and GSE27609 (RNA-seq library reads).

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We are grateful to J. Schug for assistance with Illumina analysis approaches in the early phase of this project and to S. Fayngerts, S. Rafail and members of the Mourelatos laboratory for technical help and lively discussions. The antibody to Tnp2 was a gift from W.S. Kistler (University of South Carolina, Columbia). This work was supported by the US National Institute of General Medical Sciences of the National Institutes of Health (R01GM0720777) and in part by Penn Institute for Regenerative Medicine and Perelman School of Medicine, University of Pennsylvania grants to Z.M. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Work in the Gregory lab was supported 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.

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* cross-linking 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.

Published online at http://www.nature.com/doifinder/10.1038/nsmb.2347.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Cox, D.N. *et al.* A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev.* **12**, 3715–3727 (1998).
2. Deng, W. & Lin, H. miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Dev. Cell* **2**, 819–830 (2002).
3. Kuramochi-Miyagawa, S. *et al.* Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. *Development* **131**, 839–849 (2004).
4. Boswell, R.E. & Mahowald, A.P. tudor, a gene required for assembly of the germ plasm in *Drosophila melanogaster*. *Cell* **43**, 97–104 (1985).
5. Chen, C., Nott, T.J., Jin, J. & Pawson, T. Deciphering arginine methylation: Tudor tells the tale. *Nat. Rev. Mol. Cell Biol.* **12**, 629–642 (2011).
6. Kirino, Y. *et al.* Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. *Nat. Cell Biol.* **11**, 652–658 (2009).
7. Nishida, K.M. *et al.* Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in *Drosophila* germlines. *EMBO J.* **28**, 3820–3831 (2009).
8. Vagin, V.V. *et al.* Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. *Genes Dev.* **23**, 1749–1762 (2009).
9. Kirino, Y. *et al.* Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization. *RNA* **16**, 70–78 (2010).
10. Yokota, S. Historical survey on chromatoid body research. *Acta Histochem. Cytochem.* **41**, 65–82 (2008).
11. Lau, N.C. *et al.* Characterization of the piRNA complex from rat testes. *Science* **313**, 363–367 (2006).
12. Aravin, A. *et al.* A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442**, 203–207 (2006).
13. Girard, A., Sachidanandam, R., Hannon, G.J. & Carmell, M.A. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442**, 199–202 (2006).
14. Grivna, S.T., Beyret, E., Wang, Z. & Lin, H. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev.* **20**, 1709–1714 (2006).
15. Siomi, M.C., Sato, K., Pezic, D. & Aravin, A.A. PIWI-interacting small RNAs: the vanguard of genome defence. *Nat. Rev. Mol. Cell Biol.* **12**, 246–258 (2011).
16. Kawaoka, S., Izumi, N., Katsuma, S. & Tomari, Y. 3′ End formation of PIWI-interacting RNAs in vitro. *Mol. Cell* **43**, 1015–1022 (2011).
17. Pillai, R.S. & Chuma, S. piRNAs and their involvement in male germline development in mice. *Dev. Growth Differ.* **54**, 78–92 (2012).
18. Besse, F. & Ephrussi, A. Translational control of localized mRNAs: restricting protein synthesis in space and time. *Nat. Rev. Mol. Cell Biol.* **9**, 971–980 (2008).
19. Heidaran, M.A., Showman, R.M. & Kistler, W.S. A cytochemical study of the transcriptional and translational regulation of nuclear transition protein 1 (TP1), a major chromosomal protein of mammalian spermatids. *J. Cell Biol.* **106**, 1427–1433 (1988).

20. Kleene, K.C. Patterns of translational regulation in the mammalian testis. *Mol. Reprod. Dev.* **43**, 268–281 (1996).
21. Carmell, M.A. *et al.* MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev. Cell* **12**, 503–514 (2007).
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.* **16**, 639–646 (2009).
23. Shoji, M. *et al.* The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. *Dev. Cell* **17**, 775–787 (2009).
24. Kuramochi-Miyagawa, S. *et al.* MVH in piRNA processing and gene silencing of retrotransposons. *Genes Dev.* **24**, 887–892 (2010).
25. Frost, R.J. *et al.* MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. *Proc. Natl. Acad. Sci. USA* **107**, 11847–11852 (2010).
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. USA* **107**, 11841–11846 (2010).
27. De Fazio, S. *et al.* The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature* **480**, 259–263 (2011).
28. Tanaka, T. *et al.* Tudor domain containing 7 (Tdrd7) is essential for dynamic ribonucleoprotein (RNP) remodeling of chromatoid bodies during spermatogenesis. *Proc. Natl. Acad. Sci. USA* **108**, 10579–10584 (2011).
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. *Curr. Biol.* **19**, 630–639 (2009).
30. Chi, S.W., Zang, J.B., Mele, A. & Darnell, R.B. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* **460**, 479–486 (2009).
31. Robine, N. *et al.* A broadly conserved pathway generates 3′UTR-directed primary piRNAs. *Curr. Biol.* **19**, 2066–2076 (2009).
32. Gunawardane, L.S. *et al.* A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. *Science* **315**, 1587–1590 (2007).
33. Horwich, M.D. *et al.* The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr. Biol.* **17**, 1265–1272 (2007).
34. Kirino, Y. & Mourelatos, Z. The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs. *RNA* **13**, 1397–1401 (2007).
35. Simon, B. *et al.* Recognition of 2′-O-methylated 3′-end of piRNA by the PAZ domain of a Piwi protein. *Structure* **19**, 172–180 (2011).
36. Tian, Y., Simanshu, D.K., Ma, J.B. & Patel, D.J. Structural basis for piRNA 2′-O-methylated 3′-end recognition by Piwi PAZ (Piwi/Argonaute/Zwille) domains. *Proc. Natl. Acad. Sci. USA* **108**, 903–910 (2011).
37. Kirino, Y. & Mourelatos, Z. Site-specific crosslinking of human microRNPs to RNA targets. *RNA* **14**, 2254–2259 (2008).
38. Reuter, M. *et al.* Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. *Nature* **480**, 264–267 (2011).
39. Zhang, C. & Darnell, R.B. Mapping *in vivo* protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nat. Biotechnol.* **29**, 607–614 (2011).
40. Herbert, T.P. & Hecht, N.B. The mouse Y-box protein, MSY2, is associated with a kinase on non-polysomal mouse testicular mRNAs. *Nucleic Acids Res.* **27**, 1747–1753 (1999).
41. Bagarova, J., Chowdhury, T.A., Kimura, M. & Kleene, K.C. 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* **140**, 853–864 (2010).
42. Kleene, K.C., Bagarova, J., Hawthorne, S.K. & Catado, L.M. Quantitative analysis of mRNA translation in mammalian spermatogenic cells with sucrose and Nycodenz gradients. *Reproductive biology and endocrinology* **8**, 155 (2010).
43. Giorgini, F., Davies, H.G. & Braun, R.E. Translational repression by MSY4 inhibits spermatid differentiation in mice. *Development* **129**, 3669–3679 (2002).
44. Kimura, M., Ishida, K., Kashiwabara, S. & Baba, T. Characterization of two cytoplasmic poly(A)-binding proteins, PABPC1 and PABPC2, in mouse spermatogenic cells. *Biol. Reprod.* **80**, 545–554 (2009).
45. Yang, J., Medvedev, S., Reddi, P.P., Schultz, R.M. & Hecht, N.B. The DNA/RNA-binding protein MSY2 marks specific transcripts for cytoplasmic storage in mouse male germ cells. *Proc. Natl. Acad. Sci. USA* **102**, 1513–1518 (2005).
46. Yang, J. *et al.* Absence of the DNA-/RNA-binding protein MSY2 results in male and female infertility. *Proc. Natl. Acad. Sci. USA* **102**, 5755–5760 (2005).
47. Huang, H. *et al.* piRNA-associated germline nuage formation and spermatogenesis require MitoPLD profusogenic mitochondrial-surface lipid signaling. *Dev. Cell* **20**, 376–387 (2011).
48. Watanabe, T. *et al.* MITOPLD is a mitochondrial protein essential for nuage formation and piRNA biogenesis in the mouse germline. *Dev. Cell* **20**, 364–375 (2011).
49. Hermo, L., Pelletier, R.M., Cyr, D.G. & Smith, C.E. 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.* **73**, 279–319 (2010).
50. Haraguchi, C.M. *et al.* Chromatoid bodies: aggresome-like characteristics and degradation sites for organelles of spermiogenic cells. *J. Histochem. Cytochem.* **53**, 455–465 (2005).
51. Austin, C.R. & Sapsford, C.S. The development of the rat spermatid. *J. R. Microsc. Soc.* **71**, 397–406 (1951).
52. Dostie, J. & Dreyfuss, G. Translation is required to remove Y14 from mRNAs in the cytoplasm. *Curr. Biol.* **12**, 1060–1067 (2002).
53. Faehnle, C.R. & Joshua-Tor, L. Argonaute MID domain takes centre stage. *EMBO Rep.* **11**, 564–565 (2010).
54. Liu, X., Jin, D.Y., McManus, M.T. & Mourelatos, Z. Precursor microRNA-programmed silencing complex assembly pathways in mammals. *Mol. Cell* **46**, 507–517 (2012).
55. Oko, R.J., Jando, V., Wagner, C.L., Kistler, W.S. & Hermo, L.S. Chromatin reorganization in rat spermatids during the disappearance of testis-specific histone, H1t, and the appearance of transition proteins TP1 and TP2. *Biol. Reprod.* **54**, 1141–1157 (1996).
56. Alfonso, P.J. & Kistler, W.S. Immunohistochemical localization of spermatid nuclear transition protein 2 in the testes of rats and mice. *Biol. Reprod.* **48**, 522–529 (1993).

ONLINE METHODS

Antibodies. Affinity-purified, rabbit polyclonal anti-peptide antibodies were generated: Tdrd1, EVGGSKGDRKPPTC and Tdrd6, CETKTSKFYERSTRS (custom antibodies generated by Genscript and identified as Tdrd1-88 and Tdrd6-2). Other antibodies used were anti-Mili (monoclonal antibody clone 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) and anti-Tnp2 (gift of W.S. Kistler, University of South Carolina)[19,56]. For western blots, primary antibody dilutions were 1 μg/ml, or 1 μl of ascites fluid per milliliter.

Histology and electron microscopy. Miwi mice, developed in the Lin laboratory[2], were purchased from MMRRC–University of Missouri (mouse strain 029995). Testes collected from adult mice and from 21, 22, 23, 24, 25, 26, 27, 28 d.p.p. mice (± 18 h) from *Piwil1* heterozygotes and homozygotes were fixed in Bouin's, embedded in paraffin and stained with hematoxylin and eosin. All housing, breeding and procedures were performed according to the US National Institutes of Health Guide for the Care and Use of Experimental Animals and approved by the University of Pennsylvania Institutional Animal Care and Use Committee. Electron Microscopy of testes from *Piwil1* heterozygous and *Piwil1* knockout mice was performed at the Electron Microscopy Resource Laboratory at PENN.

Preparation of recombinant Miwi and Mili, and *in vitro* cross-linking 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 cross-linking experiments between recombinant Piwi proteins and piRNA, we followed a similar strategy to that previously described[37]. Cross-linking 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 cross-linking 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 cross-linking.

Mili and Miwi HITS-CLIP. Adapted from the Argonaute HITS-CLIP[30] with stated modifications. Testes from C57BL/6J mice were collected, 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 (3×) at 254 nm (400 mJ/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 the Mili antibody, 8 μl rabbit anti-mouse IgG was incubated with the beads, in 350 μl antibody (Ab) binding buffer for 45 min at room temperature; beads were then incubated in 350 μl Ab binding buffer (0.1 M Na-phosphate pH 8 and 0.1% NP-40) with 10 μl of 17.8 ascites fluid at 4 °C for ~4 h. For Miwi antibody, 15 μg in 350 μl Ab binding buffer were added to the beads at 4 °C for ~2–3 h; Ab-bound beads were washed with 1× PXL (1× PBS (no Mg^{2+} and no Ca^{2+}), 0.1% SDS, 0.5% deoxycholate and 0.5% NP-40) or 1× PMPG (1× PBS (no Mg^{2+} and no Ca^{2+}) and 2% Empigen). Mili CLIP was performed with 1× PXL (1× PMPG is also suitable), whereas Miwi CLIP was performed with 1× PMPG.

For each IP, UV light–treated cells from two testis were lysed in 300 μl of 1× PXL or 1× PMPG with protease inhibitors and RNasin (2 u/μl); lysates were treated with DNase (Promega) for 10 min at 37 °C. For optimal retrieval of long cross-linked RNAs, lysates for Mili CLIP were treated with 5 U 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 immunoprecipitation beads were performed with 1× and 5× PXL or 1× and 5× PMPG (5× is 5× PBS, detergents same as 1×), 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′ adaptor labeling and ligation to CIP-treated RNA CLIP tags.

Immunoprecipitation beads were eluted at 70 °C for 12 min using 30 μl of 2× Novex reducing loading buffer. Samples were analyzed by NuPAGE. Cross-linked RNA–protein complexes were transferred on nitrocellulose (Invitrogen LC2001), and the membrane was exposed to film for 1–2 h. The main radioactive signal of the immunoprecipitated cross-linked 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].

Reverse-transcriptase (RT)-PCR and re-amplification were performed as described previously[30]. Reaction products from both reactions were analyzed on 3% Metaphor 1×TAE gels stained with EtBr. piRNAs form a distinct band at ~70 bp and ~120 bp after first and second PCR, respectively, whereas cDNA from larger RNAs formed 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. Two micrograms 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 10× 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. Fifty picomoles of RL3-3′ biotin adaptor (IDT) were ligated to the fragmented RiboMinus-RNA using T4 RNA ligase in 80 μl of reaction mixture at 16 °C overnight.

One hundred microliters of M280 Streptavidin Dynabeads (Invitrogen 112.05D) slurry were used. The ligation mixture was incubated with the beads for 15 min at RT in 0.5× BW buffer (1×: 10 mM Tris pH 7.5, 2 M 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 pmol in reaction) ligation mixture (80 μl) at 16 °C for 6 h, then washed twice with BW buffer and twice with H_2O. For reverse transcription, the beads were split in two aliquots, and each aliquot was mixed with Titan One Tube RT-PCR System (11888382001 Roche); reverse transcription was performed at 50 °C for 45 min, and at 55 °C for 15 min. Two PCR amplification steps and agarose gel analysis (with care to exclude adaptor 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 1× KMH (200 mM KCl, 2 mM $MgCl_2$, 20 mM HEPES pH 7.4, 0.5% NP-40, 0.1 U/μl rRNAsin and protease inhibitors) was prepared as a step gradient by overlaying five equal parts of Nycodenz solutions and was left to diffuse overnight at 4 °C. We laid 0.2 ml of post nuclear testis lysate in 1× KMH over the gradient and centrifuged it at 150,000*g* for 20 h. The gradient was collected in ten 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. One hundred microliters of Dynabeads $Oligo(dT)_{25}$ suspension (Invitrogen 610-11) were washed twice with 1× KMH, and were mixed with 100 μl of Nycodenz fraction, diluted with equal volume of 1× KMH. The suspension was rotated gently for 20 min at 4 °C, and was then washed three times with 1× KMH buffer. Two thirds of the beads were eluted with 2× SDS reducing loading buffer (Invitrogen NP0007), and the rest were used for RNA extraction with Trizol.

Quantitative RT-PCR. Equal volume of RNA extracted from each fraction was reverse transcribed by Supersript III (Invitrogen 18080-051) in the

doi:10.1038/nsmb.2347

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′-CTGCCGTCTAC TCCTCTTGG; IPH #494 F: 5′-TGTTCACTTACATATCAGGGTC, R: 5′-GTAAAGCCCAAGAGCAAGAG; IPH #204 F: 5′-AGTCTGTGTAG TAGTTTCCTGAG, 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 d.p.p. mice were analyzed by mass spectrometry at Taplin Mass Spectrometry Facility, Harvard University 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. Details of bioinformatic methods are available in the **Supplementary Note**.