Molecular Cell
Article

# An Ancient Transcription Factor Initiates the Burst of piRNA Production during Early Meiosis in Mouse Testes

Xin Zhiguo Li,[1,4] Christian K. Roy,[1,4] Xianjun Dong,[2] Ewelina Bolcun-Filas,[3] Jie Wang,[2] Bo W. Han,[1] Jia Xu,[2] Melissa J. Moore,[1] John C. Schimenti,[3] Zhiping Weng,[2,*] and Phillip D. Zamore[1,*]
[1]Department of Biochemistry and Molecular Pharmacology, Howard Hughes Medical Institute
[2]Program in Bioinformatics and Integrative Biology
University of Massachusetts Medical School, Worcester, MA 01605, USA
[3]Department of Biomedical Sciences and Center for Vertebrate Genomics, Cornell University College of Veterinary Medicine, Ithaca, NY 14850, USA
[4]These authors contributed equally to this work
*Correspondence: zhiping.weng@umassmed.edu (Z.W.), phillip.zamore@umassmed.edu (P.D.Z.)
http://dx.doi.org/10.1016/j.molcel.2013.02.016

## SUMMARY

**Animal germ cells produce PIWI-interacting RNAs (piRNAs), small silencing RNAs that suppress transposons and enable gamete maturation. Mammalian transposon-silencing piRNAs accumulate early in spermatogenesis, whereas pachytene piRNAs are produced later during postnatal spermatogenesis and account for >95% of all piRNAs in the adult mouse testis. Mutants defective for pachytene piRNA pathway proteins fail to produce mature sperm, but neither the piRNA precursor transcripts nor the trigger for pachytene piRNA production is known. Here, we show that the transcription factor A-MYB initiates pachytene piRNA production. A-MYB drives transcription of both pachytene piRNA precursor RNAs and the mRNAs for core piRNA biogenesis factors including MIWI, the protein through which pachytene piRNAs function. A-MYB regulation of piRNA pathway proteins and piRNA genes creates a coherent feedforward loop that ensures the robust accumulation of pachytene piRNAs. This regulatory circuit, which can be detected in rooster testes, likely predates the divergence of birds and mammals.**

## INTRODUCTION

P-element induced wimpy testis (PIWI)-interacting RNAs (piRNAs) can be distinguished from other animal small silencing RNAs by their longer length (typically 23–35 nt), 2′-*O*-methyl-modified 3′ termini, and association with PIWI proteins, a distinct subgroup of Argonaute proteins, the small RNA-guided proteins responsible for RNA interference and related pathways (Kumar and Carmichael, 1998; Aravin and Hannon, 2008; Farazi et al., 2008; Kim et al., 2009; Thomson and Lin, 2009; Cenik and Zamore, 2011). piRNA production does not require Dicer, the double-stranded RNA endonuclease that makes microRNAs (miRNAs) and small interfering RNAs (siRNAs), and piRNAs are thought to derive from single-stranded rather than double-stranded RNA (Vagin et al., 2006; Houwing et al., 2007).

In most bilateral animals, germline piRNAs protect the genome from transposon activation, but also have other functions (Aravin et al., 2001, 2007, 2008; Vagin et al., 2004, 2006; Brennecke et al., 2007; Carmell et al., 2007; Hartig et al., 2007; Kuramochi-Miyagawa et al., 2008; Ashe et al., 2012; Lee et al., 2012; Shirayama et al., 2012). A few days after birth, the majority of piRNAs in the mouse testis are pre-pachytene piRNAs; 25% of these piRNA species map to more than one location in the genome. A second class of piRNAs, typically derived from intergenic regions, has been reported to emerge in the mouse testis 14.5 days postpartum (dpp), when the developing spermatocytes synchronously enter the pachytene phase of meiotic prophase I. These pachytene piRNAs compose >95% of piRNAs in the adult mouse testis. Loss of genes required to make pachytene piRNAs blocks production of mature sperm (Deng and Lin, 2002; Aravin and Hannon, 2008; Reuter et al., 2011; Vourekas et al., 2012). What triggers the accumulation of pachytene piRNAs when spermatocytes enter the pachynema is unknown.

In *Caenorhabditis elegans*, each piRNA is processed from its own short RNA polymerase II (Pol II) transcript (Gu et al., 2012). In contrast, insect and mouse piRNAs are thought to be processed from long RNAs transcribed from large piRNA loci. Supporting this view, a transposon inserted into the 5′ end of the *flamenco* piRNA cluster in flies reduces the production of *flamenco* piRNAs 168 kbp 3′ to the insertion, suggesting that it disrupts transcription of the entire locus (Brennecke et al., 2007). High-throughput sequencing and chromatin immunoprecipitation (ChIP) has been used to define the genomic structure of the piRNA-producing genes of immortalized, cultured silk moth BmN4 cells (Kawaoka et al., 2013). However, for flies and mice, we do not know the structure of piRNA-producing genes, their transcripts, or the nature of the promoters that control their expression.

Instead, piRNA loci have been defined as clusters: regions of the genome with a high density of mapping piRNA sequences (Aravin et al., 2006, 2007; Girard et al., 2006; Grivna et al.,

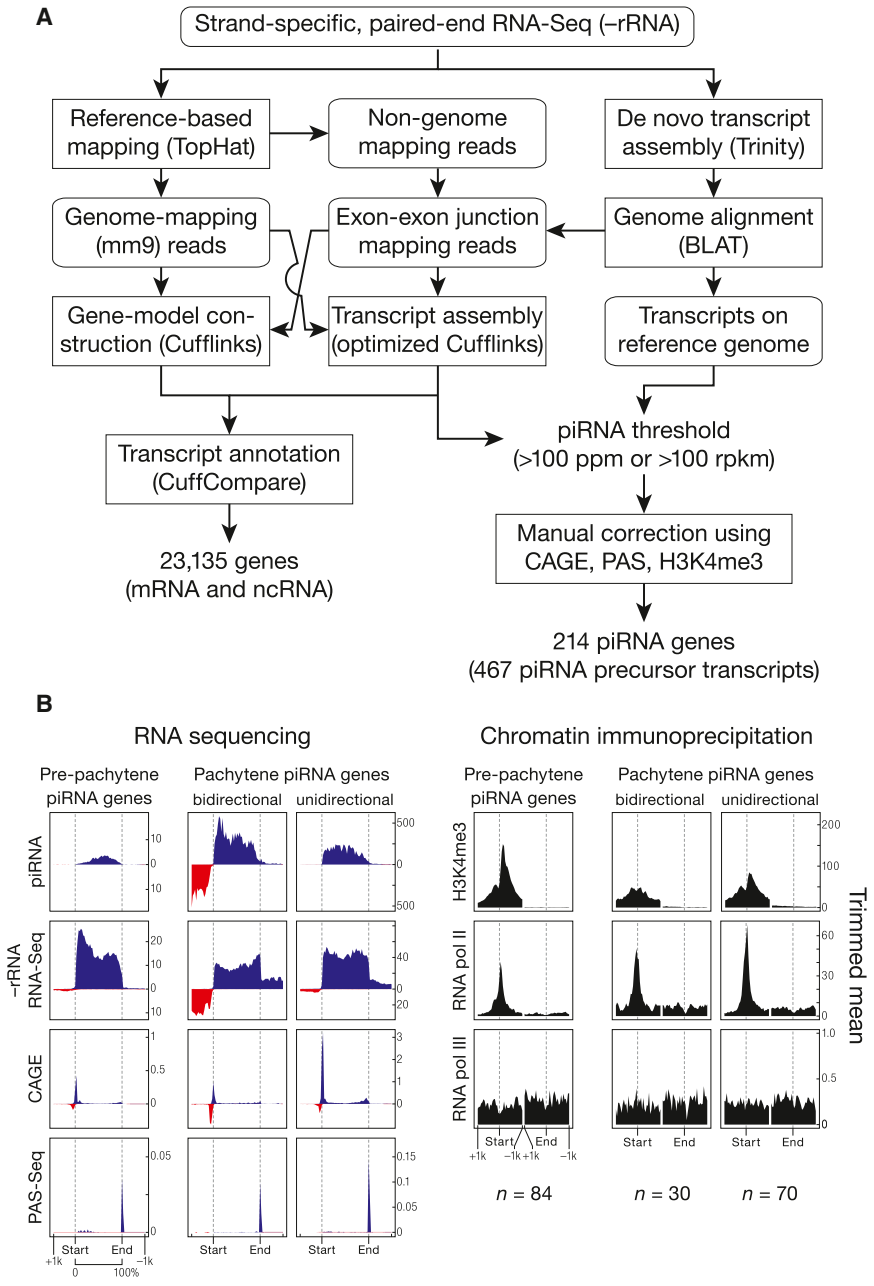


**Figure 1. piRNA Precursors Are RNA Pol II Transcripts**

(A) Strategy to assemble the mouse testis transcriptome. Rectangles with rounded corners, input or output data; rectangles, processes. Decisions are shown without boxing.

(B) Aggregated data for piRNA-producing transcripts (5% trimmed mean). Oxidized small RNA (>23 nt) sequencing data were used to detect piRNAs; transcript abundance was measured using total RNA depleted of rRNA (RNA-seq). RNA Pol III data were from SRA001030. Dotted lines show the transcriptional start site (Start) and site of polyadenylation (End). See also Figure S1 and Table S1.

2006a; Lau et al., 2006; Brennecke et al., 2007; Ro et al., 2007). In reality, piRNA-producing loci correspond to discrete transcription units that include both intergenic loci believed to encode no protein (Brennecke et al., 2007, 2008; Vourekas et al., 2012) and protein-coding genes that also produce piRNAs (Aravin et al., 2007; Robine et al., 2009; Saito et al., 2009).

We used high-throughput sequencing data to define the genes and transcripts that produce piRNAs in the juvenile and adult mouse testis. Using these data, we identified the factor that initiates transcription of pachytene piRNA genes: A-MYB (MYBL1), a spermatocyte protein that serves as a master regulator of genes encoding proteins required for cell-cycle progression through the pachytene stage of meiosis (Trauth et al., 1994; Bolcun-Filas et al., 2011). A-MYB also initiates transcription of the genes encoding many piRNA biogenesis factors. The combined action of A-MYB at the promoters of genes producing pachytene piRNA precursor transcripts and genes encoding piRNA biogenesis proteins creates a coherent feedforward loop that triggers a >6,000-fold increase in pachytene piRNA abundance during the ~5 days between the early and late phases of the pachytene stage of male meiosis. A-MYB also promotes its own transcription through a positive feedback loop. The A-MYB-regulated feedforward loop is evolutionarily conserved: A-MYB is bound to the promoters of both piRNA clusters and *PIWIL1*, *TDRD1*, and *TDRD3* in the rooster (*Gallus gallus*) testis.

## RESULTS

### Defining piRNA-Producing Transcripts in the Mouse Testis

To define the structure of piRNA-producing loci in the testis of wild-type adult mice, we assembled the transcripts detected by three biological replicates of strand-specific, paired-end, rRNA-depleted, total RNA sequencing (RNA-seq; Figure 1A). We mapped reads to the mouse genome using TopHat (Trapnell et al., 2009) and performed de novo transcriptome assembly using Trinity (Grabherr et al., 2011) to identify unannotated exon-exon junctions. We used all mapped reads, including reads corresponding to unannotated exon-exon junctions, to perform reference-based transcript assembly (Cufflinks; Trapnell et al., 2010).

To identify the transcripts that produce piRNAs, we sequenced piRNAs from six developmental stages of mouse testes (10.5 dpp, 12.5 dpp, 14.5 dpp, 17.5 dpp, 20.5 dpp, and adult) and mapped them to the assembled transcripts. The first round

of spermatogenesis proceeds synchronously among the tubules of the testis: mouse testes at 10.5 dpp advance no further than the zygotene stage (staging according to Nebel et al., 1961); 12.5 dpp to the early pachytene; 14.5 dpp to the middle pachytene; 17.5 to the late pachytene; and 20.5 dpp to the round spermatid stage. For each stage, we prepared two sequencing libraries: one comprising all small RNAs and one in which oxidation was used to enrich for piRNAs by virtue of their 2′-*O*-methyl-modified 3′ termini (Ghildiyal et al., 2008).

To qualify as a piRNA-producing transcript, an assembled RNA was required to produce either a sufficiently high piRNA abundance (>100 ppm; *p*arts *p*er *m*illion uniquely mapped reads) or density (>100 rpkm; *r*eads *p*er *k*ilobase of transcript per *m*illion uniquely mapped reads). These criteria retained both long transcripts producing an abundance of piRNAs and short transcripts generating many piRNAs per unit of length. To refine the termini of each piRNA-producing transcript, we supplemented the RNA-seq data with high-throughput sequencing of the 5′ ends of RNAs bearing an N(5′)ppp(5′)N cap structure (cap analysis of gene expression; CAGE) and the 3′ ends of transcripts preceding the poly(A) tail (polyadenylation site sequencing; PAS-seq). The assembled piRNA-producing transcripts likely correspond to continuous RNAs in vivo because the CAGE library used to annotate transcript 5′ ends was constructed after two rounds of poly(A) selection. Thus, the RNA molecules in the library derive from complete transcripts extending from the 5′ cap to the poly(A) tail (Figure 1B). Conventional 5′ and 3′ RACE (rapid amplification of cDNA ends) analysis of piRNA-producing transcripts confirmed the ends of 16 loci (data not shown). To provide additional confirmation of the 5′ end of each piRNA-producing transcript, we also determined the locations of histone H3 bearing trimethylated lysine 4 (H3K4me3), a histone modification associated with RNA Pol II transcription start sites (Guenther et al., 2007).

## piRNA Precursor RNAs Are Canonical RNA Pol II Transcripts

The presence of 5′ caps and poly(A) tails and the binding of histone H3K4me3 to the genomic DNA immediately upstream of the transcription start site of each piRNA locus suggest that piRNA transcripts are produced by RNA Pol II (Figure 1B). Moreover, using antibodies to RNA Pol II, but not RNA Pol III, ChIP-seq showed a peak at the transcription start site as well as polymerase occupancy across the entire piRNA gene (Figure 1B; Kutter et al., 2011). We conclude that piRNA transcripts are conventional RNA Pol II transcripts bearing 5′ caps and 3′ poly(A) tails.

## A Transcript-Based Set of piRNA Loci

Our transcriptome assembly yielded 467 piRNA-producing transcripts that define 214 genomic loci (Figure S1A and Table S1). Among the ~2.2 million distinct piRNA species and ~8.8 million piRNA reads from the adult mouse testis, the 214 genomic loci account for 95% of all piRNAs.

Previous studies defined piRNA clusters based solely on small RNA sequencing data (Girard et al., 2006; Lau et al., 2006; Aravin et al., 2007). Our approach differs in that it (1) uses RNA-seq data, whose greater read length facilitates the identification of introns, allowing us to define the architecture of piRNA precursor transcripts and (2) uses CAGE, PAS-seq, and H3K4me3 ChIP-seq data to refine the 5′ and 3′ ends of the piRNA transcripts. Consequently, the piRNA loci presented here account for more piRNAs using fewer genomic base pairs than those previously defined (Figures S1B and S1C; Lau et al., 2006; Girard et al., 2006). Our piRNA-producing loci include 41 piRNA loci that escaped previous detection (Girard et al., 2006; Lau et al., 2006; Aravin et al., 2007), 37 of which contain introns. The 41 loci account for 2% of piRNAs at 10.5 dpp and 0.36% in the adult testis.

## Three Classes of piRNAs during Postnatal Spermatogenesis

Mice produce three PIWI proteins: MIWI2 (PIWIL4), which binds piRNAs in perinatal testis (Carmell et al., 2007; Aravin et al., 2008); MILI (PIWIL2), which binds piRNAs at least until the round spermatid stage of spermatogenesis (Kuramochi-Miyagawa et al., 2004; Aravin et al., 2006, 2007); and MIWI (PIWIL1), which is first produced during the pachytene stage of meiosis (Deng and Lin, 2002). From 10.5 to 20.5 dpp, piRNA abundance increases and longer piRNAs appear, reflecting a switch from MILI-bound piRNAs, which have a 26–27 nt modal length (Montgomery et al., 1998; Aravin et al., 2006, 2008; Robine et al., 2009), to MIWI-bound piRNAs, which have a 30 nt modal length (Figure S2A; Reuter et al., 2009; Robine et al., 2009). This switch occurs at the pachytene phase of meiosis. MILI-bound pre-pachytene piRNAs predominate before the onset of pachynema; at the pachytene and round spermatid stages, most piRNAs are MIWI-bound pachytene piRNAs.

We used hierarchical clustering to analyze the change in piRNA abundance from 10.5 to 20.5 dpp for the 214 genes defined by our data (Figures 2A and S2A and Table S2). Three types of piRNA-producing genes were identified according to when their piRNAs first accumulate and how their expression changes during spermatogenesis: 84 pre-pachytene, 100 pachytene, and 30 hybrid loci. At 10.5 dpp, the earliest time we evaluated, 84 genes dominate piRNA production (median piRNA abundance per gene = 16 rpkm; Figure 2B). Nearly all (81 out of 84) were congruent with protein-coding genes. The 84 pre-pachytene piRNA genes account for 13% of piRNAs at 10.5 dpp, but only 0.31% of piRNAs in the adult testis. Of the pre-pachytene piRNAs accounted for by the 84 loci, 15% derive from 31 piRNA-producing genes that, to our knowledge, have not previously been described.

A parallel analysis of piRNA precursor transcription using RNA-seq (>100 nt) corroborated the classification based on piRNA abundance; of the 100 piRNA genes classified as pachytene based on the developmental expression profile of their piRNAs, 93 were grouped as pachytene according to the developmental expression profile of their transcripts. Of these 93, 89 are intergenic. All 84 piRNA genes designated pre-pachytene using piRNA data were classified as pre-pachytene according to their transcript abundance.

Despite their name, pre-pachytene piRNAs were readily detected in >90% and ~95% pure pachytene spermatocytes, as well as round spermatids (Figure S2B; Gan et al., 2011; Modzelewski et al., 2012). Transcript abundance from the 84

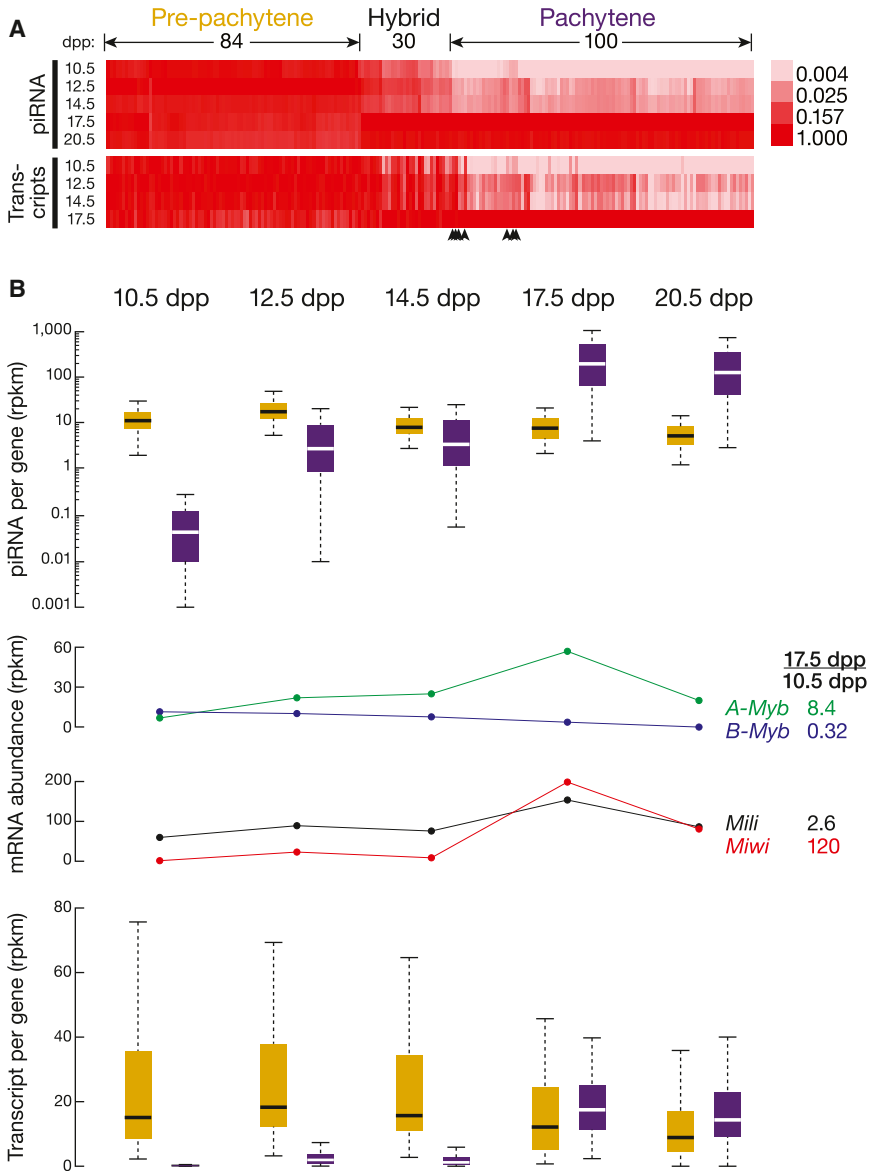


**Figure 2. Three Classes of piRNA-Generating Loci**

(A) Normalized piRNA density (rpkm) for each piRNA-producing gene is shown as a heatmap across the developmental stages. Hierarchical clustering divided the genes into three classes. Arrowheads mark seven pachytene piRNA genes that were not classified as pachytene according to the change in the abundance of their precursor RNAs from 10.5 to 17.5 dpp.

(B) Top: box plots present piRNA density per gene as spermatogenesis progresses (here and elsewhere, pre-pachytene in yellow and pachytene in purple). Middle: expression of *A-Myb*, *B-Myb*, *Mili*, and *Miwi* was measured by RNA-seq. Bottom: box plots present piRNA precursor expression per gene, measured by RNA-seq, from 10.5 to 20.5 dpp. See also Figure S2 and Table S2.

pre-pachytene loci was high at 3 dpp (median abundance = 11 rpkm), higher by 8 dpp (18 rpkm), and lower in purified leptotene/zygotene spermatocytes (3.3 rpkm; Figure S2B). Yet piRNA precursor transcripts were readily detectable in purified pachytene spermatocytes at a level (4.6 rpkm) comparable to that in purified leptotene/zygotene spermatocytes (Figure S2B; Gan et al., 2011; Modzelewski et al., 2012). From 10.5 to 20.5 dpp, the steady-state level of pre-pachytene piRNA precursor transcripts remained constant (Figure 2B).

Finally, the abundance of pre-pachytene piRNA precursor transcripts was better correlated with pre-pachytene piRNA abundance at 17.5 dpp ($\rho = 0.47$), when pachytene spermatocytes compose a larger fraction of the testis, than at 10.5, 12.5, or 14.5 dpp ($0.32 \leq \rho \leq 0.40$; Figure S2C). Our data suggest that the pre-pachytene loci continue to be transcribed and processed into piRNAs long after spermatocytes enter the pachytene stage of meiosis. Thus, the name pre-pachytene piRNA is a misnomer that should be retained only for historical reasons.

Hierarchical clustering identified 100 pachytene genes whose piRNAs emerge at 12.5 dpp, 2 days earlier than previously reported (Girard et al., 2006). Nearly all the pachytene genes are intergenic (93 out of 100). piRNA expression from pachytene piRNA genes peaks at 17.5 dpp (Figure 2B). Overall, the median abundance of piRNAs from these 100 loci increased >6,000-fold from 10.5 to 17.5 dpp. Transcripts from pachytene genes were low at 10.5 dpp (median abundance = 0.15 rpkm) and increased 116-fold from 10.5 to 17.5 dpp. From 10.5 to 20.5 dpp, the dynamics of pachytene piRNA abundance from each piRNA gene correlated with the increase in abundance of its precursor transcripts ($0.39 \leq \rho \leq 0.63$; p value $\leq 7.3 \times 10^{-5}$; Figure S2C). The 100 pachytene genes account for 92% of piRNAs in the adult testis, making it unlikely that biologically functional pachytene piRNAs originate from thousands of genomic loci (Gan et al., 2011). Figures 3 and S3 provide examples of pachytene and pre-pachytene piRNA genes defined by our data.

Hierarchical clustering detected a third class, hybrid piRNAs, which derives from 30 genes with characteristics of both pre-pachytene and pachytene piRNA loci. Like pre-pachytene, hybrid piRNAs were detected at 10.5 dpp (median abundance = 3.7 rpkm) and in purified spermatogonia (Gan et al., 2011). Like pachytene piRNAs, hybrid piRNA abundance increased during the pachytene stage of meiosis, but the increase was delayed until late (17.5 dpp) rather than early pachynema (14.5 dpp). Overall, piRNAs from hybrid genes increased >10-fold from 14.5 to 17.5 dpp. The median abundance of piRNAs from hybrid

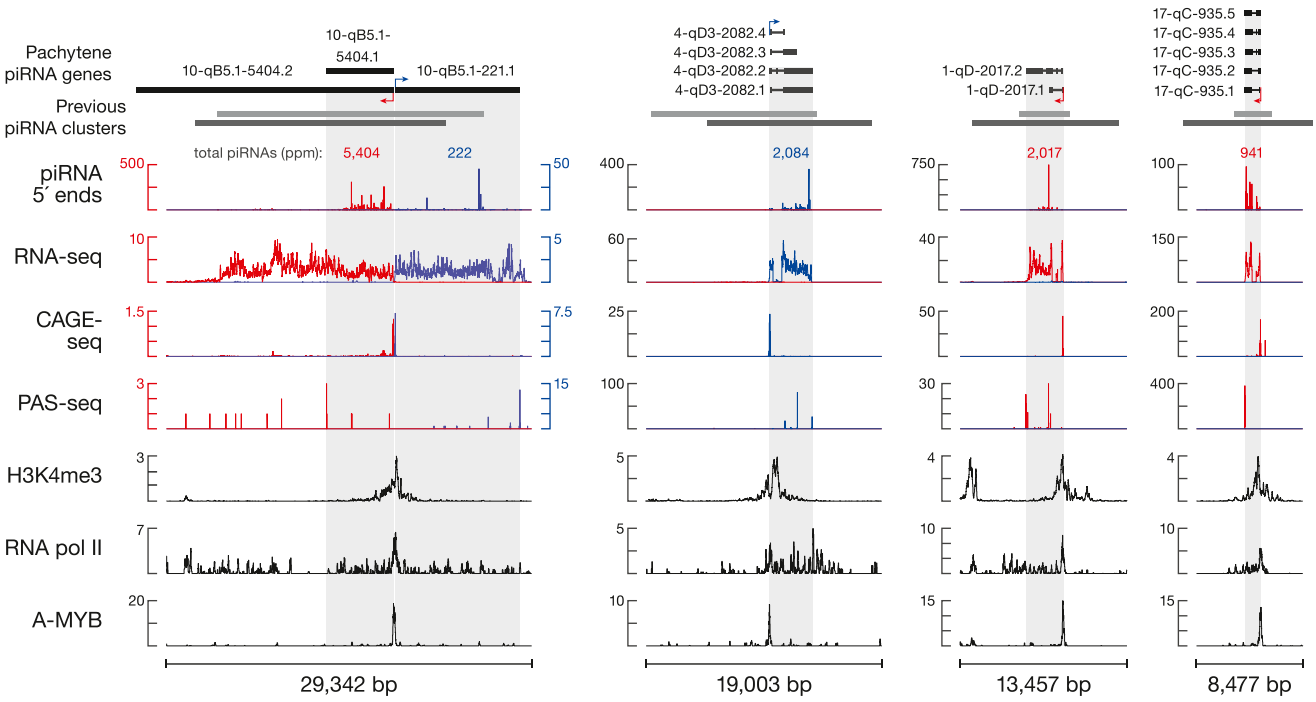


**Figure 3. Examples of Pachytene piRNA Genes**
Previous cluster boundaries are from Lau et al. (2006); gray and Girard et al. (2006); dark gray. See also Figure S3.

piRNA genes ranged from 90–120 rpkm in purified pachytene spermatocytes, >20-fold greater than their median abundance in spermatogonia (Gan et al., 2011; Modzelewski et al., 2012). Moreover, hybrid piRNA precursor transcripts were readily detected in purified pachytene spermatocytes (median abundance = 9.0 rpkm; Modzelewski et al., 2012).

## A-MYB Regulates Pachytene piRNA Precursor Transcription

The coordinated increase in pachytene piRNA precursor transcripts suggests their regulation by a common transcription factor or factors. Among the 100 pachytene piRNA genes, 15 pairs (30 genes) are divergently transcribed. The 5′ ends of the piRNA precursor RNAs from each pair are close in genomic distance (median = 127 bp), suggesting that a shared promoter lies between the two transcription start sites.

We took advantage of the unique genomic organization of these 15 pairs of divergently transcribed piRNA genes to search for sequence motifs common to their promoters. The MEME algorithm (Bailey and Elkan, 1994) revealed a motif highly enriched in these bidirectional promoters ($E = 8.3 \times 10^{-12}$; Figure 4A). This motif matches the binding site of the Myb family of transcription factors (Figure 4A; Gupta et al., 2007; Newburger and Bulyk, 2009). The Myb motif is not restricted to bidirectional promoters; MEME identified the same motif using the promoters of all pachytene piRNA genes ($E = 9.1 \times 10^{-28}$; Figure 4B).

The Myb transcription factor family is conserved among eukaryotes. Like other vertebrates, mice produce three Myb proteins, A-MYB (MYBL1), B-MYB (MYBL2), and C-MYB (MYB), each with a distinct tissue distribution (Mettus et al., 1994; Trauth et al., 1994; Latham et al., 1996; Oh and Reddy, 1999). Testes produce both A- and B-MYB proteins. Multiple lines of evidence implicate A-MYB, rather than B-MYB, as a candidate for regulating pachytene piRNA transcription. First, the expression of *A-Myb* during spermatogenesis resembles that of pachytene piRNAs: *A-Myb* transcripts appear at ~12.5 dpp and peak at 17.5 dpp (Figure 2B; Bolcun-Filas et al., 2011). The expression of *A-Myb* messenger RNA (mRNA) increases ~15-fold from 8 dpp to 19 dpp, whereas *B-Myb* mRNA expression remains constant and low during the same time frame and into adulthood (Horvath et al., 2009). Our RNA-seq data (Figure 2B) corroborate these findings. Indeed, in our RNA-seq analysis of adult testes, *A-Myb* mRNA was 24-fold more abundant than *B-Myb*. Second, a testis-specific *A-Myb* point-mutant allele, *Mybl1*$^{repro9}$, which is caused by a cytosine-to-adenine transversion that changes alanine 213 to glutamic acid, leads to meiotic arrest at the pachytene stage with subtle defects in autosome synapsis; *A-Myb* null mutant mice have defects in multiple tissues, including the testis and the mammary gland (Toscani et al., 1997; Bolcun-Filas et al., 2011). Third, our RNA-seq analysis of *A-Myb* mutant testes shows that there is no significant change in *B-Myb* expression in the mutant, compared to the heterozygous controls, at 14.5 or 17.5 dpp. Finally, B-MYB protein is not detectable in pachytene spermatocytes (Horvath et al., 2009).

To assess more directly the role of A-MYB in pachytene piRNA precursor transcription, we used anti-A-MYB antibody to perform ChIP followed by high-throughput sequencing of the A-MYB-bound DNA. The anti-A-MYB antibody is specific

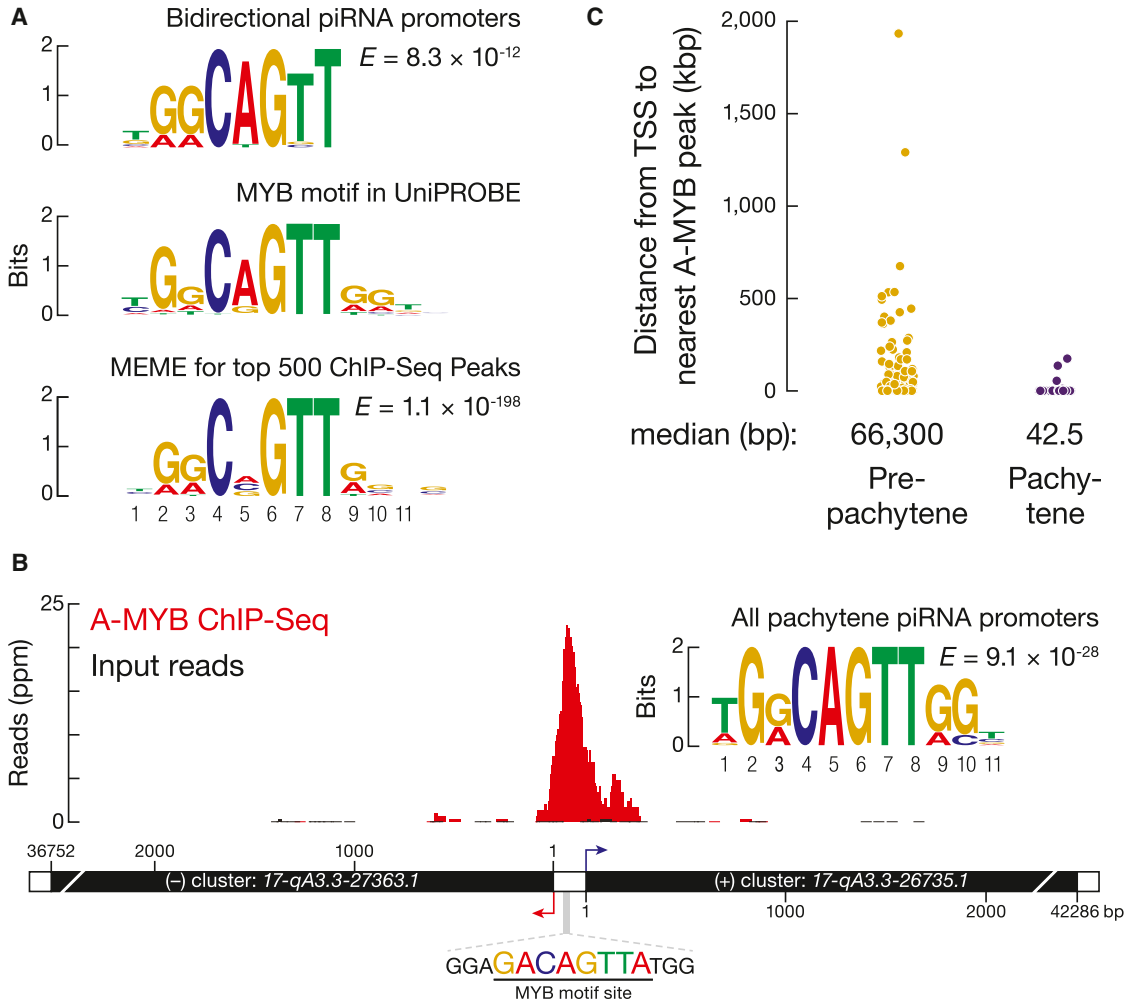


**Figure 4. A-MYB Binds the Promoters of Pachytene piRNA Genes**

(A) Top: MEME identified a sequence motif in the bidirectional promoters of the 15 pairs of divergently transcribed pachytene piRNA genes. *E* value computed by MEME measures the statistical significance of the motif. Middle: Myb motif from the mouse UniPROBE database. Bottom: MEME-reported motif for the top 500 (by peak score) A-MYB ChIP-seq peaks from adult mouse testes.

(B) A-MYB ChIP-seq data for the common promoter of the divergently transcribed pachytene piRNA genes *17-qA3.3-27363.1* and *17-qA3.3-26735.1*.

(C) The distance from the annotated transcription start site (TSS) of each piRNA gene to the nearest A-MYB peak. See also Figure S4.

for A-MYB, and the peptide used to raise the antibody is not present in B-MYB. The model-based analysis of ChIP-seq (MACS) algorithm (Zhang et al., 2008) reported 3,815 genomic regions with significant A-MYB binding (false discovery rate, FDR < $10^{-25}$); we call these regions A-MYB peaks or peaks. Among the 500 peaks with the lowest FDR values, 394 (80%) contained at least one significant site ($p < 10^{-4}$) for the MYB binding motif (Figure 4A). Figure 4B shows an example of such an A-MYB peak at the bidirectional promoter of the divergently transcribed pair of pachytene piRNA genes *17-qA3.3-27363.1* and *17-qA3.3-26735.1*. A-MYB occupancy of this genomic site was confirmed by ChIP and quantitative PCR (ChIP-qPCR) (Figure S4A).

The median distance from the transcription start site to the nearest A-MYB peak was ~43 bp for the 100 pachytene piRNA genes but >66,000 bp for the 84 pre-pachytene genes (Figure 4C). Our data suggest that during mouse spermatogenesis A-MYB binds to the promoters of both divergently and unidirectionally transcribed pachytene piRNA genes.

To test the idea that A-MYB promotes transcription of pachytene, but not pre-pachytene, piRNA genes, we used RNA-seq to measure the abundance of RNA > 100 nt long from the testes of *A-Myb* point-mutant (*Mybl1*$^{repro9}$) mice and their heterozygous littermates (Figure 5). Pachytene piRNA precursor transcripts—both divergently and unidirectionally transcribed—were significantly depleted in *A-Myb* mutant testes compared to the heterozygotes: the median decrease was 45-fold at 14.5 dpp ($q = 1.1 \times 10^{-13}$) and 248-fold at 17.5 dpp ($q = 3.9 \times 10^{-23}$). The abundance of pre-pachytene piRNA transcripts was not significantly changed ($q \geq 0.34$). The binding of A-MYB to the promoters of pachytene piRNA genes, together with the depletion of pachytene piRNA transcripts in the *A-Myb* mutant, further

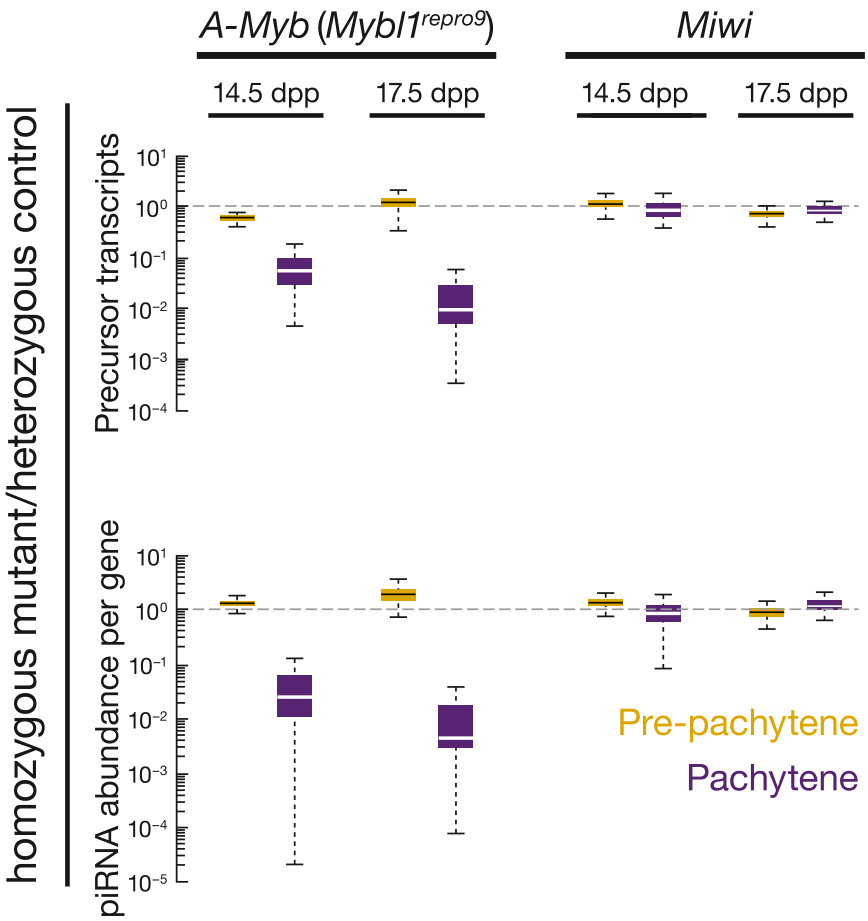


**Figure 5. Pachytene piRNAs and Precursors Decrease in *A-Myb* Mutant Testes**

The change in transcript or piRNA abundance per gene in *A-Myb* (n = 3) and *Miwi* (n = 1) mutants compared to heterozygotes in testes isolated at 14.5 and 17.5 dpp. See also Figure S5.

supports the view that A-MYB directly regulates transcription of pachytene piRNA genes.

### *A-Myb* Regulates Pachytene piRNA Production

To test the consequences of the loss of piRNA precursor transcripts, we measured piRNA abundance in the *A-Myb* mutant. Like pachytene piRNA precursor transcription, pachytene piRNA abundance significantly decreased in mutant testes. At 14.5 dpp, median piRNA abundance per pachytene gene decreased 87-fold in *A-Myb* homozygous mutant testes compared to heterozygotes ($p < 2.2 \times 10^{-16}$; Figure 5). By 17.5 dpp, median pachytene piRNA abundance was >9,000 times lower in the *A-Myb* mutant than the heterozygotes ($p < 2.2 \times 10^{-16}$). In contrast, pre-pachytene piRNA levels were essentially unaltered. Figure 6 presents examples of the effect at 14.5 and 17.5 dpp of the *A-Myb* mutant on piRNA precursor transcript and mature piRNA abundance for one pre-pachytene and three pachytene piRNA genes.

Our data show that A-MYB binds to the promoters of pachytene piRNA genes; *A-Myb*, *Miwi*, and pachytene piRNA precursor transcription begins at 12.5 dpp; and *A-Myb* mutant spermatocytes reach pachynema with subtle defects in autosome synapsis (Bolcun-Filas et al., 2011). Could pachytene piRNA depletion nonetheless be an indirect consequence of the meiotic arrest caused by the *A-Myb* mutant? To test this possibility, we sequenced small RNAs from *Spo11* mutant testes, which failed to generate double-stranded DNA breaks at the leptotene stage and display a meiotic arrest (Baudat et al., 2000; Romanienko and Camerini-Otero, 2000). The median abundance of piRNAs from pre-pachytene genes did not decrease at 14.5 dpp. By 17.5 dpp, piRNA from pachytene genes decreased just 5.9-fold in the *Spo11* mutant testes compared to the heterozygotes (Figure S5). We note that A-MYB protein abundance is reduced in the *Spo11* mutant (Bolcun-Filas et al., 2011).

*Trip13* is required to complete the repair of double-strand DNA breaks on fully synapsed chromosomes. *Trip13* mutants display a meiotic arrest similar to that in *A-Myb* mutant testes (Li and Schimenti, 2007): pachytene arrest with synapsed chromosomes. To further test whether the loss of pachytene piRNA precursor transcripts in *A-Myb* mutants reflects a general effect of meiotic arrest, we measured piRNA precursor transcript abundance in *Trip13* mutant testes at 17.5 dpp. Unlike *A-Myb*, piRNA precursor transcripts were readily detectable in the *Trip13* mutant (Figure S6). We conclude that the loss of pachytene piRNA precursor transcripts and piRNAs in *A-Myb* mutant testes is a direct consequence of the requirement for A-MYB to transcribe pachytene piRNA genes and not a general feature of meiotic arrest at the pachytene stage.

### A-MYB Regulates Expression of piRNA Biogenesis Factors

The *A-Myb* mutant more strongly affected pachytene piRNA accumulation than it did the steady-state abundance of the corresponding piRNA precursor transcripts (Figure 5); the median decrease in pachytene piRNA abundance was 2-fold greater at 14.5 dpp and 38-fold greater at 17.5 dpp than the decrease in the steady-state abundance of pachytene precursor transcripts (Table S1). These data suggest that A-MYB exerts a layer of control on piRNA accumulation beyond its role in promoting pachytene piRNA precursor transcription.

*Miwi* has previously been proposed to be a direct target of A-MYB; *Miwi* mRNA abundance is reduced in A-MYB mutant testes, and ChIP microarray data place A-MYB on the *Miwi* promoter (Bolcun-Filas et al., 2011). Our RNA-seq data confirm that accumulation of *Miwi* mRNA requires A-MYB: *Miwi* mRNA

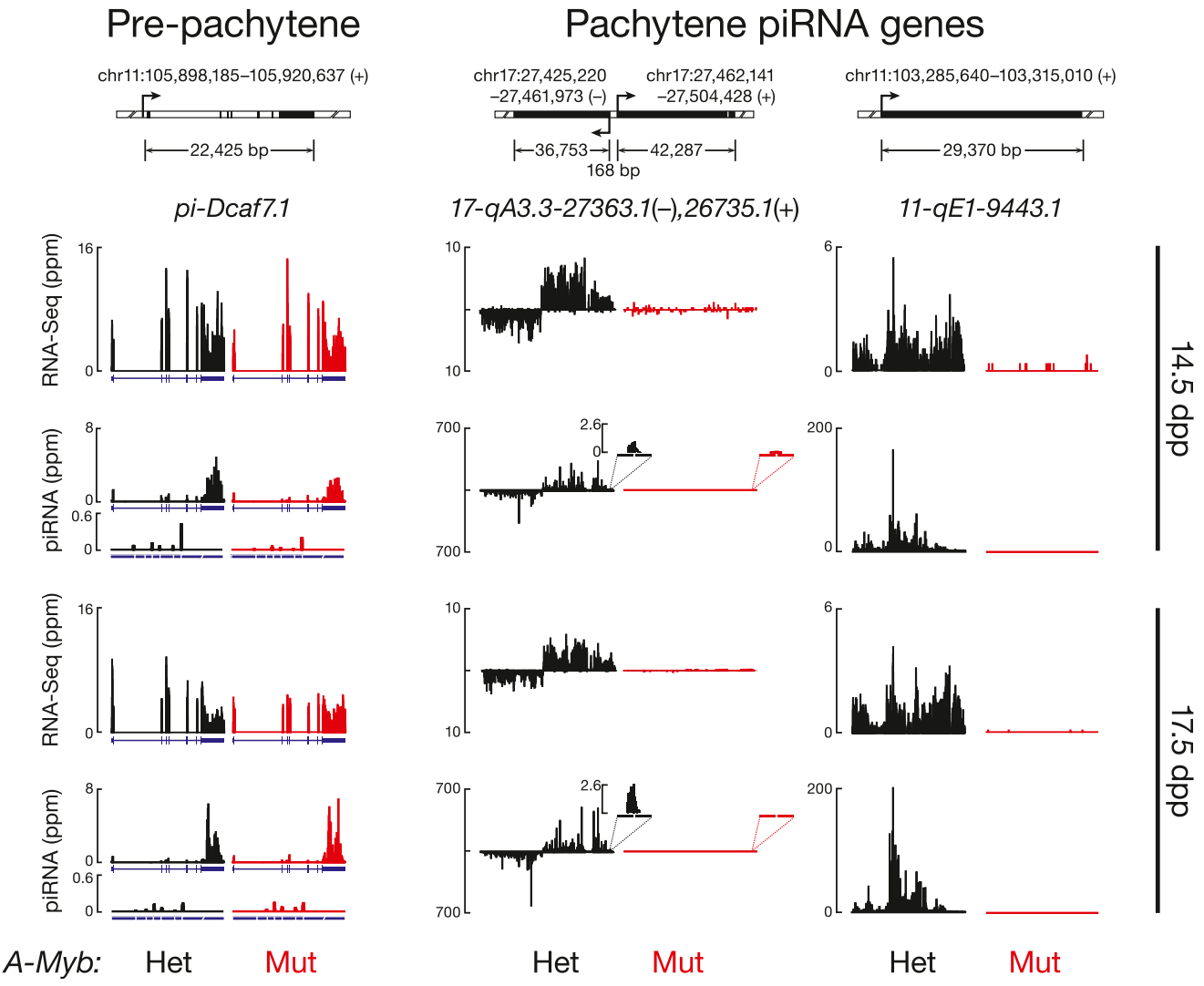


**Figure 6. Examples of the Effect of the *A-Myb* Mutation on piRNA Expression**

Transcript and piRNA abundance in heterozygous (Het) and homozygous *A-Myb* (Mut) point-mutant testes is shown for four illustrative examples at 14.5 and 17.5 dpp. Also shown is the abundance of piRNA sequencing reads that map to the exon-exon junctions. Gene *11-qE1-9443* does not have an intron. Exons, blue boxes; splice junctions, gaps; the last exon is compressed and not to scale. See also Figure S6.

decreased more than 50-fold in testes isolated from *A-Myb* mutant mice at 14.5 dpp compared to their heterozygous littermates (Figures 7A and S7 and Table S3). Furthermore, our ChIP data confirm that A-MYB binds the *Miwi* promoter in vivo (Figures 7B, S4B, and S7). Like pachytene piRNAs, *Miwi* transcripts first appear at 12.5 dpp (Figure 2B), and MIWI protein is first detected in testes at 14.5 dpp (Deng and Lin, 2002). Loss of MIWI arrests spermatogenesis at the round spermatid stage (Deng and Lin, 2002).

A previous study reported that piRNAs fail to accumulate to wild-type levels in *Miwi* mutant testes (Grivna et al., 2006b). However, our data suggest that the overall change in piRNA abundance caused by loss of MIWI is quite small: RNA-seq detected no change at 14.5 dpp (change in total piRNA abundance = 1.1; n = 2) and only a modest decrease at 17.5 dpp (change in total piRNA abundance = 0.58; n = 1). piRNAs from pachytene loci decreased just 2.7-fold at 14.5 dpp ($p = 0.0046$) and 3.5-fold at 17.5 dpp ($p = 1.8 \times 10^{-6}$) in *Miwi* mutant testes (Figure 5). By comparison, pachytene piRNAs declined 87-fold at 14.5 dpp and 9,400-fold at 17.5 dpp in the *A-Myb* mutant.

Does the loss of MIWI affect piRNA precursor transcription? We measured transcript abundance and piRNA expression in *Miwi* null mutant testes at 14.5 and 17.5 dpp. In *Miwi*$^{-/-}$ testes, pachytene piRNA precursor transcripts were present at levels indistinguishable from *Miwi* heterozygotes (median change = 1.0- to 1.4-fold; q = 1; Figure 5). Thus, loss of MIWI does not explain loss of pachytene piRNA precursor transcripts in *A-Myb* mutant testes.

In addition to *Miwi*, ChIP-seq detected A-MYB bound to the promoters of 12 other RNA-silencing-pathway genes (Figure 7B

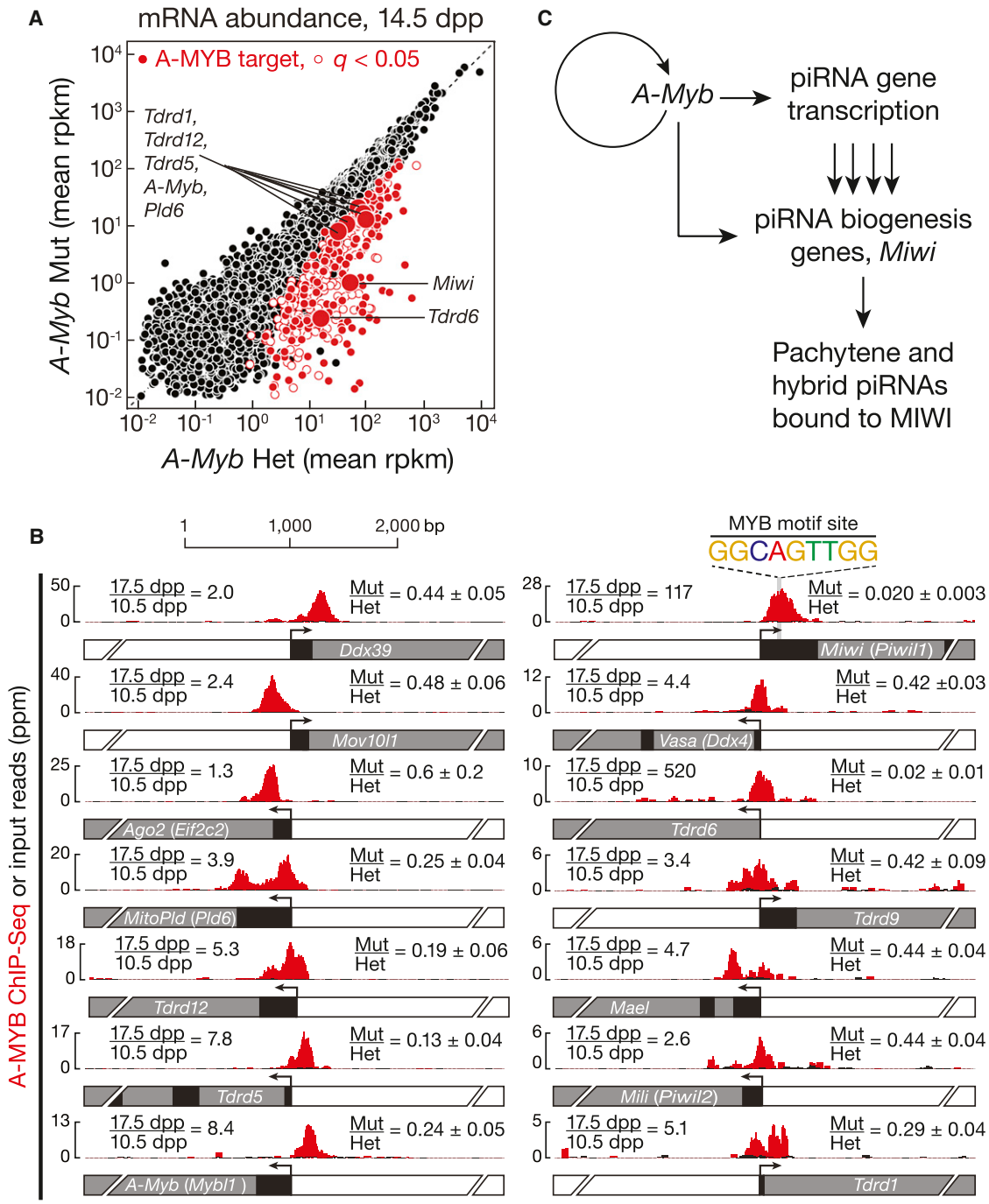


**Figure 7. A-MYB Regulates Expression of mRNAs Encoding piRNA Pathway Proteins**

(A) mRNA abundance in *A-Myb* mutant versus heterozygous testes. The 407 genes with a significant (q < 0.05) change in steady-state mRNA levels are shown as red circles. The 203 with A-MYB peaks within 500 bp of their transcription start site are filled.

(B) A-MYB ChIP-seq signal at the transcription start sites of *A-Myb* and genes implicated in RNA silencing pathways. For each, the figure reports the change in mRNA abundance between 17.5 and 10.5 dpp in wild-type testes and the mean change between *A-Myb* mutant and heterozygous testes at 14.5 dpp (mean ± SD; n = 3).

(C) A model for the regulation of pachytene piRNA biogenesis by A-MYB. See also Figure S7 and Table S3.

and Table S3). Of these, the mRNA abundance—measured by three biologically independent RNA-seq experiments—of *Ago2*, *Ddx39* (*uap56* in flies), *Mael*, *Mili*, *Mov10l1*, *Tdrd9*, and *Vasa* did not change significantly at 14.5 dpp in *A-Myb* mutant testes compared to heterozygotes (q > 0.05); except for *Ago2*, all decreased significantly in the mutant at 17.5 dpp. In contrast,

the abundance of the mRNAs encoding Tudor domain proteins decreased significantly in *A-Myb* mutant testes: *Tdrd6* (64-fold decrease; q = 3.1 × $10^{-5}$) and *Tdrd5* (7.5-fold decrease; q = 1.0 × $10^{-5}$). *Tdrd5* is expressed in embryonic testes then decreases around birth (Yabuta et al., 2011). TDRD5 protein reappears at 12 dpp, increasing throughout the pachynema (Smith et al., 2004; Yabuta et al., 2011). Our data indicate that A-MYB activates *Tdrd5* transcription at the onset of the pachytene stage of meiosis. Similarly, *Tdrd6* mRNA can be detected at the middle pachytene, but not the zygotene stage, and peaks after late pachytene; TDRD6 protein can be detected at 17 dpp and continues to increase until 21 dpp (Vasileva et al., 2009). The findings that TDRD5 and TDRD6 colocalize with MIWI in pachytene spermatocytes (Hosokawa et al., 2007; Vasileva et al., 2009; Yabuta et al., 2011) and that TDRD6 binds MIWI (Chen et al., 2009; Vagin et al., 2009; Vasileva et al., 2009) suggest a role for these Tudor domain proteins in pachytene piRNA production or function. As in *Miwi*$^{-/-}$ testes, spermatogenesis arrests at the round spermatid stage in *Tdrd5*$^{-/-}$ and *Tdrd6*$^{-/-}$ mutant testes (Vasileva et al., 2009; Yabuta et al., 2011). Loss of *Tdrd6* expression has little effect on piRNA levels (Figure S3; Vagin et al., 2009), perhaps because the functions of Tudor domain proteins overlap.

Other genes encoding piRNA pathway proteins whose promoters are bound by A-MYB and whose expression decreased significantly in *A-Myb* mutant testes include *MitoPld* (*Pld6*; 3.9-fold decrease; q = 0.0095) and *Tdrd12* (5.3-fold decrease; q = 0.0046). *MitoPld* encodes an endoribonuclease implicated in an early step in piRNA biogenesis in mice and flies (Houwing et al., 2007; Pane et al., 2007; Haase et al., 2010; Huang et al., 2011; Watanabe et al., 2011; Ipsaro et al., 2012; Nishimasu et al., 2012). The function of *Tdrd12* is not known, but its fly homologs (*Yb*, *Brother of Yb*, and *Sister of Yb*) are all required for piRNA production (Handler et al., 2011). *Tdrd1* decreased 3.4-fold, but with q value = 0.015. *Tdrd1* is first expressed in fetal prospermatogonia, then re-expressed in pachytene spermatocytes (Chuma et al., 2006). In *Tdrd1* mutant testes, spermatogenesis fails, with no spermatocytes progressing past the round spermatid stage (Chuma et al., 2006). TDRD1 binds MILI and MIWI (Chen et al., 2009; Kojima et al., 2009) and colocalizes with TDRD5 and TDRD6 in the chromatoid body (Hosokawa et al., 2007).

Together, these data support the idea that at the onset of the pachytene phase of meiosis, A-MYB coordinately activates transcription of many genes encoding piRNA pathway proteins.

## A-MYB and the Pachytene piRNA Regulatory Circuitry

A number of genes encoding known and suspected piRNA pathway proteins are bound and regulated by A-MYB (Figures 7B and S7C). Our data support a model in which A-MYB drives both the transcription of pachytene piRNA genes and the mRNAs encoding genes required for piRNA production including *Miwi*, *MitoPld*, and *Tdrd9*. Regulation by A-MYB of both the sources of pachytene piRNAs and the piRNA biogenesis machinery creates a coherent feedforward loop (Figure 7C). Feedforward loops amplify initiating signals to increase target gene expression. Furthermore, they function as switches that are sensitive to sustained signals; they reject transient signals (Shen-Orr et al., 2002; Osella et al., 2011).

A-MYB also bound to the *A-Myb* promoter (Figure 7B), and *A-Myb* transcripts decreased 4.2-fold in testes from an *A-Myb* point mutant (*Mybl1*$^{repro9}$; Figure 7B). The *A-Myb* mutant fails to produce the high level of A-MYB protein observed in wild-type testes at the late pachytene stage of meiosis (Bolcun-Filas et al., 2011). Instead, A-MYB protein never becomes more abundant than the level achieved in wild-type testes by the beginning of the pachytene stage. While the lower level of A-MYB in the *A-Myb* mutant may reflect instability of the mutant protein, a simpler explanation is that mutant A-MYB cannot activate *A-Myb* transcription.

## Feedforward Regulation of piRNA Production Is Evolutionarily Conserved

Is A-MYB-mediated, feedforward control a general feature of regulation of piRNA production among vertebrates? To test whether A-MYB control of piRNA precursor transcription is evolutionarily conserved, we used high-throughput sequencing to identify piRNAs in adult rooster testes. Birds and mammals diverged 330 million years ago (Benton and Donoghue, 2007). After removing the sequences of identifiable miRNAs (Burnside et al., 2008) and annotated noncoding RNAs, total small RNA from the adult rooster testis showed peaks at both 23 and 25 nt (Figure 8A). When the RNA was oxidized before being prepared for sequencing, only a single 25 nt peak remained, consistent with the 25 nt small RNAs corresponding to piRNAs containing 2′-*O*-methyl-modified 3′ termini. These longer, oxidation-resistant species typically began with uracil (62% of species and 65% of reads; Figure 8B), and we detected a significant Ping-Pong amplification signature (*Z* score = 31; Figure 8C). We conclude that the oxidation-resistant, 24–30 nt long small RNAs correspond to rooster piRNAs. Like piRNAs generally, rooster piRNAs are diverse, with 5,742,529 species present among 81,121,893 genome-mapping reads. Like mouse pachytene piRNAs, 70% of piRNAs from adult rooster testes mapped to unannotated intergenic regions, 19% mapped to transposons, and 14% mapped to protein-coding genes. Of the piRNAs that map to protein-coding genes, >95% derive from introns. Forty-two percent of piRNA species mapped uniquely to the *Gallus gallus* genome.

Using 24–30 nt piRNAs from oxidized libraries, we identified 327 rooster piRNA clusters (Figure S8). These account for 76% of all uniquely mapping piRNAs. Of the 327 clusters, 25 overlapped with protein-coding genes. To begin to identify the transcription start sites for the rooster piRNA clusters, we analyzed adult rooster testes by H3K4me3 ChIP-seq. More than 81% (268 out of 327) of the clusters contained a readily detectable H3K4me3 peak within 1 kbp of the piRNA cluster. In contrast, the median distance from a cluster to the nearest transcription start site of an annotated gene was 73 kbp, suggesting that the H3K4me3 peaks reflect the start sites for rooster piRNA precursor transcripts.

Next, we asked where in the genome A-MYB bound in adult rooster testes. A-MYB ChIP-seq identified 5,509 significant peaks (FDR < $10^{-25}$). MEME analysis of the top 500 peaks with the lowest FDR values identified a motif (*E* = 2.6 × $10^{-201}$; Figure 8D) similar to that found in the mouse (Figure 4A). *A-MYB* is the only one of the three chicken *MYB* genes expressed in

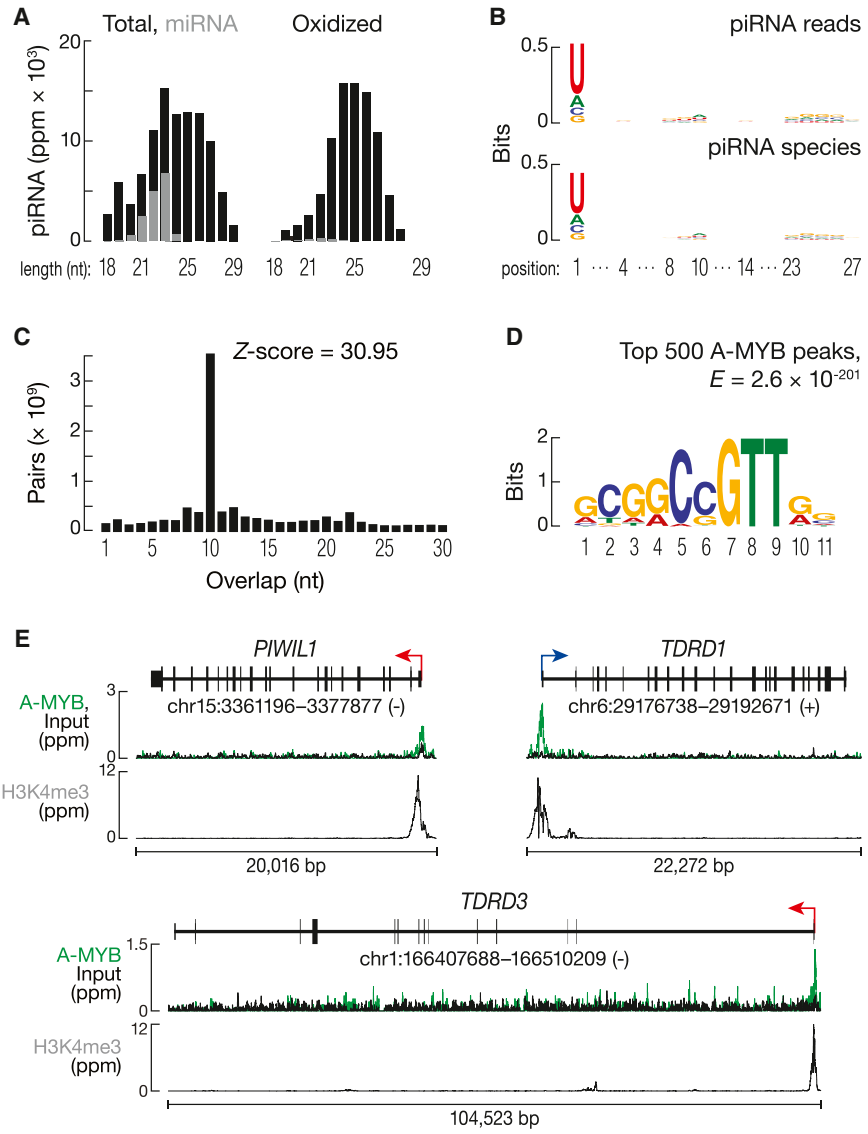


**Figure 8. Feedforward Regulation of piRNA Biogenesis by A-MYB Is Conserved in Rooster**

(A) Length distributions of total rooster testis small RNAs (black) and miRNAs (gray).

(B) Sequence logo showing the nucleotide composition of piRNA reads and species.

(C) The 5′-5′ overlap between piRNAs from opposite strands was analyzed to determine if rooster piRNAs display Ping-Pong amplification. The number of pairs of piRNA reads at each position is reported. *Z* score indicates that a significant 10 nt overlap (Ping-Pong) was detected. *Z* score > 1.96 corresponds to p value < 0.05.

(D) MEME-reported motif of the top 500 (by peak score) A-MYB ChIP-seq peaks from adult rooster testes.

(E) A-MYB, H3K4me3, and input ChIP-seq signals at the transcription start sites of rooster *PIWIL1*, *TDRD1*, and *TDRD3*. See also Figure S8.

adult testis (X.Z.L. and P.D.Z., unpublished data), supporting the view that these peaks correspond to A-MYB binding. The core sequence motif associated with A-MYB binding in mouse differs at one position (CAGTT) from that in rooster (C C/G GTT). This difference between mammalian and chicken MYB proteins has been noted previously (Weston, 1992; Deng et al., 1996).

To determine whether chicken A-MYB might regulate transcription of some piRNA clusters in the testis, we compared the A-MYB peak nearest to each piRNA cluster with the nearest H3K4me3 peak. Of the 327 rooster piRNA clusters, at least 104 were occupied by A-MYB at their promoters, as defined by an overlapping H3K4me3 peak. These 104 clusters account for 31% of uniquely mapping rooster piRNAs.

The chicken genome encodes at least two PIWI proteins: PIWIL1 and PIWIL2. Remarkably, the promoter of *Gallus gallus PIWIL1*, the homolog of mouse *Miwi*, contained a prominent A-MYB peak (Figure 8E). *TDRD1* and *TDRD3* also showed A-MYB peaks (Figure 8E). Thus, as in mice, *Gallus gallus* A-MYB controls the transcription of both piRNA clusters and genes encoding piRNA pathway proteins. We conclude that A-MYB-mediated feedforward regulation of piRNA production was likely present in the last common ancestor of birds and mammals.

In mice, we found no piRNA-producing genes on the sex chromosomes (Figure S1A), perhaps because mouse sex chromosomes are silenced during the pachytene stage (Li et al., 2009b). Birds use a ZW rather than an XY mechanism for sex determination, so roosters are homogametic (ZZ), allowing the sex chromosomes to remain transcriptionally active in males (Namekawa and Lee, 2009; Schoenmakers et al., 2009). Indeed, we find that 39 of the 327 rooster piRNA clusters are on the Z chromosome, accounting for 12% of uniquely mapping piRNAs (Figure S8). Of the 39 Z chromosome clusters, 18 had an A-MYB peak at their promoter.

## DISCUSSION

The data presented here provide strong support for the view that piRNAs in mammals begin as long, single-stranded precursors generated by testis-specific, RNA Pol II transcription of individual piRNA genes (see also Vourekas et al., 2012). Transcription by RNA Pol II affords piRNA genes the same rich set of transcriptional controls available to regulate mRNA expression. Our data establish that developmentally regulated transcription of piRNA genes determines when specific classes of piRNAs emerge during spermatogenesis.

During mouse spermatogenesis, transcription of pachytene piRNA genes begins at the onset of the pachytene stage of

meiosis; pachytene piRNAs accumulate subsequently. The presence of the MYB binding motif near the transcription start sites of pachytene piRNA genes, the physical binding of A-MYB to those genes, and the loss of pachytene piRNA precursor transcripts and piRNAs in testes from *A-Myb* mutant mice all argue that A-MYB regulates pachytene piRNA production.

A-MYB also drives increased expression of piRNA pathway genes. Among these, *Miwi* expression shows the greatest dependence on A-MYB, but A-MYB also drives transcription of genes encoding other proteins in the piRNA pathway, including *MitoPld*, *Mael*, and five genes encoding Tudor domain proteins. For example, A-MYB increases expression of *Tdrd6* more than 500-fold. Loss of A-MYB function more strongly depletes pachytene piRNAs than loss of MIWI, in part because pachytene piRNAs can still be loaded into MILI in *Miwi* mutant testes, although MILI-loaded pachytene piRNAs do not suffice to produce functional sperm. In the *A-Myb* mutant, expression of mRNAs encoding multiple piRNA pathway proteins decreases. We speculate that in wild-type male mice, the increased expression of these mRNAs at the onset of the pachytene stage of meiosis ensures that sufficient piRNA-precursor-processing and MIWI-loading factors are available to cope with the large increase in pachytene piRNA precursor transcription.

We propose that induction of A-MYB during the early pachytene stage of spermatogenesis initiates a feedforward loop that ensures the precisely timed production of these piRNAs. Coherent feedforward loops show delayed kinetics in order to reject background stimuli (Mangan and Alon, 2003). Indeed, we observed a delay from the early to middle pachytene in the accumulation of pachytene piRNAs, despite the continued increase in *A-Myb* expression (Figure 2A). Pachytene piRNA levels increase 75-fold (median for the 100 genes) from 10.5 to 12.5 dpp, coincident with increased expression of *A-Myb*. However, from 12.5 to 14.5 dpp, pachytene piRNAs increase only 1.2-fold. Pachytene piRNAs subsequently resume their accumulation, increasing 65-fold from 14.5 to 17.5 dpp. We believe this delay is a consequence of a feedforward loop that ensures the production of pachytene piRNAs only at the pachytene stage of spermatogenesis. Regulation by a feedforward loop also predicts a rapid shutdown of pachytene piRNA pathways at round spermatid stage VIII, when A-MYB protein levels decrease (Horvath et al., 2009). Supporting this idea, the abundance of MIWI decreases sharply by the elongated spermatid stage of spermatogenesis (Deng and Lin, 2002). Testing this proposal is a clear challenge for the future.

In fruit flies and zebrafish (Brennecke et al., 2007; Houwing et al., 2007), most piRNAs map to repetitive regions, whereas in mammals, uniquely mapping intergenic piRNAs predominate in the adult testis. The discovery that 70% of rooster piRNA reads map to intergenic regions suggests that the expansion of intergenic piRNAs controlled by A-MYB feedforward regulation arose before the divergence of birds and mammals. In the future, detailed analysis of piRNA production across avian spermatogenesis should provide insight into the evolutionary origins and functions of pachytene piRNAs, a class of piRNAs thus far only detected in mammals.

In summary, we have shown that mouse piRNA genes are coregulated transcriptionally, establishing that A-MYB coordinately regulates the biogenesis of an entire piRNA class, the pachytene piRNAs. The discovery that a loss-of-function *A-Myb* mutant, *Mybl1*$^{repro9}$, disrupts piRNA precursor transcription in vertebrates provides a tool to understand the transformation of long, single-stranded piRNA precursors into mature piRNAs and to explore the functions and targets of the pachytene piRNAs.

## EXPERIMENTAL PROCEDURES

### Mice

*Mybl1*$^{repro9}$, *Spo11*$^{tm1Sky}$, and *Piwil1*$^{tm1Hf}$ mice were maintained and used according to the guidelines of the Institutional Animal Care and Use Committee of the University of Massachusetts Medical School and genotyped as described (Baudat et al., 2000; Deng and Lin, 2002; Bolcun-Filas et al., 2011).

### Sequencing

Small (Ghildiyal et al., 2008; Seitz et al., 2008) and long RNA-seq (Zhang et al., 2012) and analysis (Li et al., 2009a) were as described. Reads that did not map to mouse genome mm9 were mapped to piRNA precursor transcripts to obtain splice junction mapping small RNAs. Total small RNA libraries from different developmental stages and from mutants were normalized to the sum of all miRNA hairpin mapping reads. Oxidized samples were calibrated to the corresponding total small RNA library via the abundance of shared, uniquely mapped piRNA species. piRNA expression data were grouped with Cluster 3.0. Differential gene expression was analyzed with DESeq R (Anders and Huber, 2010); ChIP-seq reads were aligned to the genome using Bowtie version 0.12.7 (Langmead et al., 2009), and peaks were identified using MACS (Zhang et al., 2008).

## ACCESSION NUMBERS

The Gene Expression Omnibus (GEO) accession number for the RNA-seq, ChIP-seq, and small RNA data reported in this paper is GSE44690.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes eight figures, three tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2013.02.016.

## ACKNOWLEDGMENTS

We thank K. Chase and K. Schimenti for help collecting tissues; C. Tipping for help with mouse husbandry; P. Johnson and B. Keagle for providing rooster testes; G. Farley for technical assistance; H. Lin for reagents; Xi Chen, Xiaotu Ma, Oliver Rando, and Benjamin Carone for advice on ChIP; and members of our laboratories for critical comments on the manuscript. X.Z.L. was supported by the Lalor Foundation and the Jane Coffin Childs Memorial Fund for Medical Research.

Received: November 20, 2012
Revised: January 17, 2013
Accepted: February 12, 2013
Published: March 21, 2013

## REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Aravin, A.A., and Hannon, G.J. (2008). Small RNA silencing pathways in germ and stem cells. Cold Spring Harb. Symp. Quant. Biol. *73*, 283–290.

Aravin, A.A., Naumova, N.M., Tulin, A.V., Vagin, V.V., Rozovsky, Y.M., and Gvozdev, V.A. (2001). Double-stranded RNA-mediated silencing of genomic

tandem repeats and transposable elements in the *D. melanogaster* germline. Curr. Biol. *11*, 1017–1027.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., et al. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203–207.

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. Science *316*, 744–747.

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol. Cell *31*, 785–799.

Ashe, A., Sapetschnig, A., Weick, E.M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., et al. (2012). piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell *150*, 88–99.

Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. *2*, 28–36.

Baudat, F., Manova, K., Yuen, J.P., Jasin, M., and Keeney, S. (2000). Chromosome synapsis defects and sexually dimorphic meiotic progression in mice lacking Spo11. Mol. Cell *6*, 989–998.

Benton, M.J., and Donoghue, P.C. (2007). Paleontological evidence to date the tree of life. Mol. Biol. Evol. *24*, 26–53.

Bolcun-Filas, E., Bannister, L.A., Barash, A., Schimenti, K.J., Hartford, S.A., Eppig, J.J., Handel, M.A., Shen, L., and Schimenti, J.C. (2011). A-MYB (MYBL1) transcription factor is a master regulator of male meiosis. Development *138*, 3319–3330.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. Cell *128*, 1089–1103.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., and Hannon, G.J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. Science *322*, 1387–1392.

Burnside, J., Ouyang, M., Anderson, A., Bernberg, E., Lu, C., Meyers, B.C., Green, P.J., Markis, M., Isaacs, G., Huang, E., and Morgan, R.W. (2008). Deep sequencing of chicken microRNAs. BMC Genomics *9*, 185.

Carmell, M.A., Girard, A., van de Kant, H.J., Bourc'his, D., Bestor, T.H., de Rooij, D.G., and Hannon, G.J. (2007). MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev. Cell *12*, 503–514.

Cenik, E.S., and Zamore, P.D. (2011). Argonaute proteins. Curr. Biol. *21*, R446–R449.

Chen, C., Jin, J., James, D.A., Adams-Cioaba, M.A., Park, J.G., Guo, Y., Tenaglia, E., Xu, C., Gish, G., Min, J., and Pawson, T. (2009). Mouse Piwi interactome identifies binding mechanism of Tdrkh Tudor domain to arginine methylated Miwi. Proc. Natl. Acad. Sci. USA *106*, 20336–20341.

Chuma, S., Hosokawa, M., Kitamura, K., Kasai, S., Fujioka, M., Hiyoshi, M., Takamune, K., Noce, T., and Nakatsuji, N. (2006). Tdrd1/Mtr-1, a tudor-related gene, is essential for male germ-cell differentiation and nuage/germinal granule formation in mice. Proc. Natl. Acad. Sci. USA *103*, 15894–15899.

Deng, W., and Lin, H. (2002). *miwi*, a murine homolog of *piwi*, encodes a cytoplasmic protein essential for spermatogenesis. Dev. Cell *2*, 819–830.

Deng, Q.L., Ishii, S., and Sarai, A. (1996). Binding site analysis of c-Myb: screening of potential binding sites by using the mutation matrix derived from systematic binding affinity measurements. Nucleic Acids Res. *24*, 766–774.

Farazi, T.A., Juranek, S.A., and Tuschl, T. (2008). The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. Development *135*, 1201–1214.

Gan, H., Lin, X., Zhang, Z., Zhang, W., Liao, S., Wang, L., and Han, C. (2011). piRNA profiling during specific stages of mouse spermatogenesis. RNA *17*, 1191–1203.

Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L., Zapp, M.L., Weng, Z., and Zamore, P.D. (2008). Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. Science *320*, 1077–1081.

Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature *442*, 199–202.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. *29*, 644–652.

Grivna, S.T., Beyret, E., Wang, Z., and Lin, H. (2006a). A novel class of small RNAs in mouse spermatogenic cells. Genes Dev. *20*, 1709–1714.

Grivna, S.T., Pyhtila, B., and Lin, H. (2006b). MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. Proc. Natl. Acad. Sci. USA *103*, 13415–13420.

Gu, W., Lee, H.C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., Jr., and Mello, C.C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. Cell *151*, 1488–1500.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77–88.

Gupta, S., Stamatoyannopoulos, J.A., Bailey, T.L., and Noble, W.S. (2007). Quantifying similarity between motifs. Genome Biol. *8*, R24.

Haase, A.D., Fenoglio, S., Muerdter, F., Guzzardo, P.M., Czech, B., Pappin, D.J., Chen, C., Gordon, A., and Hannon, G.J. (2010). Probing the initiation and effector phases of the somatic piRNA pathway in *Drosophila*. Genes Dev. *24*, 2499–2504.

Handler, D., Olivieri, D., Novatchkova, M., Gruber, F.S., Meixner, K., Mechtler, K., Stark, A., Sachidanandam, R., and Brennecke, J. (2011). A systematic analysis of *Drosophila* TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. EMBO J. *30*, 3977–3993.

Hartig, J.V., Tomari, Y., and Förstemann, K. (2007). piRNAs—the ancient hunters of genome invaders. Genes Dev. *21*, 1707–1713.

Horvath, G.C., Kistler, M.K., and Kistler, W.S. (2009). RFX2 is a candidate downstream amplifier of A-MYB regulation in mouse spermatogenesis. BMC Dev. Biol. *9*, 63.

Hosokawa, M., Shoji, M., Kitamura, K., Tanaka, T., Noce, T., Chuma, S., and Nakatsuji, N. (2007). Tudor-related proteins TDRD1/MTR-1, TDRD6 and TDRD7/TRAP: domain composition, intracellular localization, and function in male germ cells in mice. Dev. Biol. *301*, 38–52.

Houwing, S., Kamminga, L.M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., Filippov, D.V., Blaser, H., Raz, E., Moens, C.B., et al. (2007). A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell *129*, 69–82.

Huang, H., Gao, Q., Peng, X., Choi, S.Y., Sarma, K., Ren, H., Morris, A.J., and Frohman, M.A. (2011). piRNA-associated germline nuage formation and spermatogenesis require MitoPLD profusogenic mitochondrial-surface lipid signaling. Dev. Cell *20*, 376–387.

Ipsaro, J.J., Haase, A.D., Knott, S.R., Joshua-Tor, L., and Hannon, G.J. (2012). The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. Nature *491*, 279–283.

Kawaoka, S., Hara, K., Shoji, K., Kobayashi, M., Shimada, T., Sugano, S., Tomari, Y., Suzuki, Y., and Katsuma, S. (2013). The comprehensive epigenome map of piRNA clusters. Nucleic Acids Res. *41*, 1581–1590.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat. Rev. Mol. Cell Biol. *10*, 126–139.

Kojima, K., Kuramochi-Miyagawa, S., Chuma, S., Tanaka, T., Nakatsuji, N., Kimura, T., and Nakano, T. (2009). Associations between PIWI proteins and TDRD1/MTR-1 are critical for integrated subcellular localization in murine male germ cells. Genes Cells *14*, 1155–1165.

Kumar, M., and Carmichael, G.G. (1998). Antisense RNA: function and fate of duplex RNA in cells of higher eukaryotes. Microbiol. Mol. Biol. Rev. *62*, 1415–1434.

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

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

Kutter, C., Brown, G.D., Gonçalves, A., Wilson, M.D., Watt, S., Brazma, A., White, R.J., and Odom, D.T. (2011). Pol III binding in six mammals shows conservation among amino acid isotypes despite divergence among tRNA genes. Nat. Genet. *43*, 948–955.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Latham, K.E., Litvin, J., Orth, J.M., Patel, B., Mettus, R., and Reddy, E.P. (1996). Temporal patterns of A-*myb* and B-*myb* gene expression during testis development. Oncogene *13*, 1161–1168.

Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363–367.

Lee, H.C., Gu, W., Shirayama, M., Youngman, E., Conte, D.J., Jr., and Mello, C.C. (2012). *C. elegans* piRNAs mediate the genome-wide surveillance of germline transcripts. Cell *150*, 78–87.

Li, X.C., and Schimenti, J.C. (2007). Mouse pachytene checkpoint 2 (trip13) is required for completing meiotic recombination but not synapsis. PLoS Genet. *3*, e130.

Li, C., Vagin, V.V., Lee, S., Xu, J., Ma, S., Xi, H., Seitz, H., Horwich, M.D., Syrzycka, M., Honda, B.M., et al. (2009a). Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell *137*, 509–521.

Li, X.C., Barringer, B.C., and Barbash, D.A. (2009b). The pachytene checkpoint and its relationship to evolutionary patterns of polyploidization and hybrid sterility. Heredity (Edinb) *102*, 24–30.

Mangan, S., and Alon, U. (2003). Structure and function of the feed-forward loop network motif. Proc. Natl. Acad. Sci. USA *100*, 11980–11985.

Mettus, R.V., Litvin, J., Wali, A., Toscani, A., Latham, K., Hatton, K., and Reddy, E.P. (1994). Murine A-myb: evidence for differential splicing and tissue-specific expression. Oncogene *9*, 3077–3086.

Modzelewski, A.J., Holmes, R.J., Hilz, S., Grimson, A., and Cohen, P.E. (2012). AGO4 regulates entry into meiosis and influences silencing of sex chromosomes in the male mouse germline. Dev. Cell *23*, 251–264.

Montgomery, M.K., Xu, S., and Fire, A. (1998). RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA *95*, 15502–15507.

Namekawa, S.H., and Lee, J.T. (2009). XY and ZW: is meiotic sex chromosome inactivation the rule in evolution? PLoS Genet. *5*, e1000493.

Nebel, B.R., Amarose, A.P., and Hacket, E.M. (1961). Calendar of gametogenic development in the prepuberal male mouse. Science *134*, 832–833.

Newburger, D.E., and Bulyk, M.L. (2009). UniPROBE: an online database of protein binding microarray data on protein-DNA interactions. Nucleic Acids Res. *37*(Database issue), D77–D82.

Nishimasu, H., Ishizu, H., Saito, K., Fukuhara, S., Kamatani, M.K., Bonnefond, L., Matsumoto, N., Nishizawa, T., Nakanaga, K., Aoki, J., et al. (2012). Structure and function of Zucchini endoribonuclease in piRNA biogenesis. Nature *491*, 284–287.

Oh, I.H., and Reddy, E.P. (1999). The myb gene family in cell growth, differentiation and apoptosis. Oncogene *18*, 3017–3033.

Osella, M., Bosia, C., Corá, D., and Caselle, M. (2011). The role of incoherent microRNA-mediated feedforward loops in noise buffering. PLoS Comput. Biol. *7*, e1001101.

Pane, A., Wehr, K., and Schüpbach, T. (2007). zucchini and squash encode two putative nucleases required for rasiRNA production in the *Drosophila* germline. Dev. Cell *12*, 851–862.

Reuter, M., Chuma, S., Tanaka, T., Franz, T., Stark, A., and Pillai, R.S. (2009). 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.

Reuter, M., Berninger, P., Chuma, S., Shah, H., Hosokawa, M., Funaya, C., Antony, C., Sachidanandam, R., and Pillai, R.S. (2011). Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. Nature *480*, 264–267.

Ro, S., Park, C., Song, R., Nguyen, D., Jin, J., Sanders, K.M., McCarrey, J.R., and Yan, W. (2007). Cloning and expression profiling of testis-expressed piRNA-like RNAs. RNA *13*, 1693–1702.

Robine, N., Lau, N.C., Balla, S., Jin, Z., Okamura, K., Kuramochi-Miyagawa, S., Blower, M.D., and Lai, E.C. (2009). A broadly conserved pathway generates 3'UTR-directed primary piRNAs. Curr. Biol. *19*, 2066–2076.

Romanienko, P.J., and Camerini-Otero, R.D. (2000). The mouse Spo11 gene is required for meiotic chromosome synapsis. Mol. Cell *6*, 975–987.

Saito, K., Inagaki, S., Mituyama, T., Kawamura, Y., Ono, Y., Sakota, E., Kotani, H., Asai, K., Siomi, H., and Siomi, M.C. (2009). A regulatory circuit for piwi by the large Maf gene *traffic jam* in *Drosophila*. Nature *461*, 1296–1299.

Schoenmakers, S., Wassenaar, E., Hoogerbrugge, J.W., Laven, J.S., Grootegoed, J.A., and Baarends, W.M. (2009). Female meiotic sex chromosome inactivation in chicken. PLoS Genet. *5*, e1000466.

Seitz, H., Ghildiyal, M., and Zamore, P.D. (2008). Argonaute loading improves the 5′ precision of both MicroRNAs and their miRNA* strands in flies. Curr. Biol. *18*, 147–151.

Shen-Orr, S.S., Milo, R., Mangan, S., and Alon, U. (2002). Network motifs in the transcriptional regulation network of *Escherichia coli*. Nat. Genet. *31*, 64–68.

Shirayama, M., Seth, M., Lee, H.C., Gu, W., Ishidate, T., Conte, D.J., Jr., and Mello, C.C. (2012). piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. Cell *150*, 65–77.

Smith, J.M., Bowles, J., Wilson, M., Teasdale, R.D., and Koopman, P. (2004). Expression of the tudor-related gene Tdrd5 during development of the male germline in mice. Gene Expr. Patterns *4*, 701–705.

Thomson, T., and Lin, H. (2009). The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. Annu. Rev. Cell Dev. Biol. *25*, 355–376.

Toscani, A., Mettus, R.V., Coupland, R., Simpkins, H., Litvin, J., Orth, J., Hatton, K.S., and Reddy, E.P. (1997). Arrest of spermatogenesis and defective breast development in mice lacking A-*myb*. Nature *386*, 713–717.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

Trauth, K., Mutschler, B., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., and Klempnauer, K.H. (1994). Mouse A-*myb* encodes a trans-activator and is expressed in mitotically active cells of the developing central nervous system, adult testis and B lymphocytes. EMBO J. *13*, 5994–6005.

Vagin, V.V., Klenov, M.S., Kalmykova, A.I., Stolyarenko, A.D., Kotelnikov, R.N., and Gvozdev, V.A. (2004). The RNA interference proteins and vasa locus are involved in the silencing of retrotransposons in the female germline of *Drosophila melanogaster*. RNA Biol. *1*, 54–58.

Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. Science *313*, 320–324.

Vagin, V.V., Wohlschlegel, J., Qu, J., Jonsson, Z., Huang, X., Chuma, S., Girard, A., Sachidanandam, R., Hannon, G.J., and Aravin, A.A. (2009). Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. Genes Dev. *23*, 1749–1762.

Vasileva, A., Tiedau, D., Firooznia, A., Müller-Reichert, T., and Jessberger, R. (2009). Tdrd6 is required for spermiogenesis, chromatoid body architecture, and regulation of miRNA expression. Curr. Biol. *19*, 630–639.

Vourekas, A., Zheng, Q., Alexiou, P., Maragkakis, M., Kirino, Y., Gregory, B.D., and Mourelatos, Z. (2012). Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. Nat. Struct. Mol. Biol. *19*, 773–781.

Watanabe, T., Chuma, S., Yamamoto, Y., Kuramochi-Miyagawa, S., Totoki, Y., Toyoda, A., Hoki, Y., Fujiyama, A., Shibata, T., Sado, T., et al. (2011). MITOPLD is a mitochondrial protein essential for nuage formation and piRNA biogenesis in the mouse germline. Dev. Cell *20*, 364–375.

Weston, K. (1992). Extension of the DNA binding consensus of the chicken c-Myb and v-Myb proteins. Nucleic Acids Res. *20*, 3043–3049.

Yabuta, Y., Ohta, H., Abe, T., Kurimoto, K., Chuma, S., and Saitou, M. (2011). TDRD5 is required for retrotransposon silencing, chromatoid body assembly, and spermiogenesis in mice. J. Cell Biol. *192*, 781–795.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhang, Z., Theurkauf, W.E., Weng, Z., and Zamore, P.D. (2012). Strand-specific libraries for high throughput RNA sequencing (RNA-Seq) prepared without poly(A) selection. Silence *3*, 9.