

MIWI and piRNA-mediated cleavage of messenger RNAs in mouse testes

Peng Zhang^{1,*}, Jun-Yan Kang^{2,3,*}, Lan-Tao Gou^{2,3,*}, Jijia Wang¹, Yuanchao Xue^{5,6}, Geir Skogerboe⁴, Peng Dai^{2,3}, Da-Wei Huang¹, Runsheng Chen⁴, Xiang-Dong Fu^{5,6}, Mo-Fang Liu^{2,3}, Shunmin He¹

¹Key Laboratory of the Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China; ²Center for RNA Research, State Key Laboratory of Molecular Biology, University of Chinese Academy of Sciences, ³Shanghai Key Laboratory of Molecular Andrology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China; ⁴Laboratory of Bioinformatics and Noncoding RNA, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China; ⁵State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan, Hubei 430072, China; ⁶Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA 92093-0651, USA

The piRNA machinery is known for its role in mediating epigenetic silencing of transposons. Recent studies suggest that this function also involves piRNA-guided cleavage of transposon-derived transcripts. As many piRNAs also appear to have the capacity to target diverse mRNAs, this raises the intriguing possibility that piRNAs may act extensively as siRNAs to degrade specific mRNAs. To directly test this hypothesis, we compared mouse PIWI (MIWI)-associated piRNAs with experimentally identified cleaved mRNA fragments from mouse testes, and observed cleavage sites that predominantly occur at position 10 from the 5' end of putative targeting piRNAs. We also noted strong biases for U and A residues at nucleotide positions 1 and 10, respectively, in both piRNAs and mRNA fragments, features that resemble the pattern of piRNA amplification by the 'ping-pong' cycle. Through mapping of MIWI-RNA interactions by CLIP-seq and gene expression profiling, we found that many potential piRNA-targeted mRNAs directly interact with MIWI and show elevated expression levels in the testes of *Miwi* catalytic mutant mice. Reporter-based assays further revealed the importance of base pairing between piRNAs and mRNA targets and the requirement for both the slicer activity and piRNA-loading ability of MIWI in piRNA-mediated target repression. Importantly, we demonstrated that proper turnover of certain key piRNA targets is essential for sperm formation. Together, these findings reveal the siRNA-like function of the piRNA machinery in mouse testes and its central requirement for male germ cell development and maturation.

Keywords: piRNA; MIWI; sperm

Cell Research (2015) 25:193–207. doi:10.1038/cr.2015.4; published online 13 January 2015

Introduction

The PIWI proteins are a subfamily of the AGO/PIWI

proteins that are mainly expressed in animal germ cells [1]. A class of 24–31-nt PIWI-binding small noncoding RNAs (ncRNAs), termed as PIWI-interacting RNAs (piRNAs), have been identified in mammalian testes [2–5]. This class of small ncRNAs is slightly longer than microRNAs (miRNAs) and small interfering RNAs (siRNAs), exhibits great sequence variety and generally is mapped to clustered loci in the genome. It has been proposed that the biogenesis of piRNAs includes a primary processing pathway and a 'ping-pong' amplification cycle pathway [6–8]. Increasing evidence suggests that piRNAs play a critical role in both epigenetic and post-transcriptional silencing of transposons [1, 6]. Several stud-

*These three authors contributed equally to this work.

Correspondence: Shunmin He^a, Mo-Fang Liu^b, Xiang-Dong Fu^c

^aTel: 86-10-6480 7279; Fax: 86-10-6480 7279

E-mail: heshunmin@gmail.com

^bTel: 86-21-5492 1146; Fax: 86-21-5492 1011

E-mail: mliu@sibcb.ac.cn

^cTel: 1-85-8534 4937; Fax: 1-85-8822 6920

E-mail: xdfu@ucsd.edu

Received 21 April 2014; revised 11 October 2014; accepted 3 November 2014; published online 13 January 2015

ies indicate that piRNAs in *Drosophila* have the capacity to downregulate the expression of certain mRNA targets via a miRNA-like mechanism that operates by imperfect base-pairing rules [9-13]. However, it is not yet known whether and to what extent this mechanism is involved in regulating gene expression in mammals.

In mice, PIWI proteins consist of three members: MILI, MIWI and MIWI2. The expression of *Mili*, *Miwi* and *Miwi2* during male germ cell differentiation is temporally and spatially regulated, which is essential for spermatogenesis. Interestingly, mouse piRNAs are expressed in two distinct phases during spermatogenesis, termed as pre-pachytene and pachytene piRNAs, respectively. Pre-pachytene piRNAs are mainly derived from transposons and are coexpressed with MIWI2 and/or MILI in fetal and perinatal male germ cells, and function in transposon silencing, thereby protecting the genome in early-stage spermatogenic cells [14-16]. Pachytene piRNAs, on the other hand, are massively produced from non-transposon intergenic regions in later pachytene spermatocytes and post-meiotic spermatids, and are mainly associated with MIWI. However, little is known about the functional significance of MIWI and the MIWI-bound pachytene piRNAs.

A recent study indicates that the slicer activity of MIWI is responsible for cleavage of transposon transcripts in mouse testes [17]. However, it has remained unclear whether MIWI exerts a similar piRNA-guided endonuclease activity on piRNA-targeted mRNA transcripts. In this study, we analyzed the published global 5'-RACE libraries from mouse testes for cleaved RNA fragments that contain a 5'-monophosphate, a structural feature likely to be generated by certain RNA slicer activity [17]. We identified approximately 200 potential piRNA target sites on specific mRNAs, thus implicating piRNAs in mediating cleavage of these mRNAs in mouse male germ cells. Following this initial finding, we performed anti-MIWI crosslinking immunoprecipitation coupled with deep sequencing (CLIP-seq) on enriched mouse round spermatids and demonstrated that ~43% of the identified piRNA target sites were physically associated with MIWI. Subsequently, we validated the MIWI/piRNA-mediated repression of target mRNAs with reporter-based assays and showed that both the slicer activity and the piRNA-loading ability of MIWI are critical for the piRNA-mediated target repression. Importantly, ectopic expression of piRNA-resistant forms of target genes in spermatids prevented the formation of sperms, suggesting that proper turnover of certain piRNA targets is essential for normal spermiogenesis. Together, these results reveal a pachytene piRNA-directed mRNA cleavage program that is critical for mouse male germ cell

development.

Results

Identification of potential MIWI-cleaved mRNA fragments

The slicer activity of PIWI proteins is required for piRNA amplification by the 'ping-pong' mechanism and post-transcriptional silencing of transposons [7, 8, 17, 18]. Taking advantage of 5'-RACE cleavage fragment libraries derived from adult testes of both control (*Miwi*^{+/-}) and MIWI slicer activity-deficient mutant (*Miwi*^{-/ADH}) mice [17], we selected fragments that show perfect matches with the sense strands of RefSeq mRNAs and considered these as potential MIWI-cleaved products. To ensure that all selected fragments are derived only from mRNAs, we only included fragments that correspond to unique loci in the mouse genome. Together with a library of MIWI-associated piRNAs from wild-type adult mouse testes [17], we noted a positional relationship between potential piRNA-binding sites and a subset of mRNA fragments. As illustrated on *Psm8*, *BC026590* and *Mdc1*, the putative cleavage sites appear to overlap with potential piRNA-binding sites with the cleavage position precisely located 10 nt from the 5' end of the matching piRNAs (Figure 1 and Supplementary information, Figure S1).

The *Psm8* mRNA fragments are mainly derived from its 3'-UTR and represent the most abundant cleavage products (645 reads per million (RPM)) in *Miwi*^{+/-} testes (Figure 1A, top). Interestingly, the 3'-UTR of *Psm8* appears to harbor potential target sites for several piRNAs (~280 reads), and importantly, the 5' ends of ~96% of these piRNAs lay 10 nt downstream of the putative mRNA cleavage site (Figure 1A and 1B). Consistent with these mRNA fragments being piRNA-mediated cleavage products, their abundance was drastically reduced in the library derived from *Miwi*^{-/ADH} testes (Figure 1A, bottom), and the *Psm8* mRNA level was indeed elevated in such mutant testes (Figure 1C). We similarly identified multiple mRNA fragments from the 5'-RACE library, which overlap with potential piRNA-binding sites on the *BC026590* (Figure 1D and 1E) and *Mdc1* mRNAs (Supplementary information, Figure S1A and S1B), and observed elevated expression of these mRNAs in *Miwi*^{-/ADH} mutant testes (Figures 1F and Supplementary information, Figure S1C). These data strongly suggest that MIWI is involved in piRNA-guided mRNA cleavage in mouse testes.

Widespread MIWI/piRNA-mediated mRNA cleavage events in mouse testes

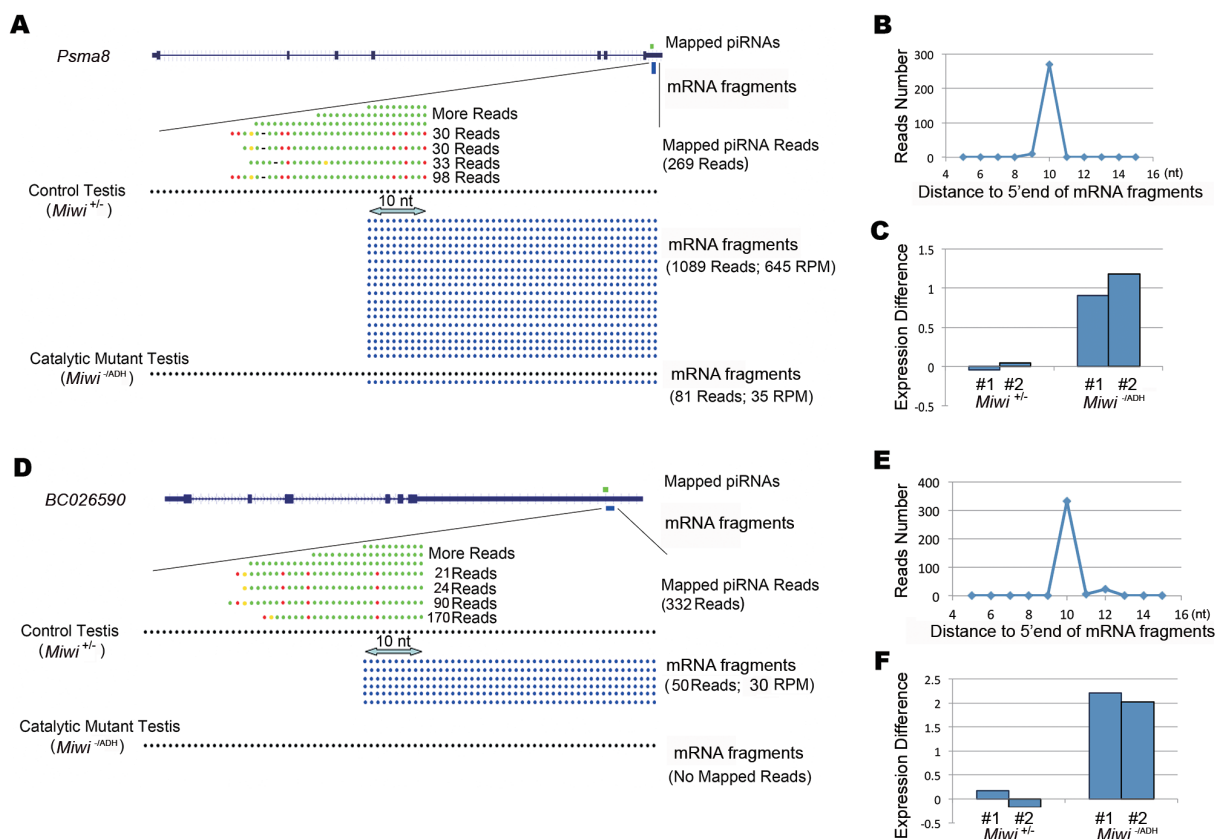


Figure 1 Evidence for piRNA-guided cleavage of target mRNAs. **(A)** Sequence overlap between 5'-RACE-identified mRNA fragments and predicted targeting piRNAs in the 3'-UTR of *Pisma8* transcripts. The upper part shows the location of *Pisma8* mRNA fragments. The piRNA data are from normal mouse testes (GSM822760, *Miwi*^{+/-}), and the mRNA fragment data are from normal (GSM822765, *Miwi*^{+/-}) or catalytic mutant (GSM822766, *Miwi*^{-ADH}) mice [17]. Blue dots: mRNA fragments at the single-nucleotide resolution; green dots: piRNAs with extensive base pairing with the mRNA; red dots: mismatches; yellow dots: G-U wobble pairs; short lines: gaps. **(B)** The distribution of matched piRNA reads near the mRNA fragment in **A**. **(C)** Comparison of *Pisma8* mRNA levels (\log_2 transferred followed by subtracting the mean of the control values) in *Miwi* slicer activity-deficient mutant (*Miwi*^{-ADH}) mouse testes relative to *Miwi*^{+/-} controls (data from GEO: GSE32180 [17]). **(D)** mRNA fragment reads and targeting piRNAs at 3'-UTR of *BC026590* similarly illustrated as in **A**. **(E)** The distribution of matched piRNA reads near the mRNA fragment as in **B**. **(F)** Expression difference of *BC026590* in *Miwi*^{-ADH} mutants relative to *Miwi*^{+/-} controls as described in **C**.

The above examples of potential MIWI/piRNA-cleaved mRNA fragments prompted us to ask how common such cases exist in mouse testes. To this end, we systematically screened potential mRNA targets for piRNAs expressed in the mouse testes. Given that perfect or near-perfect base pairing between piRNA and mRNA is required for efficient cleavage by MIWI [17], we first searched for base pairing with less than 3 mismatches between the piRNA from position 2 to 21 and the mRNA (see the base-pairing rule for piRNA targeting below). Next, we examined the distribution of potential piRNA-binding sites near the 5' end of mRNA fragments identified in the 5'-RACE library. Consistent with the previous observation that MIWI cuts targets at the posi-

tion 10 nt away from the 5' end of targeting piRNAs [17], we observed a strong enrichment of matching piRNAs whose 5' ends fall in the position 10 nt downstream of the cleavage site (Figure 2A). In contrast, such enrichment was absent among mRNA fragments derived from the *Miwi*^{-ADH} mouse testes (Figure 2B), suggesting that the slicer activity of MIWI is essential for the production of such mRNA fragments. We obtained similar results if we only counted mRNA fragments instead of the targeting piRNAs (Supplementary information, Figure S2A). Altogether, we identified a total of 1 878 piRNA:mRNA interactions based on the peak at position 10 in *Miwi*^{+/-} mouse testes (Figure 2A, Table 1 and Supplementary information, Table S1), which involves 1 295 non-redun-

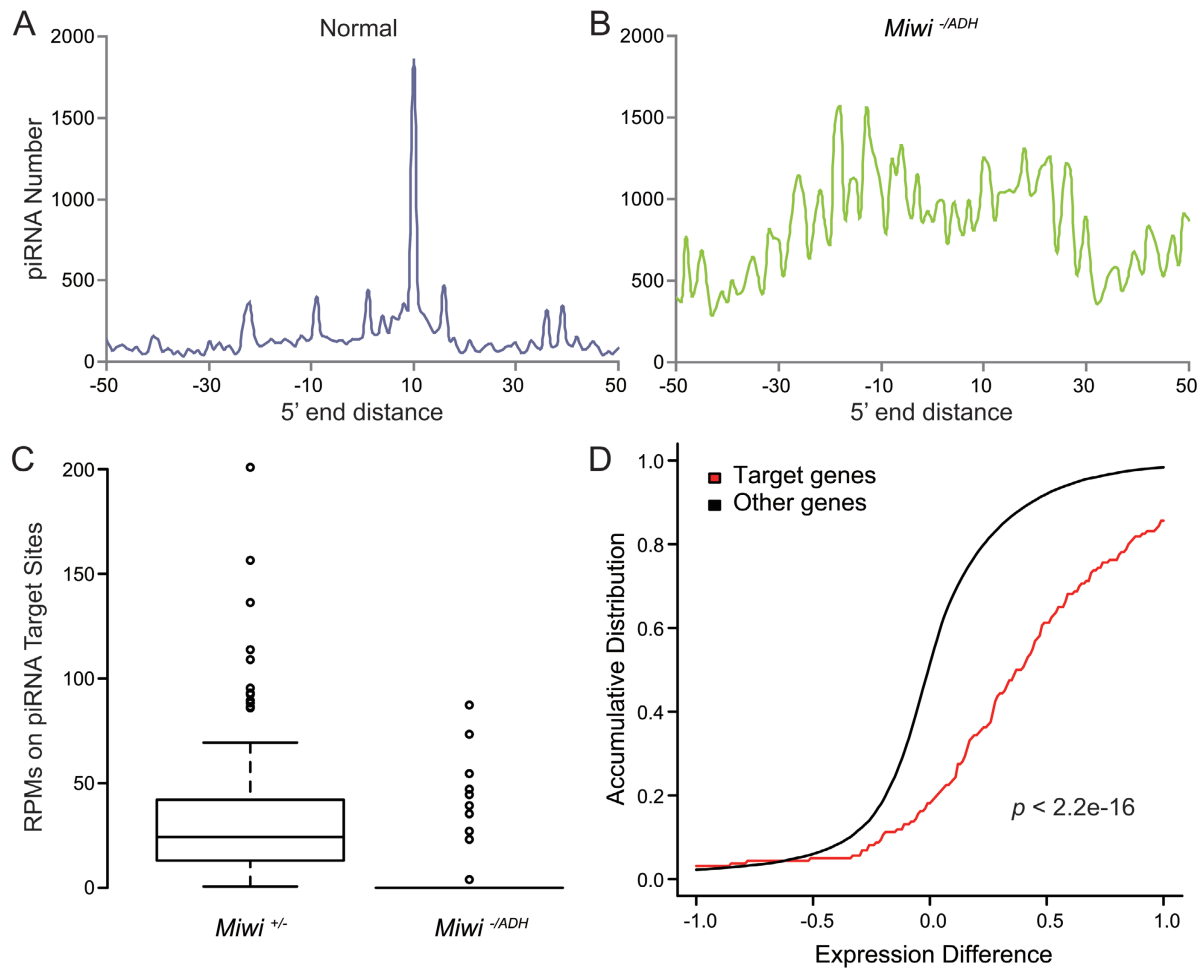


Figure 2 Meta-analyses of piRNA target mRNAs in mouse testes. **(A, B)** The distribution of potential piRNA target sites around cleaved mRNA fragments. The piRNA data are from normal (GSM822760, *Miwi*^{+/+}) or the corresponding catalytic mutant (GSM822764, *Miwi*^{-ADH}) mice. The mRNA fragment data are from normal (GSM822765, *Miwi*^{+/+}) or catalytic mutant (GSM822766, *Miwi*^{-ADH}) mice [17]. The X-axis shows the position of the 5' end of potential mRNA-targeting piRNAs relative to the 5' end of mRNA fragments. **(C)** Boxplot of distribution of mRNA fragment reads at potential MIWI/piRNA cleavage sites in control and catalytic mutant testes. **(D)** Accumulative distribution of expression difference of predicted piRNA target genes and non-targeted genes between the catalytic mutant (*Miwi*^{-ADH}) and control (*Miwi*^{+/+}) samples (gene expression profiling data, GSE32180). The gene expression difference is presented as the averaged log₂ expression of individual genes in the catalytic mutants minus the averaged log₂ expression in controls. The *P*-value was calculated by Wilcoxon rank sum test.

Table 1 Summary of potential mRNA-targeting piRNAs and their target sites in mRNAs

Target sites per gene	Target genes	Target sites	Unique targeting piRNAs	Cleaved fragment reads
1	150	150	1 066	9 090
2	15	30	202	1 274
3 or more	4	13	46	638
Total	169	193	1 295	11 002

dant piRNAs and 193 mRNA fragments. We regard these 193 fragments residing in 169 specific protein-coding genes as potential piRNA target sites.

The above results strongly implicate an essential role

of MIWI slicer activity in the regulation of mRNA stability beyond the two specific mRNAs shown in Figure 1. To further substantiate this, we compared the abundance of mRNA fragments with 5' ends at the predicted

piRNA-guided cleavage sites in *Miwi*^{+/-} vs *Miwi*^{-ADH} 5'-RACE libraries [17]. We found that the reads of such mRNA fragments were sharply reduced in *Miwi*^{-ADH} mice (Figure 2C). Moreover, ~94% of the predicted piRNA target sites were absent at the 5' ends of mRNA fragments in the 5' RACE library from *Miwi*^{-ADH} mice, supporting an essential role of MIWI slicer activity in generating mRNA fragments cleaved at such piRNA target sites. We further compared the expression levels of the predicted piRNA target genes between *Miwi*^{+/-} and *Miwi*^{-ADH} testes, and found that the levels of these mRNAs were significantly upregulated relative to other non-target genes in the *Miwi* catalytic mutant (Figure 2D and Supplementary information, Figure S2B). Gene Set Enrichment Analysis (GSEA) [19] also showed significant upregulation ($P < 0.001$) of the predicted target gene set in the catalytic mutant (Supplementary information, Figure S3). Collectively, these results suggest that piRNAs may regulate their target mRNAs via a siRNA-like mechanism in mouse testes.

Analysis of MIWI-mediated regulation of mRNA cleavage by CLIP-seq

To provide experimental evidence for MIWI/piRNA-mediated regulation of mRNA cleavage in mouse testes, we performed MIWI CLIP-seq, as we recently described [20], to determine the physical interactions between MIWI and RNAs. Considering our previous results showing abundant MIWI and its associated piRNAs in round spermatids [21], we performed MIWI CLIP-seq on round spermatids isolated from adult mouse testes (Supplementary information, Figure S4A). We identified ~48 million reads that could be mapped to the mouse genome. Considering reads with 25-33 nt in length as piRNAs and longer reads (sequenced up to 36 nt) as MIWI-bound target RNAs (Figure 3A), we found that a total of ~17 million reads correspond to piRNAs and ~10 million reads correspond to MIWI targets. Importantly, while the majority of MIWI-bound piRNAs and targets are derived from various repeat-containing transcripts and long noncoding RNAs (lncRNAs), a substantial fraction of reads are also mapped to the sense strand of protein-coding genes (Figure 3B).

Interestingly, of 193 predicted piRNA target sites in 169 genes shown in Table 1, 164 target sites in 147 genes could be matched with MIWI CLIP-identified piRNAs at the expected cleavage position. Fragments of almost all target genes (168/169) were detected by MIWI CLIP-seq assay. Among these 168 genes, 83 predicted piRNA-target sites in 73 genes fall within 15 nt from the 5'-RACE-mapped cleavage sites. Under this strict condition, the predicted piRNA target sites on 65 of these 73 genes

were supported by both MIWI CLIP-identified piRNAs and mRNA targets (Figure 3C). Using the mean coverage rate of CLIP-identified mRNA targets (including the flanking 15 nt) as the expected probability to perform a binomial test, we found that the coincidence of 193 predicted piRNA target sites relative to CLIP-identified targets within 169 genes was highly significant ($P = 0.00018$).

We noted that the expression levels of the predicted piRNA target genes that also show MIWI binding evidence were higher than those of other predicted target genes (*t*-test, $P = 0.0028$; Supplementary information, Figure S4B), indicating that CLIP-seq tends to catch target genes with higher expression levels. We also observed that the putative cleavage sites predicted from the 5'-RACE library exhibited the highest coincidence with the 5' ends of MIWI CLIP-identified mRNA fragments (Figure 3D), indicating that the cleaved products are physically associated with MIWI. Similarly to the observation depicted in Figure 2A, we also observed a significant enrichment of piRNAs (obtained from the *Miwi*^{+/+} dataset GSM822760 [17]) whose 5' ends match the position 10 from the 5' ends of CLIP-identified mRNA fragments, with ~500 genes presented in the peak at position 10, suggesting that the number of piRNA target genes may be more than that predicted by the 5'-RACE library (Supplementary information, Figure S4C). It is worth noting that compared with the results obtained from the 5'-RACE library (Figure 2A), the CLIP data showed higher background noise, which is likely contributed by MIWI-bound, but uncleaved mRNAs.

Validation of piRNA regulation of mRNA expression

Previous studies showed specific expression of *Miwi* during the developmental transition from meiotic spermatocytes to elongating spermatids and spermatogenic arrest in round spermatids in *Miwi*-null mice [22]. To experimentally determine the regulation of mRNA expression by MIWI and piRNAs, we selected 7 (*Atr*, *BC026590*, *Gfpt1*, *Mdc1*, *Ppp1cb*, *Psmad8* and *Tox4*) from the 169 predicted target genes for quantitative analysis. Among these genes, four (*Atr*, *BC026590*, *Mdc1* and *Ppp1cb*) contain both MIWI CLIP-supported targeting piRNAs and target sites in their mRNAs. We isolated round spermatids (RS) and spermatocytes (SC) from adult *Miwi*^{+/-} and *Miwi*^{-/-} testes by unit gravity sedimentation as we recently described [21], and determined the expression of these 7 target genes by RT-qPCR. We found that all the 7 tested genes were expressed at significantly elevated levels in round spermatids from *Miwi*^{-/-} compared to *Miwi*^{+/-} mice, with the largest differences observed for *BC026590* and *Psmad8* (Figure 4A

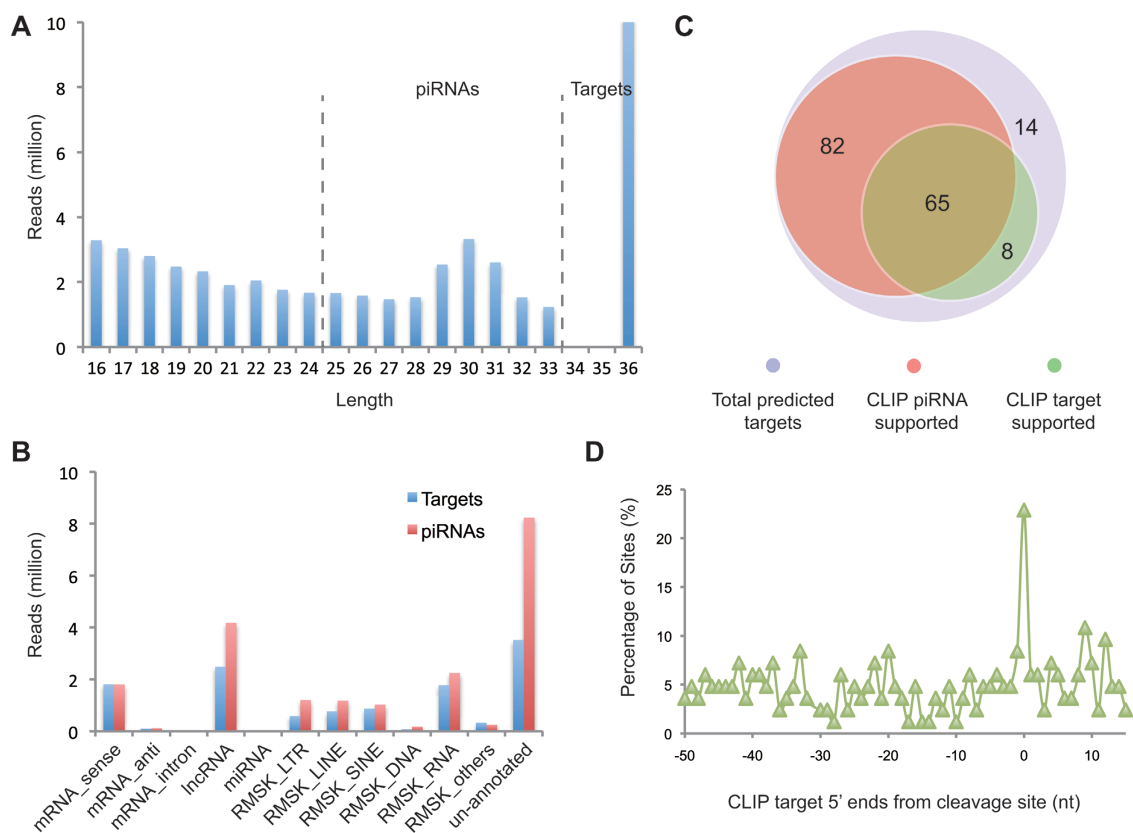


Figure 3 piRNAs and piRNA targets identified by MIWI CLIP-seq. **(A)** The length distribution of mapped RNAs detected by MIWI CLIP-seq. Short sequences of 25–33 nt likely correspond to piRNAs, while longer sequences (only 36 nt were sequenced) correspond to MIWI targets. **(B)** Genome annotations for MIWI CLIP-identified targets and piRNAs. **(C)** Overlaps between genes that contain target sites bound by MIWI CLIP-identified piRNAs (red) and genes that contain target sites directly bound by MIWI (green) in 169 piRNA target genes (blue) predicted from the 5'-RACE library. Only the MIWI CLIP-identified piRNAs bound at the expected cleavage sites and the MIWI CLIP-identified mRNA sequences fall within 15 nt from the cleavage sites were taken into account. **(D)** Alignment of the 5' ends of MIWI CLIP-identified mRNA targets (> 33 nt) with the predicted cleavage sites from the 5'-RACE library. The Y-axis represents the percentage of predicted target sites with the existence of 5' ends of MIWI CLIP-identified targets mapped at positions, as indicated by the X-axis, relative to the predicted cleavage sites.

and Supplementary information, Figure S5A). The data thus support a direct role of MIWI in the regulation of mRNA stability in mouse testes.

We next performed reporter-based assays to determine the requirement for specific piRNA:mRNA base-pairing interactions in MIWI-dependent regulation of mRNA stability. We selected 3 piRNA:mRNA pairs (*Psmad8*, *BC026590* and *Tox4*) for the analysis and placed the predicted target sequences in the 3'-UTR of the luciferase reporter. We tested two controls for each reporter, one with an 8-nt deletion in the 5' portion of piRNA-targeting region and the other with an 8-nt mutation at the 5' end of piRNAs (Supplementary information, Figure S5B). A mouse spermatocyte cell line GC-2spd(ts) [23], which

we recently showed to be a suitable *in vitro* model for probing piRNA functions [20], was used for reporter assays. Intriguingly, we found that the selected piRNAs for reporter assays (piR-mmu-77755 [24], piR-671 and piR-125394) were all absent in GC-2spd(ts) cells, despite the presence of piRNA and *Miwi* expression in these cells (NCBI-SRA: SRR1482465) [20]. When the reporters were co-transfected along with the cognate chemically synthesized piRNAs (containing a phosphate at the 5' end and a 2'-O-methyl group at the 3' end) into GC-2spd(ts) cells, we found that all wild-type reporters, but not their mutant versions, were significantly repressed by co-transfected wild-type, but not mutant, piRNAs (Figure 4B and Supplementary information, Figure S5C). RT-qP-

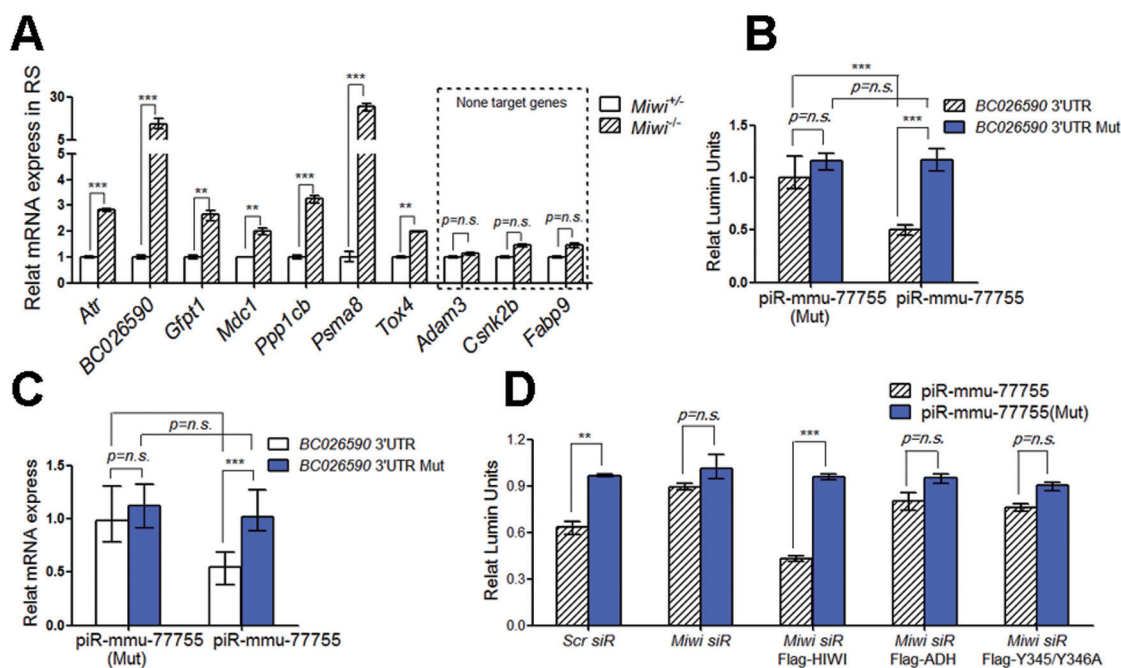


Figure 4 piRNA-induced mRNA degradation requires the slicer and piRNA-loading activities of MIWI. **(A)** RT-qPCR analyses of the mRNA levels of *Atr*, *BC026590*, *Gfp1*, *Mdc1*, *Ppp1cb*, *Psm8*, *Tox4*, *Adam3*, *Csnk2b*, and *Fabp9* in enriched *Miwi*^{-/-} round spermatids relative to those in control *Miwi*^{+/+} round spermatids. **(B, C)** Dual-luciferase reporter **(B)** or qRT-PCR assays **(C)** of the effect of piR-mmu-77755 [24] on the expression of *BC026590* 3'-UTR reporter in GC-2spd(ts) cells. A mutant piR-mmu-77755 (Mut) served as a negative control. **(D)** Ectopic Flag-HIWI, but not Flag-ADH or Flag-Y345/346A HIWI mutant, restored piR-mmu-77755-induced *BC026590* reporter repression in *Miwi* siR-treated GC-2spd(ts) cells. Data shown represent mean \pm SEM of triplicate experiments. ** $P < 0.01$, *** $P < 0.001$ (Student's *t*-test). Results shown are representative of three independent experiments.

CR confirmed piRNA-dependent downregulation of the reporter mRNA in each case (Figure 4C and Supplementary information, Figure S5D).

We next examined the requirement for the MIWI protein in piRNA-induced target repression. To this end, we knocked down endogenous *Miwi* in GC-2spd(ts) cells by siRNAs. As expected, depletion of *Miwi* strongly attenuated the effect of piR-mmu-77755 on the *BC026590* reporter activity (Figure 4D). We further examined the requirement for the MIWI slicer activity in piRNA-mediated target repression by complementing *Miwi* siRNA-treated GC-2spd(ts) cells with human PIWI (HIWI), which is highly homologous (~94% identity) to MIWI [22], or its slicer activity-deficient ADH mutant [17]. We found that, while wild-type HIWI was able to restore the piRNA function in *Miwi*-depleted GC-2spd(ts) cells, the ADH mutant HIWI failed to do so (Figure 4D). We also tested another mutant HIWI (Y345/346A) known to be defective in piRNA loading [21], and found that the ability to load piRNA is also essential for piRNA-dependent, HIWI-mediated repression of the reporter (Figure 4D).

These data unequivocally demonstrate that both the slicer activity and the piRNA-loading ability of MIWI/HIWI are indispensable for piRNA-induced target repression.

Rules for piRNA binding to target mRNA

Previous *in vitro* assays suggest that the MIWI catalytic activity requires continuous complementarity of ≥ 20 nt between piRNA and target [17]. However, both our computational analyses and reporter assays (Figure 4 and Supplementary information, Table S1 and Figure S5) suggest that a somewhat less strict base-pairing rule might suffice for MIWI/piRNA-mediated degradation of mRNA targets in the cell. We thus further explored the rules that govern piRNA-guided cleavage of mRNAs.

1. Length of complementarity We first determined the minimal continuous base-pairing length, starting from the second nucleotide of piRNA, and found that the enrichment for complementary piRNAs at the position 10 from the 5' end of mRNA cleavage sites persistently remained under the perfect match (PM) of 16-22 nt in length (PM16-PM22; Supplementary information, Figure

S6A). To estimate the false positive rate, we calculated the signal:noise ratio, in which the signal represents the peak value of piRNA counts at position 10, and the background noise represents average of values at positions ± 5 nt from position 10 (not including the position 10). Signal:noise ratios considerably above 1 were obtained with complementarity lengths as short as 16 base pairs (Supplementary information, Figure S6B). In sharp contrast, no significant enrichment was observed with the expressed piRNAs and mapped mRNA fragments from *Miwi*^{-/-ADH} testes, even when we only counted piRNAs with 20 continuous base pairs (Supplementary information, Figure S6C). These results suggest that MIWI-mediated target cleavage can be achieved with shorter continuous base pairing between piRNAs and their targets than previously thought.

2. Allowable mismatches We next tested the effect of mismatches (allowing 1-3 mismatches) in a continuous sequence of 20 nt starting from position 2 on piRNAs. Intriguingly, we still observed a significant enrichment of piRNAs at position 10 from the cleavage sites, even with 3 mismatches (Supplementary information, Figure S6D). However, as the number of mismatches increased, the signal:noise ratio was dramatically reduced (Supplementary information, Figure S6E), indicating that piRNAs with more than 3 mismatches are less likely to induce target cleavage.

3. Base-pairing potential from the 5' end of piRNA Previous studies indicate that mismatches at the first base of the piRNA do not affect piRNA-guided target cleavage by MIWI [17]. To further examine positional effects of mismatches, we examined piRNA-binding events around the mRNA cleavage sites with progressively increased mismatches from the 5' end of piRNAs and < 3 mismatches in the subsequent 19-nt sequences. We obtained the highest piRNA-binding rate with paired base beginning at position 1, indicating that the first base-pair interaction is important for piRNA-mediated mRNA cleavage. We continued to detect an appreciable peak when paired base begins at position 2 or 3, but no peak was observed when paired base starts at position 4 or beyond (Supplementary information, Figure S6F and S6G).

These results suggest a general rule for predicting piRNA targeting regions on mRNAs, which requires a relatively strict base pairing at the 5' end of piRNA with mismatches of < 3 in the following 20-nt sequence. Notably, the mRNA cleavage site at position 10 did not shift to a further downstream position when base pairing starts at position 2 or 3, indicating that the cleavage site is still present between the 10 and 11 nt relative to the piRNA 5' end, rather than counting from the start base-pairing position between piRNAs and mRNAs.

Features of piRNA target sites

We further analyzed the 193 predicted piRNA target sites (Supplementary information, Table S1) and found that most of these sites (~75%) were targeted by more than one piRNA (Figure 5A). Intriguingly, we found that more than 70% of these potential target sites were located in the 3'-UTRs (Figure 5B), despite that 3'-UTR sequences only account for ~30% of all mRNA fragments in the 5'-RACE library, suggesting that piRNA target sites are preferentially located in the 3'-UTRs of target mRNAs.

We further examined the conservation of potential piRNA target sites among placental mammals [25], and found that the sequences around piRNA target sites are more conserved than surrounding sequences (Figure 5C), indicating that such piRNA-mediated regulation of mRNA stability might be an evolutionarily conserved mechanism in placental mammals.

Features of mRNA-targeting piRNAs

We next analyzed 1 295 predicted mRNA-targeting piRNAs (Supplementary information, Table S1) and found that these piRNAs range from 28 to 32 nt in length (Supplementary information, Figure S7A), consistent with the established preference of MIWI for piRNAs of ~30 nt [2-3]. We observed that ~60% of these piRNAs map to one unique locus (Supplementary information, Figure S7B), while a small number of these piRNAs map to more than 100 loci (some even map to > 1 000 loci) in the mouse genome. Moreover, we found that the majority of these 1 295 piRNAs originate from loci that are annotated as repeat elements by RepeatMasker, while a small fraction of them map to lncRNAs or coding gene loci (Supplementary information, Figure S7C). Intriguingly, the predicted target sites exhibit more matches to non-repeat loci in mRNAs, although non-repeat-derived piRNAs only represent a small portion of the 1 295 predicted mRNA-targeting piRNAs (Supplementary information, Figure S7D). In contrast, many repeat-derived piRNAs seem to target a small number of the same sites in mRNAs (Supplementary information, Table S1).

Previous studies reported that piRNAs from adult mouse testes generally do not show signature of ping-pong cycle biogenesis [26-27]. Unexpectedly, we found that this group of mRNA-targeting piRNAs possesses both key signatures of the ping-pong mechanism, i.e., the typical bias of U at the 5' end and A at position 10 of piRNAs (Figure 5D, top), which differ from the majority of MIWI-associated piRNAs in adult mouse testes [26-27]. Moreover, consistent with these features in mRNA-targeting piRNAs, we observed the corresponding bias for U at the mRNA cleavage site and for A 10 nt downstream

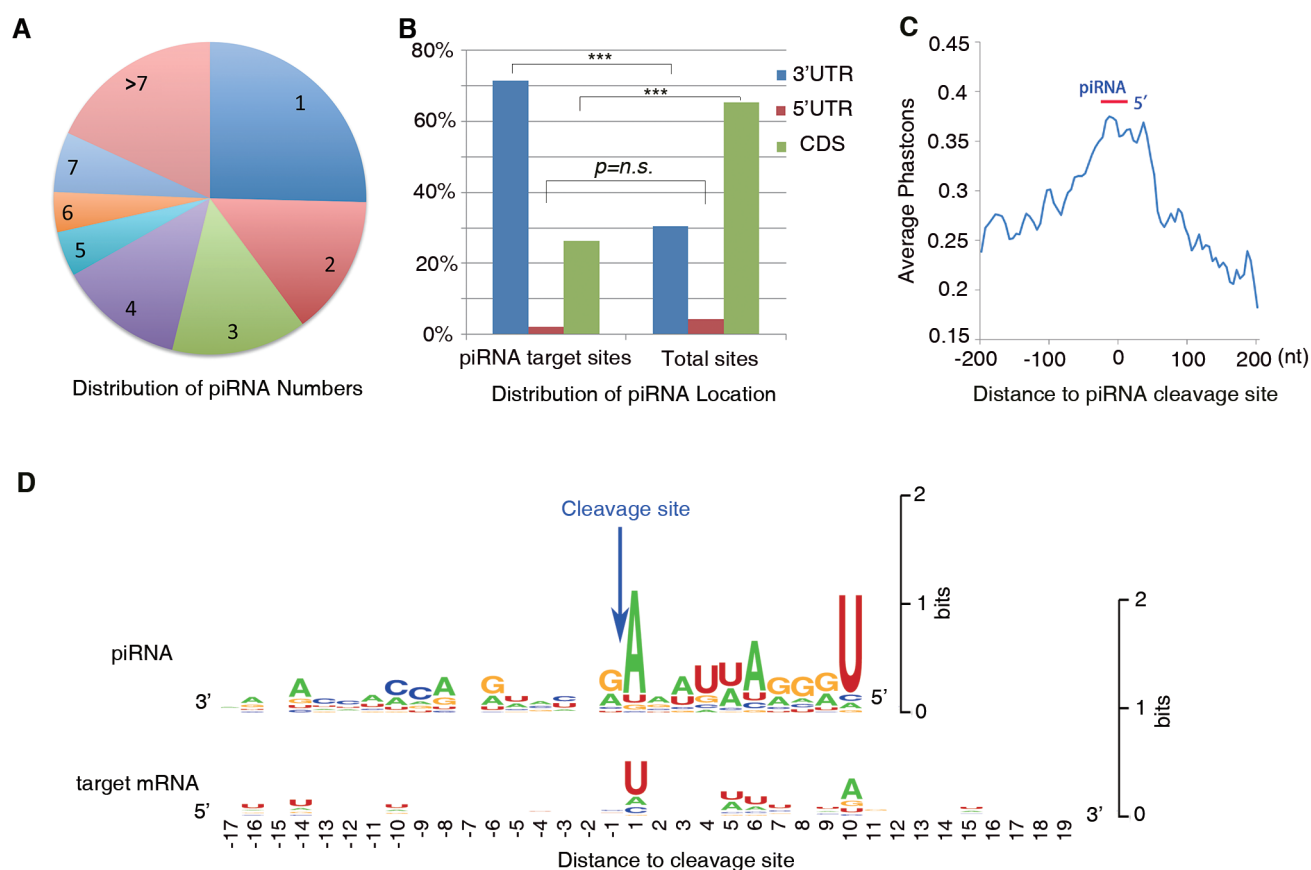


Figure 5 Features of piRNA target sites and corresponding targeting piRNAs. **(A)** Grouping the 193 predicted piRNA target sites according to the number of their corresponding targeting piRNAs. **(B)** The distribution of the 193 predicted piRNA target sites in mRNAs in comparison with that of all mRNA fragments in the 5'-RACE library. $***P < 0.001$ (Fisher's exact test). **(C)** Conservation of non-redundant piRNA target sites among placental mammals. **(D)** Sequence logo of mRNA-targeting piRNAs and their target sites in mRNAs. The logos were generated by the WebLogo web server [42] using sequences of 1 295 non-redundant piRNAs and 193 target sites in mRNAs. The blue arrow marks the mRNA cleavage site between positions 10 and 11 in piRNAs. Numbers under the target mRNA logo represent the distances from the cleavage site.

of the U residue (Figure 5D, bottom). Finally, we found a few piRNAs corresponding to the predicted target sites in mRNAs. Together, these observations indicate that a 'ping-pong'-like mechanism may be responsible for generation of piRNAs from their target mRNAs.

Temporal regulation of piRNA target genes during mouse spermatogenesis

In *Miwi*^{-ADH} and *Miwi*^{-/-} mice, germ cell development was uniformly blocked at the round spermatid stage [17], suggesting that the MIWI slicer activity-mediated target repression may represent a critical mechanism in regulating spermatogenesis. In support of this notion, we found a significant increase in the mRNA levels of 169 predicted piRNA target genes in *Miwi*^{-ADH} round spermatids compared to *Miwi*^{+/-} cells (Figure 2D and Supplementary

information, Figure S2B). Also, mRNA levels of several piRNA target genes were significantly decreased in round spermatids compared with those in spermatocytes (Supplementary information, Figure S5A). Moreover, global expression profiling analysis showed that the mRNA levels of 169 predicted piRNA target genes in elongating spermatids (GSE59291) were significantly lower than those in round spermatids (GSE32180; Supplementary information, Figure S8), suggesting a temporal regulation of piRNA target genes during mouse spermatogenesis.

Functional importance of temporal regulation of piRNA target genes for spermiogenesis in mice

To probe the functional importance of temporal regulation of piRNA target genes, we first analyzed annotated functions for the 169 predicted target genes using DA-

VID [28]. Interestingly, we found that these genes are enriched in categories of DNA damage and repair, vacuole, zinc binding, cell aging and ATP binding (Supplementary information, Table S2). Importantly, 14 out of these 169 predicted piRNA target genes have been shown to be essential for spermatogenesis in mouse models (Supplementary information, Table S3). These results suggest that silencing of genes by MIWI/piRNAs is functionally linked to male germ cell development.

To experimentally prove the biological requirement for proper regulation of piRNA targets during mouse spermatogenesis, we further characterized two specific piRNA target genes, *Rad1* and *Xrcc2*, which are known to play critical roles in DNA recombination and repair [29-30]. We first validated the predicted piRNA:mRNA interactions using a reporter-based assay, confirming *Rad1* and *Xrcc2* as authentic piRNA targets (Figure 6A and Supplementary information, Figure S9A). We next demonstrated in wild-type mice that *Rad1* and *Xrcc2* protein levels were greatly reduced in round spermatids compared to those in spermatocytes (Figure 6B), indicating a temporal regulation of these two piRNA targets during mouse spermatogenesis. Conversely, the expression of *Rad1* and *Xrcc2* was upregulated in either *Miwi*^{-/-} or *Miwi*^{-ADH} round spermatids (Figure 6C, 6D and Supplementary information, Figure S9B), suggesting that the slicer activity of MIWI is required for downregulation of these genes in mouse spermatids.

To determine the functional requirement of regulated expression of *Rad1* and *Xrcc2* for the formation of sperms in mice, we respectively cloned their cDNAs into a lentiviral vector upstream of the IRES-GFP (Supplementary information, Figure S10A), and used the spermatid-specific *protamine 1* promoter (*Prm*) to drive their expression in spermatids through testis transduction. To confirm the expression of transduced *Rad1* or *Xrcc2*, we isolated haploid spermatids from the transduced testes using the unit gravity sedimentation procedure as we described previously [21] and by western blotting, we found that the *Rad1* or *Xrcc2* protein levels were significantly increased in spermatids isolated from the expression vector-transduced testes compared to control vector-transduced testes (Supplementary information, Figure S10B; *n* = 5 in each group), demonstrating successful lentiviral transduction as well as spermatid expression of *Rad1* or *Xrcc2* protein in transduced mice.

We next scored the effects of lentiviral transduced *Rad* or *Xrcc2* on the formation of sperms in epididymides by PCR as we described previously [21]. We reasoned that if the transduced GFP-containing vector does not affect normal sperm maturation, GFP DNA should be readily detectable in genomic DNA from epididymal sperma-

tozoa by PCR. Consistent with our previous observation [21], GFP DNA was readily detected in genomic DNA from epididymal spermatozoa in five out of five mice transduced with control vector (Figure 6E, lanes 4-8; *n* = 5). In sharp contrast, we could not detect any GFP DNA signal in sperm genomic DNA prepared from mice transduced with either *Rad1* or *Xrcc2* vector (Figure 6E, lanes 10-20; *n* = 5), indicating that spermatids expressing exogenous piRNA-resistant *Rad1* or *Xrcc2* failed to progress to mature sperm. Collectively, these results suggest that the temporal regulation of piRNA targets in spermatids is essential for subsequent sperm formation.

Discussion

The role of piRNAs in transposon silencing and epigenetic regulation is well established. Recent studies show that some piRNAs in *Drosophila* appear to have the capacity to regulate the stability of several mRNAs [31]. However, in light of two recent studies suggesting that MIWI may bind and translationally repress spermatogenic mRNAs in a piRNA-independent manner in mouse testes [27, 32], it has been unclear whether mammalian piRNAs may participate in post-transcriptional regulation of protein-coding genes and, if so, whether this represents another major mode of piRNA actions. Based on our computational analyses, genome-wide mapping of MIWI-RNA interactions, and reporter-based assays, we present a series of findings that demonstrate the function of a subset of piRNAs in the regulation of mRNA stability in mouse testes by a mechanism that requires the slicer activity of the MIWI protein.

First, we identified potential MIWI-cleaved mRNA fragments in a published 5'-RACE library [17]. We also deduced that potential piRNAs may be target specific sites in these mRNAs based on base pairing from nucleotides 2 to 21 in the piRNAs. Consistent with the action site of the MIWI slicer activity on its targets [17], we found that the 5' ends of the predicted targeting piRNAs frequently match position 10 from the 5' ends of mRNA fragments. This correlation was not observed in libraries from testes of MIWI catalytic mutant mice, and the expression levels of the predicted piRNA target genes were significantly upregulated in the mutant testes. Importantly, our MIWI CLIP-seq data indicate that the MIWI protein is physically associated with the predicted piRNA target sites in mRNAs. Our reporter assays further demonstrated that piRNAs could induce downregulation of their targets, and both the slicer activity and piRNA-loading ability of MIWI are required for such piRNA-guided target repression. Taken together, these data strongly suggest that piRNA-guided mRNA

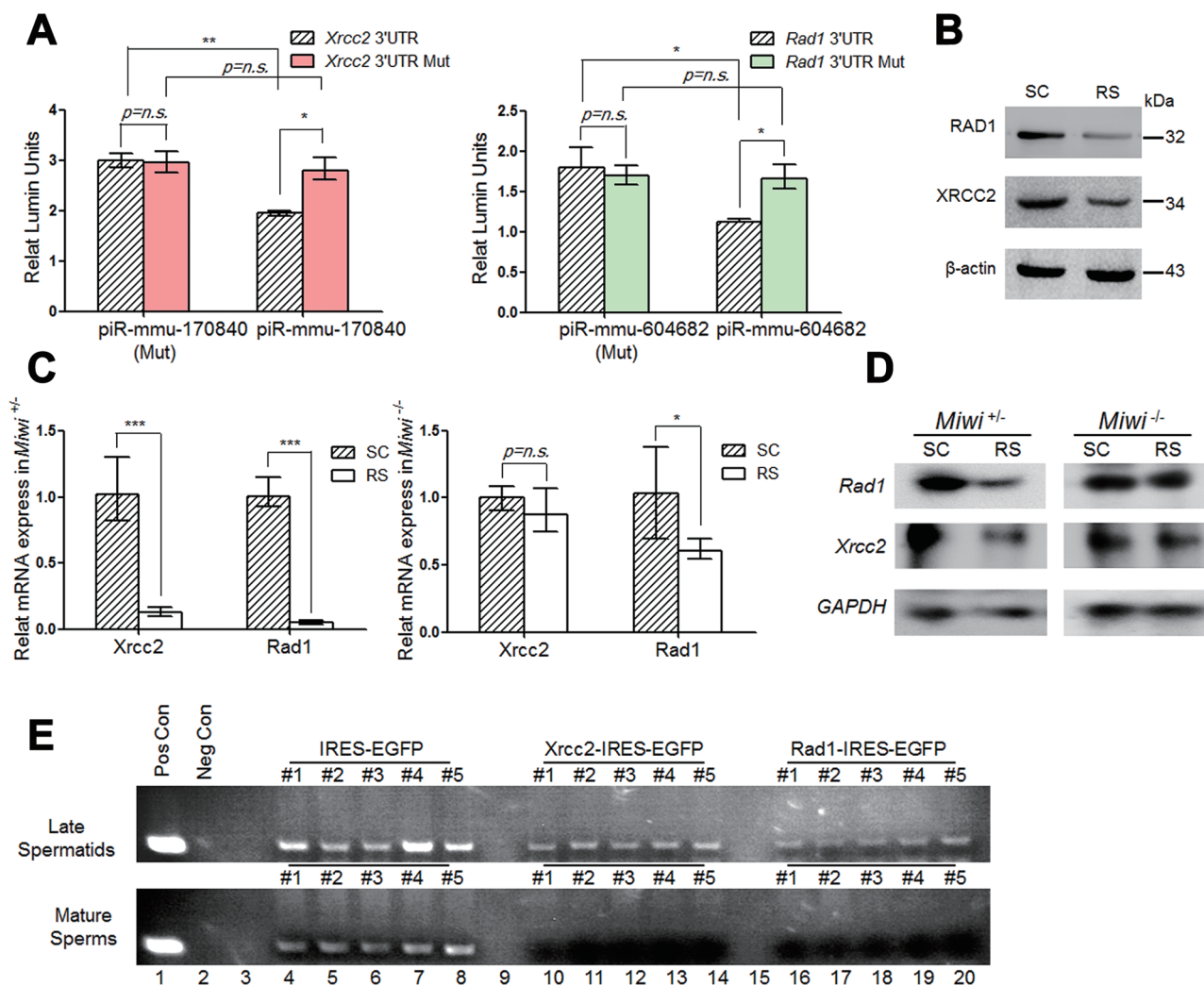


Figure 6 Temporal regulation of the piRNA target genes *Xrcc2* and *Rad1* is essential for sperm formation. **(A)** Dual-luciferase reporter assays to determine the effects of piR-mmu-170840 [24] on the *Xrcc2* reporter (left), and piR-mmu-604682 [24] on the *Rad1* reporter (right) in transfected GC-2spd(ts) cells. A mutant piRNA (Mut) served as negative control in each case. **(B)** Western blot analyses of *Xrcc2* and *Rad1* expression in enriched SC and RS, with β -actin serving as the loading control. **(C, D)** qRT-PCR **(C)** or northern blot **(D)** analyses of the mRNA levels of *Xrcc2* and *Rad1* in *Miwi*^{+/-} and *Miwi*^{-/-} mouse spermatogenic cells, with *GAPDH* serving as an internal control. Mean \pm SEM of three separate experiments were plotted. Student's *t*-test significance: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. **(E)** PCR analyses of GFP DNA in extracted genomic DNA from late spermatids (top) and epididymal sperms (bottom) from mouse testes transduced with IRES-EGFP (lanes 4-8), *Xrcc2*-IRES-EGFP (lanes 10-14), or *Rad1*-IRES-EGFP (lanes 16-20), respectively. Genomic DNA samples from Oct4-GFP transgenic mice (lane 1) and wild-type background mice (lane 2) served as positive and negative controls, respectively. Results shown are representative of three independent experiments.

degradation is a predominant mode of MIWI/HIWI functions in mammals. Considering that cleaved mRNA fragments are rapidly degraded by other cellular nucleases, we predict that the piRNA system may play widespread roles in mediating mRNA degradation in male germ cells.

Remarkably, we found that base pairing at the first

base of piRNAs is important for piRNA-mediated mRNA cleavage *in vivo*, based on our computational analyses of mRNA cleavage sites in mouse testicular 5'-RACE cleavage fragment library [17]. Consistently, our observation of 1U and 10A signatures at the cleavage sites suggests a ping-pong amplification-like mechanism for piRNA action on mRNA, further supporting the importance of the

first base pairing in this process. In contrast, based on the structure of AGO silencing complexes, the first base of guide RNAs interacts with the Mid domain and does not form duplex with target RNAs [33], suggesting that the first base pairing of guide RNAs to targets is dispensable for AGO/siRNA-mediated gene silencing. Moreover, while previous biochemical studies showed that base pairing beyond position 16 in AGO silencing complexes appears to be dispensable [33], base pairing at position 20 and beyond still significantly affects MIWI-mediated cleavage of target RNAs [17]. Future studies of the crystal structure of MIWI silencing complexes will provide critical information to understand the molecular mechanism underlying piRNA-mediated gene silencing.

We note that another mouse Piwi protein MILI has also been shown to possess slicer activity critical for post-transcriptional silencing of retrotransposon LINE1 in spermatocytes [18, 34]. However, our results demonstrate that the expression levels of the predicted piRNA target genes are significantly upregulated in *Miwi*^{-ADH} testes (Figure 2D), in which *Mili* is normally expressed. This finding implies that MILI may not be involved in the regulation of this group of piRNA targets identified in our study. Nevertheless, it remains to be determined whether MILI plays a role in post-transcriptional regulation of other coding genes in mouse testes.

In summary, our present study combined computational and experimental approaches to deduce rules for piRNA:mRNA interactions through analyzing sequence complementarity, location of target sites, as well as piRNA reads that match their target sites. Even though the rules that we deduced are subject to further refinement, our data suggest that the productive piRNA:mRNA interactions require (1) a minimal ~16-nt continuous complementarity between piRNAs and mRNAs, (2) strict base pairing at the 5' end of piRNAs, (3) < 3 mismatches in subsequent 20 nt sequences, (4) conserved 3'-UTRs as preferred piRNA target sites, and (5) U and A residues enriched at the 1st and 10th positions in piRNAs that target specific mRNAs. Given that millions of piRNAs are expressed in mouse testes and their roles in silencing transposons have been well-documented, it will be of significance in future studies to determine which portion of the piRNA machinery is devoted to protecting the genome against transposable elements and which portion is involved in the regulation of protein-coding genes, as well as these two processes is coordinated in different stages of male germ cell development in mammals.

Materials and Methods

Plasmids, RNA oligonucleotides, qPCR primers and anti-

bodies

For reporter pmirGLO-*Psm8* 3'-UTR, the 860-bp mouse *Psm8* 3'-UTR was cloned into pmirGLO vector (Promega) downstream of the *Firefly* luciferase gene. Eight nucleotides in *Psm8* 3'-UTR complementary to the 5' portion of piR-671 were deleted using KOD-Plus-mutagenesis kit (Toyobo) to generate reporter *Psm8* 3'-UTR Mut. The wild-type and mutant 3'-UTR reporters for other target candidates, including *Tox4*, *BC026590*, *Xrcc2*, and *Rad1*, were similarly constructed (Figures 4B, 4C, 6A and Supplementary information, Figures S5B-S5D and S9A). p3×Flag-MIWI, p3×Flag-HIWI, and p3×Flag-HIWI mutants were constructed as recently described [20]. pLV-*Prm*-RAD1-IRES-GFP and pLV-*Prm*-XRCC2-IRES-GFP were derived from vector Lenti-EF1α-Oct4-IRES-EGFP (SIDANSAI) as previously described [21], in which the *Oct4* promoter was replaced by the 0.6 kb of the *protamine1* promoter (*Prm*) sequences (Supplementary information, Figure S10A). All constructs were confirmed by Sanger sequencing. All RNA oligonucleotides, including piRNAs and siRNAs, were synthesized at Ribobio (Guangzhou, China), and the sequences are provided in Supplementary information, Table S4. Primer sequences for RT-qPCR are shown in Supplementary information, Table S5.

Mouse monoclonal anti-β-actin (A3854) antibody was purchased from Sigma, and anti-XRCC2 (sc-5895) was from Santa Cruz. Rabbit anti-Rad1 polyclonal antibodies (A1047) was from ABclonal Technology and anti-GFP (MBL-598) was from MBL.

Data sources

The global 5'-RACE cleavage fragment data, gene expression profile, and MIWI-immunoprecipitated piRNA data in the control and catalytic mutant mouse testes were obtained from the SuperSeries in GEO: GSE32183 [17]. We used the processed sequencing data generated from adult mice. Mouse genome data (mm9) and RefSeq mRNA data were downloaded from UCSC [35-36].

Sequence mapping

Global 5'-RACE sequences were mapped to the mouse genome and RefSeq mRNAs using bowtie [37] with no mismatches. When the 5'-RACE tags were mapped to mRNA isoforms of the same gene, only one mapping result was kept for further analyses. For piRNA target mapping, when applying the rule that 1-3 mismatches were allowed in piRNA positions 2-21, we extracted sequences from positions 2 to 21 of piRNA and aligned them to mRNAs using bowtie, allowing up to 3 mismatches. Otherwise, we used blastn with word size -W 8 and output format -m 1 and then summarized the details of the matches, mismatches and gaps using custom Perl scripts. G-U wobble base pairs were regarded as mismatches in our analyses.

Distribution of the number of potential piRNAs binding in the vicinity of mRNA fragments

For each cleaved mRNA fragment, we counted the number of unique piRNAs that could potentially bind to the mRNA at different distances from the 5' end of the mRNA fragment to the 5' end of the piRNA.

We summed up the number of the piRNAs that could potentially bind at each position relative to all mRNA fragments. The piRNA data were obtained from controls (GSM822760 *Miwi*^{+/+}) or MIWI catalytic mutants (GSM822762-4 *Miwi*^{-ADH}, 3 datasets).

The mRNA fragment data were also extracted from controls (GSM822765 *Miwi*^{+/+}) or MIWI catalytic mutants (GSM822766 *Miwi*^{-/ADH}) [17].

Analysis of public microarray data

Raw data (CEL file format) for 8 samples of round spermatids from adult mouse testes in GSE32180 [17] were loaded into R [38] using the affy package, normalized, and converted to log₂ gene expression by rma using the “mogene10stv1mmentrezgcdf” cdf package downloaded from customCDF (version 15) on the brainarray website [39]. The raw data for 6 samples of elongating spermatids (GSE59291) (Microarray performed in Shanghai Biotechnology Corporation) [20] were processed in the same way using the “mouse4302mmentrezgcdf” cdf package. For the comparison of gene expression levels between these two platforms, expression levels were further normalized by the “normalizeBetweenArrays” tool in limma package.

GSEA

Probe level expression of the GSE32180 microarray data was calculated using rma in R, and all MIWI catalytic mutant (*Miwi*^{-/ADH}) and control (*Miwi*^{+/+}) samples were extracted and converted to gct format for GSEA. MoGene_1_0_st.chip was used as platform annotation. Probe level expressions were collapsed to the gene level according to the gene annotation. A new set of potential piRNA target genes was created. Gene symbols were modified manually to fit the platform annotation. For permutation type, we chose gene_set according to the GSEA suggestion. Other parameters were set as default.

MIWI CLIP-seq and data processing

MIWI CLIP-seq was performed as described [40]. In brief, isolated round spermatids from adult C57BL/6J mice were UV-irradiated (254 nm) at 400 mj in a 15-cm plate before immunoprecipitation using a highly specific anti-MIWI antibody. Immunoprecipitated RNA-protein complexes were digested with micrococcal nuclease, labeled with [γ-³²P]-ATP (PerkinElmer) by T4 PNK (Fermentas), and isolated by SDS-PAGE, from which [³²P]-labeled RNA-protein bands were cut for extracting RNAs for linker ligation, PCR amplification, and deep sequencing.

For each read, the first 4 random index sequences were removed and appended to the read name. The resulting read was 36 nt in length. The 3'-adaptor sequence (CTCGTATGCCGCTCTTCTGCTTG) was trimmed and short reads (< 16 nt) were filtered out. Reads were then mapped to the mouse genome (mm9) using bowtie [37] with parameters: “-l25 -n2 -k101 -m100 -e200 --best --strata --sam --phred33-quals”. Mapped reads with ≤ 1 mismatch were chosen for downstream analysis. Reads with lengths of 25-33 nt were classified as piRNAs, and longer reads are considered as MIWI targets.

Cell culture and transfection

GC-2spd(ts) cells were obtained from the American Type Culture Collection (ATCC) and cultured with the medium and serum as ATCC recommended. Transfection was performed using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. For transfection of the RNA oligonucleotides, 100 nM of piRNA or scrambled siRNA oligonucleotides were used. For plasmid transfection, 4 μg DNA was used in a 6-well plate.

Luciferase reporter assay

Luciferase reporter assays were carried out as described previously [20]. In brief, each 3'-UTR luciferase reporter construct was cotransfected into GC-2spd(ts) cells in 24-well plates with RNA oligonucleotides or vectors as indicated in Figure 4 and Supplementary information, Figure S5. piRNA negative controls were designed as the mutated forms of their cognate piRNAs (designated as piR-xxx(Mut)) by altering the first 8-nt sequences at the 5' end to disrupt base pairing between piRNAs and target mRNAs. The sequences of piRNAs and their corresponding mutants are provided in Supplementary information, Figures S5B and S9A. All experiments were triplicated and *P* values were calculated by Student's *t*-test.

Northern blot

Northern blot was carried out according to standard protocols. In brief, 15 μg of total RNAs were loaded per lane on 1% formaldehyde agarose gel. After electrophoresis, samples were transferred to nylon membrane (Hybond N+, GE, USA) and probed with a ~1-kb ORF fragment of target genes that were labeled with [α-³²P]dCTP.

Ectopic expression of *Rad1* or *Xrcc2* in mouse spermatids

Lentivirus packaging and testis transduction were performed as described previously [21]. In brief, under a microscope, testes of pentobarbital sodium-anesthetized mice were pulled out and ~10 μl of high-titer lentivirus were injected into each seminiferous tubule through efferent duct by a sharp glass capillary with a tip diameter of 50 μm. The testes were then returned to the abdominal cavity, and the abdominal wall and skin were closed with sutures. The testes were harvested 4-5 weeks after transduction of ectopic *Rad1* or *Xrcc2* expression vectors.

Isolation of mouse spermatogenic cells and sperms

Mouse total germ cells were extracted from the seminiferous tubules of adult male mice, and spermatocytes and haploid spermatids were isolated using a unit gravity sedimentation procedure as described previously [21]. Epididymal sperm isolation was carried out following standard procedures. In brief, epididymides were removed immediately after male mice were sacrificed, and placed into sterile Petri dishes containing ice-cold PBS. By gently shearing the tissues and rupturing the tubules, sperms were isolated from epididymides and diffused into the medium. Finally, spermatozoa were obtained by gently filtering on ice through fine nylon mesh to remove tissue debris.

Conservation analysis

All potential piRNA target sites in the mouse genome were located. Duplicate sites located too close to one another (within 10 nt) were removed manually. The non-redundant sites obtained were uploaded to cistrome [41], and conservation analysis was performed using Conservation Plot in Integrative Analysis. Since Conservation Plot does not consider strand information, sites belonging to the plus and minus strands of individual chromosomes were separated before conservation analysis, and then the minus strand results were reversed and averaged with the plus strand results.

Availability of supporting data

The names of piRNA sequences used in the reporter assays, such as piR-mmu-77755, are available at piRBase [24] (<http://www.regulatoryrna.org/database/piRNA/>). The additional files and information are available at: http://www.regulatoryrna.org/pub/miwi_cleave/index.html.

Acknowledgments

We are grateful to Xiaole Shirley Liu and Jianhua Yang for thoughtful discussions and valuable comments on the manuscript. This work was supported by grants from the Ministry of Science and Technology of China (2011CB504605 and 2014AA021103), the National Natural Science Foundation of China (31090253, 31210103912, 31325008 and 91419307), the Chinese Academy of Sciences (CAS; KJZD-EW-L01-2), the Key Laboratory of the Zoological Systematics and Evolution of the Chinese Academy of Sciences (O529YX5105), and the National Science Fund for Fostering Talents in Basic Research (Special subjects in animal taxonomy, NSFC-J1210002). This work was also supported by NIH grants (GM052872 and HG004659). The computing resource was supported by HPC Platform, Scientific Information Center, Institute of Zoology, CAS.

References

- Thomson T, Lin H. The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. *Annu Rev Cell Dev Biol* 2009; **25**:355-376.
- Girard A, Sachidanandam R, Hannon GJ, Carmell MA. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* 2006; **442**:199-202.
- Aravin A, Gaidatzis D, Pfeffer S, *et al.* A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* 2006; **442**:203-207.
- Lau NC, Seto AG, Kim J, *et al.* Characterization of the piRNA complex from rat testes. *Science* 2006; **313**:363-367.
- Grivna ST, Beyret E, Wang Z, Lin H. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev* 2006; **20**:1709-1714.
- Siomi MC, Sato K, Pezic D, Aravin AA. PIWI-interacting small RNAs: the vanguard of genome defence. *Nat Rev Mol Cell Biol* 2011; **12**:246-258.
- Brennecke J, Aravin AA, Stark A, *et al.* Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 2007; **128**:1089-1103.
- Gunawardane LS, Saito K, Nishida KM, *et al.* A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* 2007; **315**:1587-1590.
- Aravin AA, Naumova NM, Tulin AV, Vagin VV, Rozovsky YM, Gvozdev VA. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr Biol* 2001; **11**:1017-1027.
- Nishida KM, Saito K, Mori T, *et al.* Gene silencing mechanisms mediated by Aubergine piRNA complexes in *Drosophila* male gonad. *RNA* 2007; **13**:1911-1922.
- Saito K, Inagaki S, Mituyama T, *et al.* A regulatory circuit for piwi by the large Maf gene traffic jam in *Drosophila*. *Nature* 2009; **461**:1296-1299.
- Rouget C, Papin C, Boureux A, *et al.* Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo. *Nature* 2010; **467**:1128-1132.
- Juliano C, Wang J, Lin H. Uniting germline and stem cells: the function of Piwi proteins and the piRNA pathway in diverse organisms. *Annu Rev Genet* 2011; **45**:447-469.
- Aravin AA, Sachidanandam R, Bourc'his D, *et al.* A piRNA pathway primed by individual transposons is linked to *de novo* DNA methylation in mice. *Mol Cell* 2008; **31**:785-799.
- Kuramochi-Miyagawa S, Watanabe T, Gotoh K, *et al.* DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev* 2008; **22**:908-917.
- Watanabe T, Tomizawa S, Mitsuya K, *et al.* Role for piRNAs and noncoding RNA in *de novo* DNA methylation of the imprinted mouse Rasgrf1 locus. *Science* 2011; **332**:848-852.
- Reuter M, Berninger P, Chuma S, *et al.* Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. *Nature* 2011; **480**:264-267.
- De Fazio S, Bartonicek N, Di Giacomo M, *et al.* The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature* 2011; **480**:259-263.
- Subramanian A, Tamayo P, Mootha VK, *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005; **102**:15545-15550.
- Gou LT, Dai P, Yang JH, *et al.* Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Res* 2014; **24**:680-700.
- Zhao S, Gou LT, Zhang M, *et al.* piRNA-Triggered MIWI ubiquitination and removal by APC/C in late spermatogenesis. *Dev Cell* 2013; **24**:13-25.
- Deng W, Lin H. Miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Dev Cell* 2002; **2**:819-830.
- Hofmann MC, Abramian D, Millan JL. A haploid and a diploid cell coexist in an *in vitro* immortalized spermatogenic cell line. *Dev Genet* 1995; **16**:119-127.
- Zhang P, Si X, Skogerboe G, *et al.* piRBase: a web resource assisting piRNA functional study. *Database (Oxford)* 2014 Nov 25. doi: 10.1093/database/bau110
- Siepel A, Bejerano G, Pedersen JS, *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 2005; **15**:1034-1050.
- Beyret E, Liu N, Lin H. piRNA biogenesis during adult spermatogenesis in mice is independent of the ping-pong mechanism. *Cell Res* 2012; **22**:1429-1439.
- Vourekas A, Zheng Q, Alexiou P, *et al.* Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. *Nat Struct Mol Biol* 2012; **19**:773-781.
- Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009; **4**:44-57.
- Deans B, Griffin CS, O'Regan P, Jasim M, Thacker J. Homologous recombination deficiency leads to profound genetic instability in cells derived from Xrcc2-knockout mice. *Cancer Res* 2003; **63**:8181-8187.
- Lyndaker AM, Lim PX, Mleczko JM, *et al.* Conditional in-

- activation of the DNA damage response gene Hus1 in mouse testis reveals separable roles for components of the RAD9-RAD1-HUS1 complex in meiotic chromosome maintenance. *PLoS Genet* 2013; **9**:e1003320.
- 31 Peng JC, Lin H. Beyond transposons: the epigenetic and somatic functions of the Piwi-piRNA mechanism. *Curr Opin Cell Biol* 2013; **25**:190-194.
- 32 Nishibu T, Hayashida Y, Tani S, *et al.* Identification of MIWI-associated Poly(A) RNAs by immunoprecipitation with an anti-MIWI monoclonal antibody. *Biosci Trends* 2012; **6**:248-261.
- 33 Wang Y, Juranek S, Li H, *et al.* Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* 2009; **461**:754-761.
- 34 Di Giacomo M, Comazzetto S, Saini H, *et al.* Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. *Mol Cell* 2013; **50**:601-608.
- 35 Waterston RH, Lindblad-Toh K, Birney E, *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002; **420**:520-562.
- 36 Karolchik D, Hinrichs AS, Furey TS, *et al.* The UCSC Table Browser data retrieval tool. *Nucleic Acids Res* 2004; **32**:D493-D496.
- 37 Langmead B. Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics* 2010; **Chapter 11**:Unit 11.7.
- 38 R_Development_Core_Team. R: A language and environment for statistical computing 2009.
- 39 Dai M, Wang P, Boyd AD, *et al.* Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 2005; **33**:e175.
- 40 Xue Y, Ouyang K, Huang J, *et al.* Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. *Cell* 2013; **152**:82-96.
- 41 Liu T, Ortiz JA, Taing L, *et al.* Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biol* 2011; **12**:R83.
- 42 Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res* 2004; **14**:1188-1190.

(Supplementary information is linked to the online version of the paper on the *Cell Research* website.)



This work is licensed under the Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0>