doi:10.1038/nature04916

LETTERS

A novel class of small RNAs bind to MILI protein in mouse testes

Alexei Aravin[1]*†, Dimos Gaidatzis[2]*, Sébastien Pfeffer[1]†, Mariana Lagos-Quintana[1], Pablo Landgraf[1], Nicola Iovino[1], Patricia Morris[3], Michael J. Brownstein[4], Satomi Kuramochi-Miyagawa[5], Toru Nakano[5], Minchen Chien[6], James J. Russo[6], Jingyue Ju[6,7], Robert Sheridan[8], Chris Sander[8], Mihaela Zavolan[2] & Thomas Tuschl[1]

Small RNAs bound to Argonaute proteins recognize partially or fully complementary nucleic acid targets in diverse gene-silencing processes[1–4]. A subgroup of the Argonaute proteins—known as the 'Piwi family'[5]—is required for germ- and stem-cell development in invertebrates[6,7], and two Piwi members—MILI and MIWI—are essential for spermatogenesis in mouse[8,9]. **Here we describe a new class of small RNAs that bind to MILI in mouse male germ cells, where they accumulate at the onset of meiosis. The sequences of the over 1,000 identified unique molecules share a strong preference for a 5′ uridine, but otherwise cannot be readily classified into sequence families. Genomic mapping of these small RNAs reveals a limited number of clusters, suggesting that these RNAs are processed from long primary transcripts. The small RNAs are 26–31 nucleotides (nt) in length—clearly distinct from the 21–23 nt of microRNAs (miRNAs) or short interfering RNAs (siRNAs)—and we refer to them as 'Piwi-interacting RNAs' or piRNAs. Orthologous human chromosomal regions also give rise to small RNAs with the characteristics of piRNAs, but the cloned sequences are distinct. The identification of this new class of small RNAs provides an important starting point to determine the molecular function of Piwi proteins in mammalian spermatogenesis.**

We immunoprecipitated MILI-containing ribonucleoprotein complexes from testis lysate of adult mice and purified the associated RNAs. Although 5′ ^{32}P-labelling of the isolated molecules revealed a distinct population of RNAs 26–28 nt in length, total testis RNA from fertile adults—but not 10-day-old mice—showed a prominent RNA species of approximately 30 nt in length (Fig. 1a). Labelling of the RNA remaining in the supernatant of the MILI immunoprecipitation showed that the 30-nt fraction was intact, and that the 26–28-nt MILI-interacting RNAs were unlikely to represent degradation products of the 30-nt-long RNAs. To determine the identity of the different small-RNA size populations, we cloned and sequenced the MILI-interacting small RNAs, as well as small RNAs from testis ranging in size between 18–26 nt and 24–33 nt.

The over 15,000 sequences identified in three small-RNA libraries suggested three distinct size populations of small RNAs. The 18–26-nt-fractionated library showed a peak at 21 nt and 22 nt, corresponding to miRNAs, and also revealed a 26-nt shoulder (Fig. 1b). The 24–33-nt fraction showed a bimodal distribution with a strong peak at 29–31 nt, corresponding to the small RNAs detected by 5′-labelling of total testis RNA, and a small peak at 26–28 nt. MILI-interacting small RNAs demonstrated a unimodal length-distribution, with a peak at 26–28 nt. The 27-nt shoulders in the size distribution profiles of the smaller and larger size fractions thus probably represent the MILI-interacting subpopulation. Whereas ~60% of small RNA clones from the 18–26-nt library can be annotated as miRNAs and degradation products of abundant cellular RNAs, more than 80% of the sequences in the MILI immunoprecipitate and the 24–33-nt libraries did not derive from known transcripts or genomic repeats (Supplementary Table 1).

Sequence analysis indicated that 85% of the MILI-interacting RNAs and 88% of the 24–33-nt RNAs contain a 5′ uridine residue (Supplementary Table 1, Supplementary Fig. 1). The bias for 5′ uridine is one of the characteristics of miRNAs and repeat-associated siRNAs (rasiRNAs) produced from double-stranded RNA (dsRNA) precursors by RNase III enzymes[10–12]. The majority of mouse small RNA sequences (92.9% and 97.7%, respectively) obtained from libraries of 24–33-nt testis and MILI immunoprecipitation were cloned only once. Genomic mapping of cloned sequences showed that less than 15% of clones in both libraries are derived from annotated repetitive regions (Supplementary Tables 1 and 2). Furthermore, less than 3% of the sequences match the genome more than ten times. The relative frequency of sequences matching different types of repeats roughly corresponds to the relative proportion of those repeat elements in the genome.

Clustering the genomic loci corresponding to the more than 1,500 MILI-interacting sequences, on the basis of a maximum distance of 15 kilobases (kb) between two consecutive loci, indicated that 81% of clones are derived from only 42 genomic regions (Fig. 2a; Supplementary Tables 3 and 4). Moreover, 19% of the sequences in the 18–26-nt library also originate in these regions, indicating that the 26-nt shoulder in the size distribution profile of this library indeed corresponds to MILI-interacting RNAs. Even more surprisingly, we found that 76% of all sequences identified in the 24–33-nt library map to the same regions, which thus seem to produce both the 26–28-nt MILI-interacting RNAs and the abundant 29–31-nt RNAs that are present in total testis RNA. On the basis of the interaction of these small RNAs with the MILI member of the Piwi protein family, we refer to them as 'Piwi-interacting RNAs' (piRNAs).

[1]Howard Hughes Medical Institute, Laboratory of RNA Molecular Biology, The Rockefeller University, 1230 York Avenue, Box 186, New York, New York 10021, USA. [2]Biozentrum, Universität Basel, Klingelbergstr 50-70, CH-4056 Basel, Switzerland. [3]Population Council, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA. [4]J. Craig Venter Institute, Functional Genomics, 9704 Medical Center Drive, Rockville, Maryland 20850, USA. [5]Department of Pathology, Medical School, Graduate School of Frontier Biosciences, Osaka University, Yamada-oka 2-2 Suita, Osaka 565-0871, Japan. [6]Columbia Genome Center, Russ Berrie Pavilion, 1150 St. Nicholas Avenue, New York, New York 10032, USA. [7]Department of Chemical Engineering, Columbia University, 500 West 120 Street, New York, New York 10027, USA. [8]Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA. †Present addresses: Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA (A.A.); CNRS-UPR 2357, IBMP, 12 rue du Général Zimmer, 67084 Strasbourg Cedex, France (S.P.).
*These authors contributed equally to this work.

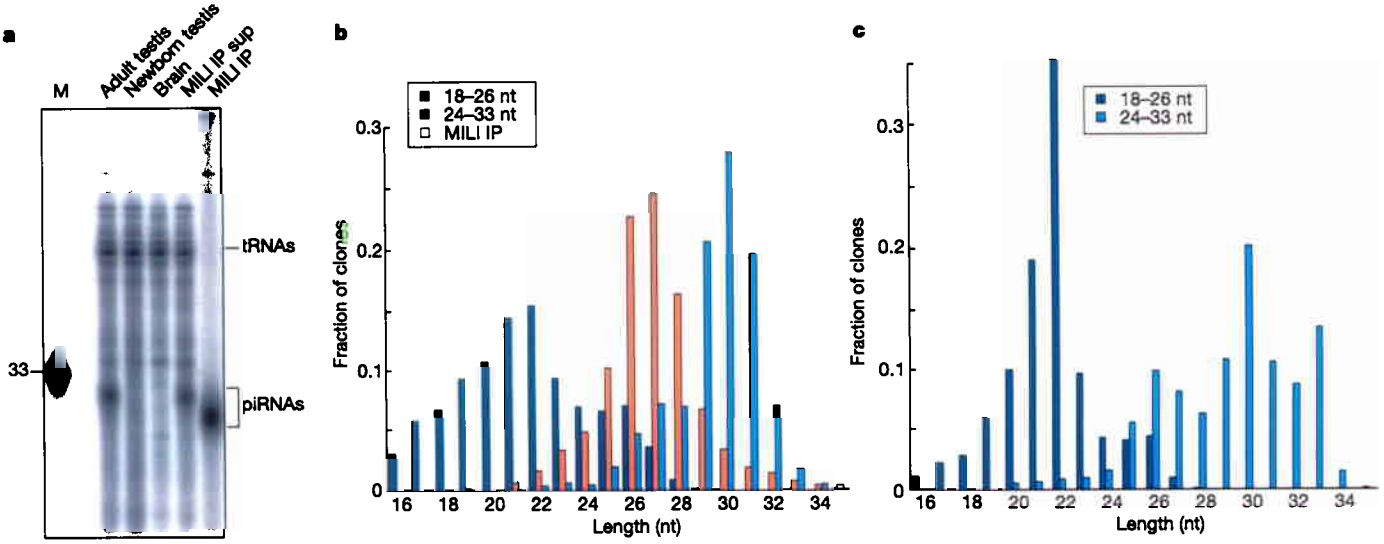

Figure 1 | MILI interacts with 26–28-nt RNAs. **a**, Radioactive labelling of total RNA isolated from adult mouse testis or 10-day-old newborn mouse testis, and adult mouse brain. Adult testis shows an abundant 29–31-nt small-RNA fraction. MILI-immunoprecipitated (IP) RNAs are predominantly 26–28 nt long, whereas 29–31-nt-long RNAs remain in the supernatant (sup). The size and mobility of the oligoribonucleotide marker (M) is indicated on the left; the mobility of transfer RNAs (tRNAs) and 'Piwi- interacting RNAs' (piRNAs) is indicated on the right. **b**, **c**, The size distribution (in nucleotides (nt)) of small RNAs cloned from adult mouse (**b**) and human (**c**) testis total RNA in the 18–26-nt (dark blue bars), 24–33-nt (light blue bars), and MILI-immunoprecipitated (orange bars) RNA fractions.

The pronounced clustering strongly suggests that multiple piRNAs are processed from long primary transcripts—a hypothesis supported by testis-specific expressed sequence tags (ESTs) and messenger RNAs mapping to these regions. The identified piRNA-encoding regions are distributed over most mouse chromosomes and they range in size from 0.9 to 127 kb (Fig. 2a; Supplementary Tables 3 and 4). Although piRNAs map exclusively to one chromosomal strand in many regions, some regions—including the region on mouse chromosome 17 that is one of the major sources of piRNAs—encode piRNAs in both orientations. Notably, in all cases

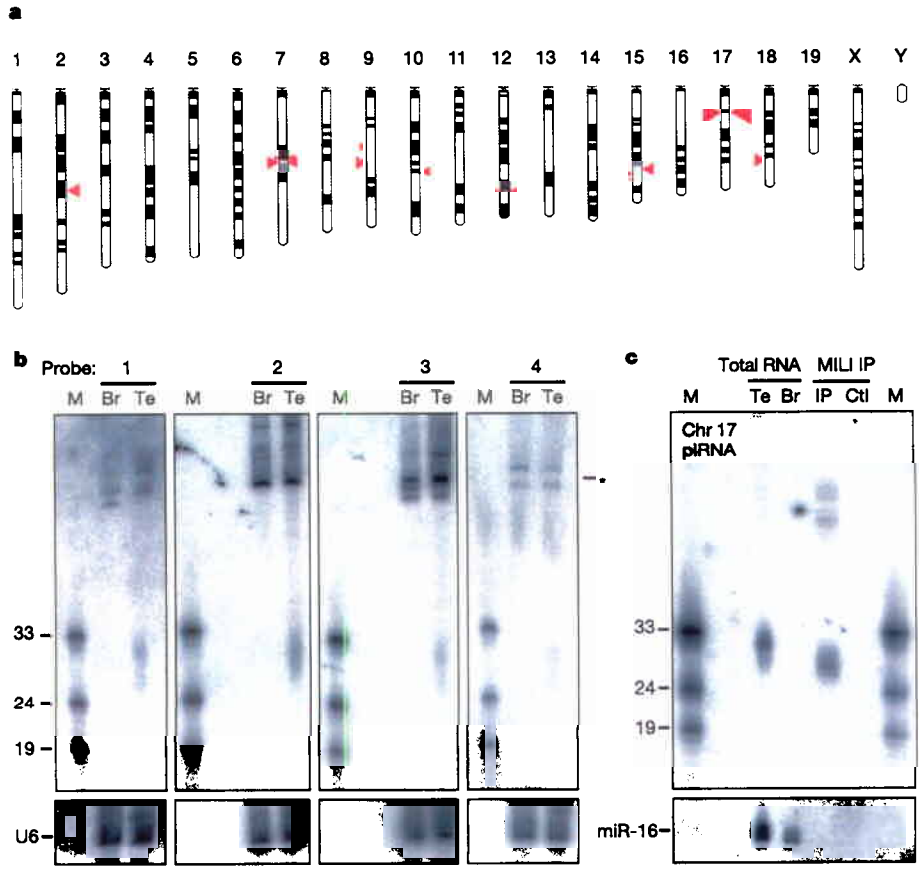

Figure 2 | MILI-interacting piRNAs are encoded in clustered genomic loci. **a**, Location of piRNA-encoding regions on mouse chromosomes. The size of the triangles is proportional to the number of cloned MILI-interacting sequences. The left or right position of the triangles indicates mapping of the clones to the minus or plus strand, respectively. **b**, **c**, Northern blot hybridization for mouse chromosome (Chr) 17 piRNAs using partially hydrolysed RNA probes against four distinct 500-nt regions (**b**) or a 27-nt oligodeoxynucleotide probe complementary to a single piRNA clone (**c**) (top panels). Mouse total RNA from adult testis (Te) or brain (Br) was examined, and the size and mobility of the RNA size marker (M) is specified. The asterisk marks tRNA cross-hybridization signals. Blots were re-probed with a U6 antisense probe (bottom panels). MILI-immunoprecipitated RNAs (IP) are absent in control (Ctl) immunoprecipitation experiments; miR-16 is detectable in total RNA from testis or brain, but is undetectable in the MILI-immunoprecipitated sample.

this arrangement was caused by two closely spaced clusters: one that almost exclusively contained sequences mapped to the sense strand and the other containing sequences mapped to the antisense strand, suggesting bidirectional divergent transcription from a central promoter (Supplementary Table 5; Supplementary Fig. 2). Thus, although piRNAs resemble in size the rasiRNAs of fruitfly and zebrafish[10,12], they differ from rasiRNAs in several respects. First, piRNAs map to the genome in a highly strand-specific manner, in contrast to rasiRNAs that map to repeat regions in sense or antisense orientation as if randomly generated from long dsRNA precursors[10,12]. Second, piRNAs predominantly map to single genomic loci, whereas rasiRNAs map by definition to repetitive sites including transposable elements. In fact, the proportion of repeat elements is smaller within the piRNA regions than in the 100-kb flanking regions (29% versus 38%, P-value $< 2.2 \times 10^{-16}$).

The spacing between cloned piRNAs within each genomic region has no apparent pattern. To assess whether the piRNAs show any evidence of specific processing, we aligned the piRNAs whose loci are partially overlapping (52% of the clones in the piRNA regions) and evaluated the precision with which their 5′ and 3′ ends have been processed. We found frequent examples of piRNAs whose 5′ or 3′ ends coincide (Supplementary Tables 6 and 7), indicating that they are not random degradation products of long transcripts (Supplementary Fig. 3). Additionally, the 5′ end seems to be more precisely processed than the 3′ end, similar to what has been observed among miRNA clones[11]. To examine if piRNAs may be—like miRNAs—excised from dsRNA fold-back precursor structures, we investigated if any base-paired regions emerged in approximately 100 nt from each side of the piRNA. Although our computational approach was able to reveal the loop and stem regions of miRNA precursors, no clear hybridization pattern involving the piRNAs or the sequences flanking them was found (Supplementary Fig. 4). It is possible that long-range dsRNA structures or sequence-specific protein machinery are involved in guiding the maturation process. The strong preference for a 5′ uridine in piRNAs suggests the involvement of Drosha or Dicer RNase III, but additional structural or sequence-specific determinants are yet to be identified.

We further examined the processing of piRNA primary transcripts by northern blotting of total testis RNA using probes specific to four 500-nt regions within the largest piRNA cluster on mouse chromosome 17. All four probes antisense to cloned piRNAs detected a signal at 26–31 nt in testis, but not in brain, RNA (Fig. 2b). None of the sense probes yielded any signal, supporting the strand-specific accumulation of piRNAs (data not shown). Surprisingly, even a single oligodeoxynucleotide probe antisense to an individual piRNA was sufficient to detect a 26–28-nt-size signal in MILI-immunoprecipitated RNA and a 29–31-nt-size signal in total testis RNA (Fig. 2c). Oligodeoxynucleotide probes to sequences immediately flanking the isolated piRNA failed to produce a signal (data not shown), suggesting that the processing of piRNAs occurs in a directed fashion. Examination of the 5′ and 3′ ends of a 30-nt cloned piRNA by rapid amplification of cDNA ends (RACE) showed that the 5′ ends of the piRNA RACE clones were invariant in both libraries, whereas the 3′-end clones were truncated by 2 nt in the MILI immunoprecipitate small-RNA library (Supplementary Fig. 5). It is possible that the more abundant 29–31-nt fraction represents an intermediate processing product or that it corresponds to small RNAs interacting with another Piwi protein expressed in testis. Re-probing of the RNA blot for the ubiquitously expressed 22-nt miR-16 produced a miRNA signal for testis total RNA but none for MILI-immunoprecipitated RNA, indicating that MILI was specifically loaded with piRNAs but not miRNAs.

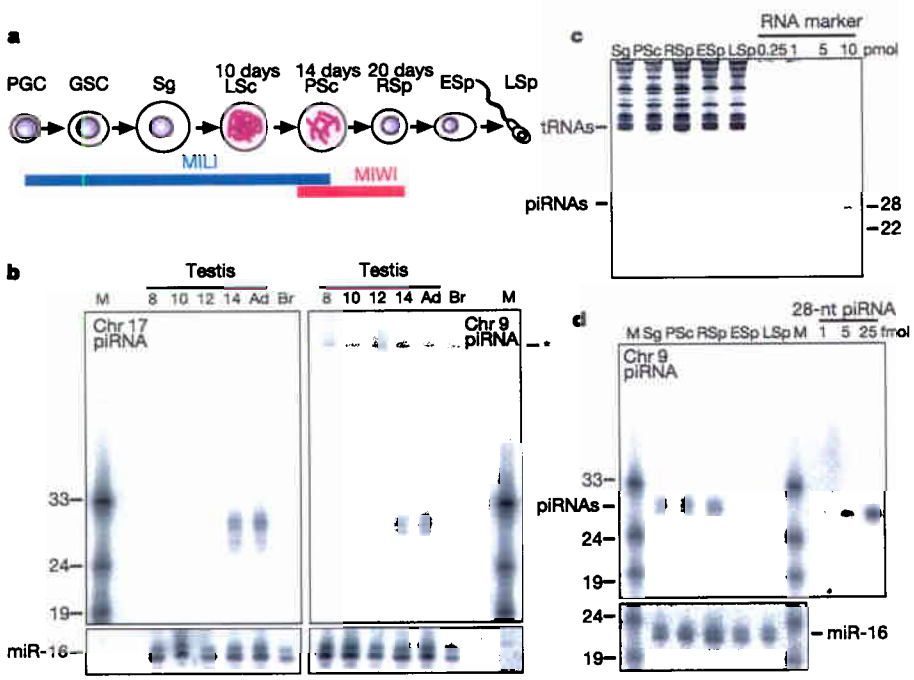

Figure 3 | Temporal expression of piRNAs during mouse spermatogenesis. **a**, Sketch of mouse spermatogenesis with the temporal expression patterns of MILI (blue) and MIWI (red) indicated. Abbreviations: PGC, primordial germline cells; GSC, germline stem cells; Sg, spermatogonia; LSc, leptotene spermatocytes; PSc, pachytene spermatocytes; RSp, round spermatids; ESp, elongating spermatids; LSp, late spermatids. **b**, Northern blot hybridization of testis total RNA from 8-, 10-, 12- and 14-day-old newborn and 3-month-old adult (Ad) mice and brain (Br) total RNA with oligodeoxynucleotide probes complementary to piRNAs (top panels) from chromosome (Chr) 17 (left panel) and chromosome 9 (right panel). The asterisk marks a cross-hybridization signal to a larger RNA also present in brain. Expression of miR-16 is monitored and serves as loading control (bottom panels). **c**, SYBR Green II staining of total RNA from elutriator-enriched male germ cells separated on a denaturing polyacrylamide gel. As reference RNA markers, two 22-nt and two 28-nt oligoribonucleotides of distinct sequences were loaded. **d**, Northern blot hybridization for samples shown in **c** using antisense probes specific for the piRNA cluster on chromosome 9 (top panel) and for miR-16 (bottom panel). A synthetic 28-nt chromosome 9 piRNA was loaded to quantify the amount of piRNA expressed in germ cells.

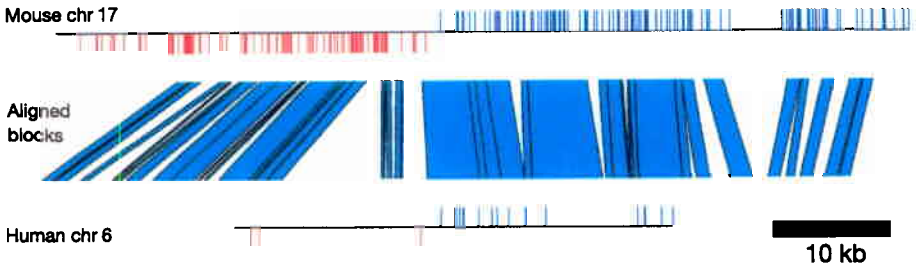

Figure 4 | Predominant mouse piRNA cluster and its orthologous cluster in human. Alignment view of the most highly expressed mouse piRNA cluster and its corresponding human orthologue. The positions of cloned sequences indicate divergent transcription from a central promoter in both species.

To obtain insight into the temporal expression of piRNAs during mouse spermatogenesis, we isolated total testis RNA at different time points of postnatal development. In mice, mitotically active spermatogonia represent the principal developing germ cells in testis up to day 6 after birth. Meiosis I is initiated on day 10, with germ cells reaching the preleptotene/leptotene, zygotene and early pachytene stages by days 10, 12 and 14, respectively[13] (Fig. 3a). MILI is expressed in male germ cells from primordial germ cells until the pachytene stage of meiosis[8], whereas MIWI is expressed in pachytene-stage spermatocytes and round spermatids[9]. Northern blotting of testis total RNA for two distinct piRNAs from mouse chromosome 9 and 17 revealed piRNA accumulation starting at day 14, when the first spermatocytes reach the pachytene stage (Fig. 3b). To assess the presence of piRNAs in specific germ-cell types, RNA was isolated from cells enriched for different stages of spermatogenesis after elutriation purification. Notably, the 30-nt piRNAs were so abundant that they were visible by SYBR Green II staining, and were estimated to be present at about 1 million piRNA molecules per mouse spermatocyte or round spermatid (Fig. 3c). Quantitative northern blotting for an individual piRNA from mouse chromosome 9 indicated about 8,000 copies per pachytene spermatocyte and about 2,000 copies per haploid round spermatid (Fig. 3d). The piRNA level was reduced by about tenfold in RNA isolated from later stages of germ-cell development.

Mouse piRNA sequences are well conserved and cluster within the closely related rat genome. However, the alignments of the mouse genome with eight other genomes[14] indicated that piRNA regions are poorly conserved between more distant species (Supplementary Fig. 6a), and that conserved elements are present with similar frequency within piRNA clusters and within introns of protein-coding genes (Supplementary Fig. 6b). For seven of the ten mouse piRNA clusters that seem to contain a bidirectional promoter, we found short regions of homology with the human genome. Moreover, we found that the frequency of ESTs that overlap with these human genomic regions orthologous to the mouse piRNA clusters was 9–21-times higher compared with the representation of testis ESTs among all the GenBank ESTs (Supplementary Table 5).

To provide experimental support for human piRNAs, we prepared and sequenced 18–26-nt and 24–33-nt small-RNA libraries from human testis total RNA (Supplementary Table 1; Fig. 1c). The small-RNA composition shows the expected enrichment for 5′ uridine, especially in the longer-size library where a 5′ uridine bias is not introduced by the presence of miRNAs. Using the same clustering criteria as for the mouse sequences, we were able to define 14 human piRNA clusters. They, together, contain 8.5% of all the cloned human sequences, and 24% of the human clones that were not derived from other functional RNAs (Supplementary Tables 8 and 9). The divergently transcribed piRNA cluster with the strongest expression in mouse is orthologous to the divergently transcribed cluster with the strongest expression in human (Fig. 4), and two additional human regions that are orthologous to divergently transcribed mouse piRNA clusters are experimentally supported (Supplementary Table 5).

Although Piwi proteins were shown to be important for stem- and germ-cell development in different animals[6–9], the underlying biochemical pathways are unknown. The identification of piRNAs provides an important molecular link regarding the function of Piwi proteins. Given the timing of the maturational arrest in *mili* and *miwi* knockout mice at the pachytene spermatocyte[8] and the spermatid steps[9], respectively, it is conceivable that piRNAs and germline-specific Piwi proteins regulate the timing of meiotic and postmeiotic events through transcriptional and translational repression. Argonaute proteins have been implicated in diverse processes such as genome rearrangement in *Tetrahymena*[15] or heterochromatic silencing and chromosome segregation in fission yeast[16], and we are only beginning to develop an understanding of the molecular mechanisms mediated by the diverse group of Argonaute ribonucleoprotein complexes.

METHODS

A detailed description of the methods used in this study is provided in Supplementary Information. Germ cells were obtained from the seminiferous tubules of 3-month-old C57BL/6J male mice using centrifugal elutriation as described previously[17]. Immunoprecipitation of MILI ribonucleoprotein complexes from mouse whole-testis whole-cell lysates was carried out using affinity-purified anti-MILI-pepN2 antibody raised against the peptide VRKDREEPRSSLPDPS (amino acids 107–122). Nucleic acids that co-immunoprecipitated with MILI were isolated, cloned and sequenced as described[18]. Small RNAs from total RNA of mouse or human testes were gel-purified, cloned and sequenced[18–20]. Cloned small RNAs were mapped to the mm6 and hg17 assemblies of the mouse and human genomes, respectively. The functional annotation was done as described previously[12,18,21], except that we updated our data set of non-coding RNAs and included predicted miRNA sequences[22–24]. For the repeat annotation, we used the repeat-masker results from the UCSC database. Clones were considered to be repeat-associated if they overlapped by at least 15 nucleotides with an annotated repeat element. The assessment of cross-species conservation of piRNAs and piRNA clusters was based on phastCons[14] conservation scores. The coverage of piRNA clusters by phastCons[14] conserved elements was also compared to the coverage of coding sequences and intronic regions of mouse RefSeq mRNAs[25] by conserved elements. Northern blots were performed as described previously[20], and the preparation and sequences of reference standards and probes are provided in Supplementary Information.

Received 18 February; accepted 17 May 2006.
Published online 4 June 2006.

1. Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**, 343–349 (2004).
2. Mello, C. C. & Conte, D. Jr. Revealing the world of RNA interference. *Nature* **431**, 338–342 (2004).
3. Matzke, M. A. & Birchler, J. A. RNAi-mediated pathways in the nucleus. *Nature Rev. Genet.* **6**, 24–35 (2005).
4. Zamore, P. D. & Haley, B. Ribo-gnome: the big world of small RNAs. *Science* **309**, 1519–1524 (2005).
5. Carmell, M. A., Xuan, Z., Zhang, M. Q. & Hannon, G. J. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev.* **16**, 2733–2742 (2002).
6. 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).
7. Reddien, P. W., Oviedo, N. J., Jennings, J. R., Jenkin, J. C. & Sanchez Alvarado, A. SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science* **310**, 1327–1330 (2005).

8. Kuramochi-Miyagawa, S. *et al. Mili*, a mammalian member of *piwi* family gene, is essential for spermatogenesis. *Development* **131**, 839–849 (2004).
9. Deng, W. & Lin, H. *miwi*, a murine homolog of *piwi*, encodes a cytoplasmic protein essential for spermatogenesis. *Dev. Cell* **2**, 819–830 (2002).
10. Aravin, A. A. *et al.* The small RNA profile during *Drosophila melanogaster* development. *Dev. Cell* **5**, 337–350 (2003).
11. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
12. Chen, P. Y. *et al.* The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes Dev.* **19**, 1288–1293 (2005).
13. Bellve, A. R. *et al.* Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization. *J. Cell Biol.* **74**, 68–85 (1977).
14. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005).
15. Mochizuki, K. & Gorovsky, M. A. Small RNAs in genome rearrangement in *Tetrahymena*. *Curr. Opin. Genet. Dev.* **14**, 181–187 (2004).
16. Grewal, S. I. & Rice, J. C. Regulation of heterochromatin by histone methylation and small RNAs. *Curr. Opin. Cell Biol.* **16**, 230–238 (2004).
17. Jenab, S. & Morris, P. L. Testicular leukemia inhibitory factor (LIF) and LIF receptor mediate phosphorylation of signal transducers and activators of transcription (STAT)-3 and STAT-1 and induce c-*fos* transcription and activator protein-1 activation in rat Sertoli but not germ cells. *Endocrinology* **139**, 1883–1890 (1998).
18. Pfeffer, S. *et al.* Identification of microRNAs of the herpesvirus family. *Nature Methods* **2**, 269–276 (2005).
19. Pfeffer, S., Lagos-Quintana, M. & Tuschl, T. in *Current Protocols in Molecular Biology* (eds Ausubel, F. M. *et al.*) 26.4.1–26.4.18 (John Wiley and Sons, New York, 2003).
20. Lagos-Quintana, M. *et al.* Identification of tissue-specific microRNAs from mouse. *Curr. Biol.* **12**, 735–739 (2002).
21. Pfeffer, S. *et al.* Identification of virus-encoded microRNAs. *Science* **304**, 734–736 (2004).
22. Lim, L. P., Glasner, M. E., Yekta, S., Burge, C. B. & Bartel, D. P. Vertebrate microRNA genes. *Science* **299**, 1540 (2003).
23. Berezikov, E. *et al.* Phylogenetic shadowing and computational identification of human microRNA genes. *Cell* **120**, 21–24 (2005).
24. Legendre, M., Lambert, A. & Gautheret, D. Profile-based detection of microRNA precursors in animal genomes. *Bioinformatics* **21**, 841–845 (2005).
25. Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **33**, D501–D504 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Brennecke, M. Landthaler, Y. Pei, K. Fejes Toth and A. Sewer for discussion and help with the manuscript. We also thank L. Mitchell and K. Hwang for expert assistance in germ-cell purification, M. Poy and M. Stoffel for providing mouse tissues, and D. Weir and R. Choksi for assistance with sequencing. This work was supported by a FRAXA Research Foundation postdoctoral fellowship to A.A., an NIH grant to P.M., NIH grants to T.T., and an SNF grant to M.Z.

Author Contributions A.A., S.P. and M.L.-Q. prepared the mouse, and P.L. and N.I. the human, testes small RNA libraries. A.A. recognized the presence of piRNAs, performed the MILI IPs, prepared the MILI-interacting small RNA library, and performed, together with N.I., the northern blotting analysis. S.K.-M. and T.N. produced, characterized and purified the MILI antibody. T.T. developed the concept of cloning from Ago/Piwi IPs. P.M. isolated germline cells. M.C., J.J.R. and J.J. performed the large-scale sequencing. The bioinformatic analyses of piRNAs were designed and carried out by D.G. and M.Z. with input from A.A. and T.T. The database of RNAs with known function was compiled by M.Z., R.S., P.L. and S.P., the software used for small RNA annotation was developed by M.Z., R.S. and C.S., and the manuscript was written by A.A., M.Z. and T.T.

Author Information Sequences of the piRNAs determined in this paper are given in Supplementary Tables 4 and 9. Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.Z. (Mihaela.Zavolan@unibas.ch) or T.T. (ttuschl@rockefeller.edu).