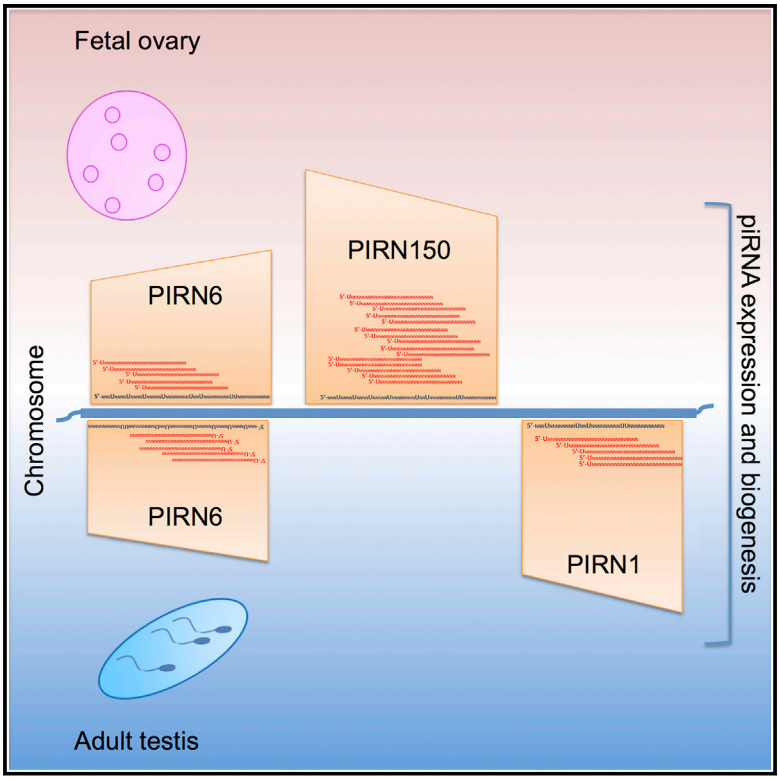
Resource

Cell Reports

Discovery and Characterization of piRNAs in the Human Fetal Ovary

Graphical Abstract



Highlights

- 150 piRNA genes were identified in human adult testis and fetal ovary
- Adult testis and fetal ovary differentially express piRNA genes
- piRNAs were not identified in non-germline tissue
- Non-repetitive piRNA genes account for over 90% of piRNAs

Authors

Zev Williams, Pavel Morozov, Aleksandra Mihailovic, ..., Stefan Juranek, Zev Rosenwaks, Thomas Tuschl

Correspondence

zev.williams@einstein.yu.edu (Z.W.), ttuschl@mail.rockefeller.edu (T.T.)

In Brief

Williams et al. identify and characterize 150 piRNA genes in the adult human testis and fetal ovary. Expression of piRNA genes differs between the sexes, and repetitive elements, including those for retrotransposons, are underrepresented among piRNA genes. No piRNA expression was detected in somatic tissue.

Williams et al., 2015, Cell Reports *13*, 854–863
October 27, 2015
http://dx.doi.org/10.1016/j.celrep.2015.09.030

CellPress

OPEN ACCESS
CellPress

Cell Reports
Resource

Discovery and Characterization of piRNAs in the Human Fetal Ovary

Zev Williams,[1,2,4,*] Pavel Morozov,[2,4] Aleksandra Mihailovic,[2] Carolina Lin,[2] Pavan Kumar Puvvula,[2] Stefan Juranek,[2] Zev Rosenwaks,[3] and Thomas Tuschl[2,*]

[1]The Program for Early and Recurrent Pregnancy Loss (PEARL), Department of Obstetrics & Gynecology and Women's Health, The Albert Einstein College of Medicine, Bronx, NY 10461, USA

[2]Howard Hughes Medical Institute and Laboratory for RNA Molecular Biology, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA

[3]The Ronald O. Perelman and Claudia Cohen Center for Reproductive Medicine, Weill Cornell Medical College, New York, NY 10021, USA

[4]Co-first author

*Correspondence: zev.williams@einstein.yu.edu (Z.W.), ttuschl@mail.rockefeller.edu (T.T.)

http://dx.doi.org/10.1016/j.celrep.2015.09.030

SUMMARY

Piwi-interacting RNAs (piRNAs), a class of 26- to 32-nt non-coding RNAs (ncRNAs), function in germline development, transposon silencing, and epigenetic regulation. We performed deep sequencing and annotation of untreated and periodate-treated small RNA cDNA libraries from human fetal and adult germline and reference somatic tissues. This revealed abundant piRNAs originating from 150 piRNA-encoding genes, including some exhibiting gender-specific expression, in fetal ovary and adult testis—developmental periods coinciding with mitotic cell divisions expanding fetal germ cells prior to meiotic divisions. The absence of reads mapping uniquely to annotated piRNA genes demonstrated their paucity in fetal testis and adult ovary and absence in somatic tissues. We curated human piRNA-expressing regions and defined their precise borders and observed piRNA-guided cleavage of transcripts antisense to some piRNA-producing genes. This study provides insights into sex-specific mammalian piRNA expression and function and serves as a reference for human piRNA analysis and annotation.

INTRODUCTION

Small non-coding RNAs (ncRNAs) of 20 to 32 nt in length assemble into ribonucleoprotein effector complexes that target partially complementary RNAs, including mRNAs, and subsequently trigger diverse biological functions (Farazi et al., 2008; Ghildiyal and Zamore, 2009; Malone and Hannon, 2009). These small-RNA-bound effector complexes comprise members of either Argonaute (EIF2C/AGO) or Piwi (PIWIL) proteins (Cenik and Zamore, 2011; Kawamura et al., 2008; Meister, 2013). The human genome encodes four Piwi proteins, PIWIL1/HIWI, PIWIL2/HILI, PIWIL3/HIWI3, and PIWIL4/HIWI2 (Bartel, 2009; Sasaki et al., 2003), while the mouse genome only encodes three Piwi proteins, PIWIL1/MIWI, PIWIL2/MILI, and PIWIL4/MIWI2 (Kuramochi-Miyagawa et al., 2001). Mammalian PIWIL proteins associate with 26- to 32-nt Piwi-interacting RNAs (piRNAs), while members of the AGO subfamily associate with 20- to 23-nt microRNAs (miRNAs) and small interfering RNAs (siRNAs) (Bartel, 2009; Le Thomas et al., 2014; Meister, 2013).

piRNAs are a distinct class of small RNAs that function in germline development, transposon silencing, and epigenetic regulation (Ashe et al., 2012; Houwing et al., 2008; Juliano et al., 2011; Malone and Hannon, 2009; Pillai and Chuma, 2012; Ross et al., 2014; Vagin et al., 2006; Rajasethupathy et al., 2012). piRNAs have remarkable diversity, with over 1.5 million unique sequences in *Drosophila* and tens of thousands of unique sequences in mammals (Aravin et al., 2006; Kawamura et al., 2008). piRNAs are distinguished from other mammalian small silencing RNAs by their longer length (26–32 nt) and 2′-*O*-methyl modification by HENMT1 methylase to their 3′ end (Horwich et al., 2007; Ishizu et al., 2012; Kirino and Mourelatos, 2007a, 2007b). The 5′ nucleotide of piRNAs is phosphorylated and is typically a uridine (5′ U) (Gunawardane et al., 2007). piRNAs are produced by RNase-III-independent processing of RNA Pol II primary piRNA transcripts (Ishizu et al., 2012; Li et al., 2013). These transcripts of up to 200 kb in length are generated from over 100 genes (Le Thomas et al., 2014)

In contrast to miRNAs that are expressed in all cell types, piRNA expression in mammals is largely restricted to the germline. piRNAs have been described in the male and female germline of *Caenorhabditis elegans*, *Drosophila melanogaster*, *Danio rerio*, and *Xenopus laevis* but, notably, only in the male germline of mammals and birds (Ashe et al., 2012; Ha et al., 2014; Houwing et al., 2008; Juliano et al., 2011; Lau et al., 2009; Li et al., 2013; Wilczynska et al., 2009; Yang et al., 2013). Within mouse testis, piRNAs are highly expressed; the total number of piRNAs are 10-fold higher than miRNAs with two million molecules per cell in mouse testis (Aravin et al., 2006; Beyret et al., 2012; Girard et al., 2006; Kawamura et al., 2008; Lau et al., 2006). In mouse, the piRNA pathway is required during germline reprogramming for re-establishment of epigenetic silencing and post-transcriptional cleavage of long interspersed elements (LINE-1) and intracisternal A-particle (IAP) sequences (Aravin et al., 2008, 2007; De

Fazio et al., 2011; Kuramochi-Miyagawa et al., 2008) and subsequent maintenance of LINE-1 silencing of spermatogonia, meiocytes, and spermatids (Di Giacomo et al., 2013, 2015; Reuter et al., 2011). Deficiency of genes required for piRNA biogenesis (e.g., PIWIL proteins MILI and MIWI2, the phospholipase MITOPLD [a phospholipase D family member], helicase MOV10L1, the piRNA biogenesis factor MAEL, and the Tudor domain protein TDRD9) is characterized by arrest of spermatogenesis and infertility in males (Carmell et al., 2007; Castañeda et al., 2014; Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2004; Shoji et al., 2009; Vasileva et al., 2009; Watanabe et al., 2011; Zheng and Wang, 2012). In contrast, females with deficiency in these genes remain fertile and without an apparent phenotype. An interesting exception is the recently described MAEL$^{-/-}$ female that displays increased fetal oocyte attrition correlated with increased LINE1 expression (Malki et al., 2014). The apparent absence of piRNAs and PIWIL proteins in mammalian ovaries has forced the question of how the putative function of piRNAs and PIWIL proteins could be carried out in the ovary in their absence.

Piwi protein and piRNA expression within the testis is under tight developmental regulation; for example, in mouse, Mili is present in mitotic spermatogonia and then absent from developing sperm in the preleptotene and leptotene stages before being present again in late zygotene until early round spermatids (Beyret and Lin, 2011; Di Giacomo et al., 2013). While these stages of spermatogenesis continuously occur in the adult testis, the corresponding stages of oogenesis occur only in the human fetal ovary and only up to 20 weeks gestation (Hartshorne et al., 2009). We therefore investigated whether piRNAs and PIWIL proteins were expressed in the human *fetal* ovary.

We performed deep sequencing of RNA extracted from human fetal and adult testis and ovary and somatic tissues and used bioinformatic analysis of small RNA libraries, PIWIL-protein immunoprecipitations, biochemical modifications, and manual curation to define and characterize human piRNA genes in the testis and ovary. We identified piRNA genes with abundant expression in human fetal ovaries but nearly undetectable expression in adult ovaries. Subsets of piRNA genes are either expressed in both human fetal ovaries and adult testis or are uniquely expressed in either ovaries or testis. The accurate annotation of piRNAs and their genes facilitates the design of functional studies and highlights the extraordinary specificity of piRNA expression in germline tissues.

RESULTS

Preparation of Small RNA cDNA Libraries

We generated barcoded 19- to 35-nt short RNA cDNA libraries from 31 different tissue sources or samples prepared from 19 subjects consisting of 16- to 23-week gestational age fetal ovaries and testis and adult ovaries and testis, as well as fetal brain, heart, and liver. Sample information and respective small RNA composition are summarized in Table S1. Collectively, we obtained 188,438,956 small RNA cDNA sequence reads, which were first annotated against public databases, while piRNAs were identified by an iterative annotation and curation process as detailed below.

Genomic Characteristics of piRNA Regions

Defined characteristics of piRNAs include a predominant length of 26–32 nt, presence of a 5′ phosphate, a 5′ U, and clustering of reads to distinct genomic locations. To identify piRNA genes, we used a seed-building approach and aligned all reads to the genome (hg19) and then computed length and 5′ U content of reads appearing in clusters. Cutoff values were established based on well-annotated and well-covered piRNA genes and gene borders were individually curated as described in the methods section. We identified 150 piRNA genes including 42 genes comprising sense and antisense transcripts arising from one central divergent or two convergent promoters. Together, these genes collect 31% and 2.5% of total reads in adult testis and fetal ovary, respectively (Tables S1 and S2). We named the genes PIRN1 to 150, and transcripts arising from bidirectional-expressed genes are indicated as PIRN#-S or AS (# symbol referring to gene number and S and AS referring to sense and antisense), whereby gene numbers were assigned according to the rank of piRNA read coverage in adult testis followed by fetal ovary. These piRNA genes were distributed across all 22 autosomal chromosomes and the X chromosome, but not the Y chromosome (Figure 1A), as also observed for marmoset (Hirano et al., 2014). Chromosomes 9 and 15 together yielded 34% of all reads. 72% of piRNA genes were expressed from unidirectional-promoter-containing genes, while 28% of piRNA genes were expressed from bidirectional-promoter-containing genes.

Comparing total piRNA to total miRNA read counts, adult testis show a ratio of ~1:1, while fetal ovary displays a ratio of 1:20, respectively (Table S1). The actual ratio of piRNAs to miRNAs is ~10-fold higher, considering the bias a 2′-*O*-methyl group introduces for 3′ adaptor ligation during cDNA library preparation (our unpublished data; Ebhardt et al., 2005). The abundance of piRNAs in fetal testis compared to adult testis and in adult ovary compared to fetal ovary was ~25-fold lower in both instances, and piRNAs were undetectable in fetal brain, fetal heart, or adult liver (Figures 1B and 1C).

piRNAs are protected from periodate oxidation by their terminal 2′-*O*-methyl modification, while RNAs with a terminal 2′,3′-hydroxyl are modified and depleted during cDNA library preparation. Periodate treatment of total RNA resulted in 5- to 100-fold enrichment of piRNAs over miRNAs in fetal and adult ovary or testis small RNA cDNA libraries (Table S3). The functionality of our annotated piRNAs is further substantiated by their enrichment in Piwi protein immunoprecipitates using lysates prepared from adult testis. Small RNA cDNA libraries prepared from immunoprecipitates of anti-PIWIL1/HIWI, PIWIL2/HILI, and PIWIL4/HIWI2 displayed a 2- to 4-fold enrichment in piRNAs over miRNAs compared to cell lysate (Table S1).

The mean length of a human piRNA gene was 25,284 nt (median 14,385), with a range of 570–359,021 nt and up to 38,944 unique piRNA reads per gene (Table S2). The genomic coordinates of piRNA genes comprise poorly annotated genomic regions (73.5%), already annotated piRNA-producing non-coding regions (13.8%), and protein-coding regions (0.8% CDS, 3.6% as 5′ UTR, 2.8% 3′ UTR, 5.5% intronic) (hg19, GRCh37, Known Genes annotation track) (Figures S1A and S1B).

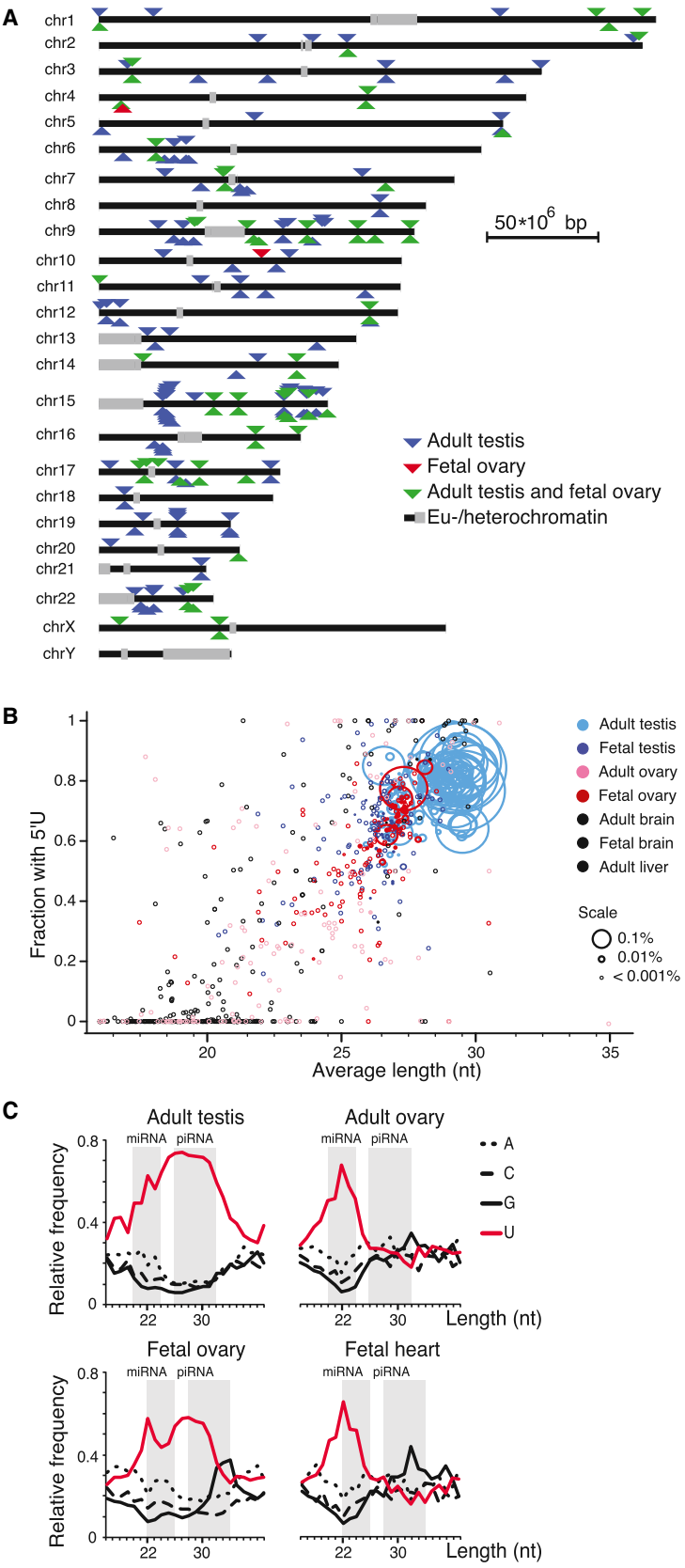


Figure 1. piRNA Expression in Fetal Ovary and Adult Testis

(A) Location of piRNA genes on human chromosomes (black bars). Telomeric and centromeric regions are drawn to scale (gray bars). piRNA regions expressed from the positive DNA strand are shown as triangles above the respective chromosome, and piRNA regions expressed from the negative DNA strand are shown as triangles below the respective chromosome. The color of the triangle indicates whether the region is expressed exclusively in adult testis (blue), fetal ovary (red), or in both testis and ovary (green). Scale bar is indicated.

(B) For each piRNA gene, the average read length (nt) and percentage of reads beginning with a uridine are shown with the circle area proportional to total small RNA reads from adult testis (light blue), fetal testis (green), adult ovary (pink), fetal ovary (red), adult brain, fetal brain, and adult liver (all indicated in black). Scale is indicated. A higher proportion of reads from fetal ovary and adult testis have a longer read length and a higher percentage of reads beginning with a 5′ U than do reads from adult and fetal brain, adult liver, adult ovary, and fetal testis.

(C) Relative frequency of each of the four nucleotides (adenine, dotted line; cytosine, dashed line; guanine, solid black line; and uridine, red line) in the 5′ position for RNA reads of the indicated nucleotide length on the x axis. Regions corresponding to miRNAs and piRNAs are shaded gray. While all tissues displayed a preference for a 5′ U in reads of ~21 nt in length, corresponding to miRNAs, only fetal ovary and adult testis displayed a second peak in reads of 26–32 nt in length. This second peak corresponds to piRNAs. Only untreated samples are included. For periodate-treated samples, see Figure S3.

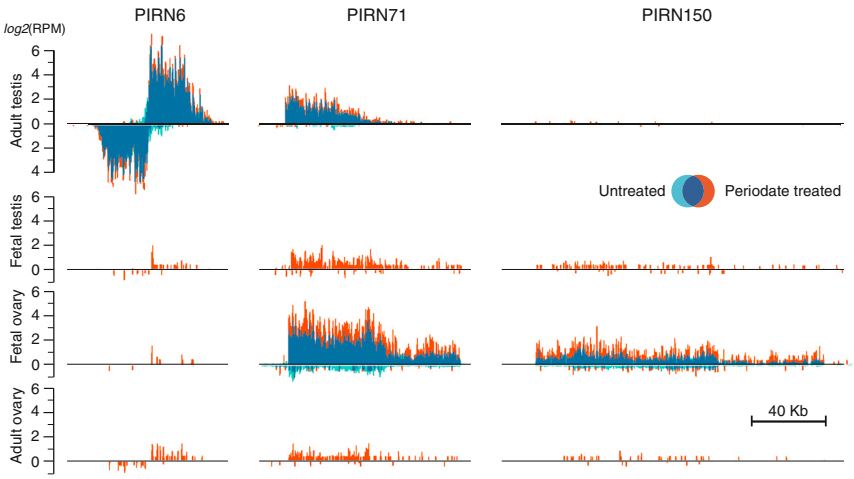


Figure 2. Representative Examples of Expression of piRNA Clusters in Adult Testis and Fetal Ovary

Mapping of untreated (blue) and periodate-treated (orange) small RNA sequence reads (19–35 nt) from adult testis, fetal testis, fetal ovary, and adult ovary for three piRNA genes (genomic coordinates indicated according to Hg19). Areas of overlap between untreated and periodate-treated reads are indicated via shading. Illustrated are PIRN6, a bidirectional piRNA cluster expressed exclusively in the adult testis; PIRN71, a piRNA cluster expressing piRNAs in both adult testis and fetal ovary but with expanded borders in the fetal ovary; and a PIRN150, a unidirectional piRNA cluster expressed exclusively in fetal ovary. Read coverage (in reads per million) is displayed in log2 scale for each of the four tissue samples. RPKM values are calculated using total library reads, and coverage is averaged over a 200-nt sliding window. A scale bar is shown.

Adult Testis and Fetal Ovary Differentially Express piRNA Genes

To calculate piRNA gene expression levels, we only considered bona fide piRNAs, which we required to be ≥26 nt in length, begin with a 5′ U, and map to ten or fewer genomic locations. Reads mapped to more than ten genomic locations often resulted in unexpectedly high read coverage spikes surrounded by otherwise evenly distributed piRNA reads. Such spikes typically correspond to pseudogenes or interspersed low-complexity sequences within genes. The reads per kilobase of transcript per million mapped reads (RPKM) of a piRNA gene was calculated by dividing by its length and the total number of bona fide piRNA reads of the library.

A subset of piRNA genes is differentially expressed between fetal ovary and adult testis. 97 piRNA genes were expressed exclusively in the testis, two piRNA genes were expressed exclusively in the ovary, and 51 piRNA genes were expressed in both (Figures 1A and 2; Table S1). In adult testis, the piRNA gene with the highest piRNA total read density was PIRN1, which was expressed at 11,396 RPKM but only at 8 RPKM in fetal ovary. Conversely, the most expressed piRNA gene in fetal ovary was PIRN71, which was expressed at 6,148 RPKM but only at 87 RPKM in adult testis. In some cases, the length of a piRNA gene from which piRNAs were expressed also differed between the sexes. For example, the PIRN71 gene displayed fetal ovary piRNAs across 92,421 nt, but in testis, expression was confined to the first 60,000 nt.

Distinct Sex-Specific piRNA Size Distributions

Testicular piRNAs were, on average, 1–2 nt longer in adult testis than in fetal ovary (29.6 nt versus 28.4 nt), suggesting association of piRNAs to differentially expressed PIWIL proteins. This was particularly apparent for piRNAs originating from genes uniformly expressed in both sexes, such as PIRN4.

To determine whether the gender differences in piRNA sizes were associated with differences in PIWIL gene expression, we performed poly(A) RNA sequencing (RNA-seq) on fetal and adult testis, ovary, liver, and brain. While we readily observed expression of PIWIL1, 2, and 4, expression of PIWIL3 was at least 10-fold lower than the least abundant PIWIL transcript in the same sample (Table S4). PIWIL1 and PIWIL2 were specifically expressed in ovary and testis, but not somatic tissue, while PIWIL4 appeared more ubiquitously expressed. PIWIL1 is over 10-fold-higher expressed in adult testis compared to its family members, while PIWIL2 is over 3-fold-higher expressed in fetal ovary compared to its members.

In order to assess if the length differences are a function of PIWIL protein association, we performed immunoprecipitation from human testis cell lysates using antibodies specific to the three major PIWIL proteins followed by small RNA cDNA library preparation and sequencing. We observed an average size distribution of 26 nt for PIWIL2, 28 nt for PIWIL4, 29 nt for PIWIL1, and 29 nt for piRNAs present in unfractionated testis tissue lysates. The dominance of long (29 nt) piRNA-binding PIWIL1 protein in adult testis and short (26 nt) piRNA-binding PIWIL2 in fetal ovary, respectively, are therefore in agreement with the observed differences in piRNA size distribution between sexes.

Absence of piRNAs and piRNA Biogenesis Factors in Non-germline Tissues

While piRNAs were clearly detectable in fetal ovary and adult testis, non-germline tissues completely lacked piRNA expression despite similar (within a factor of 10) PIWIL4 expression levels in these tissues. This prompted us to evaluate the expression of piRNA biogenesis factors across the same tissues (Gerstberger et al., 2014). While subsets of PIWIL genes and general piRNA biogenesis factors were well expressed in adult testis and slightly less in fetal ovaries, piRNA biogenesis factors such as MAEL and HENMT1 parallel the also reduced abundance of PIWIL1 and 2 proteins (Table S4). These observations are in agreement with the reduced piRNA abundance in meiotically inactive fetal testis and adult ovary and also the absence of piRNAs in non-germline tissue samples.

Repetitive Sequence piRNAs and Their Corresponding Genomic Locations

piRNA genes may be composed of predominantly unique sequence, predominantly repetitive sequence, or a combination

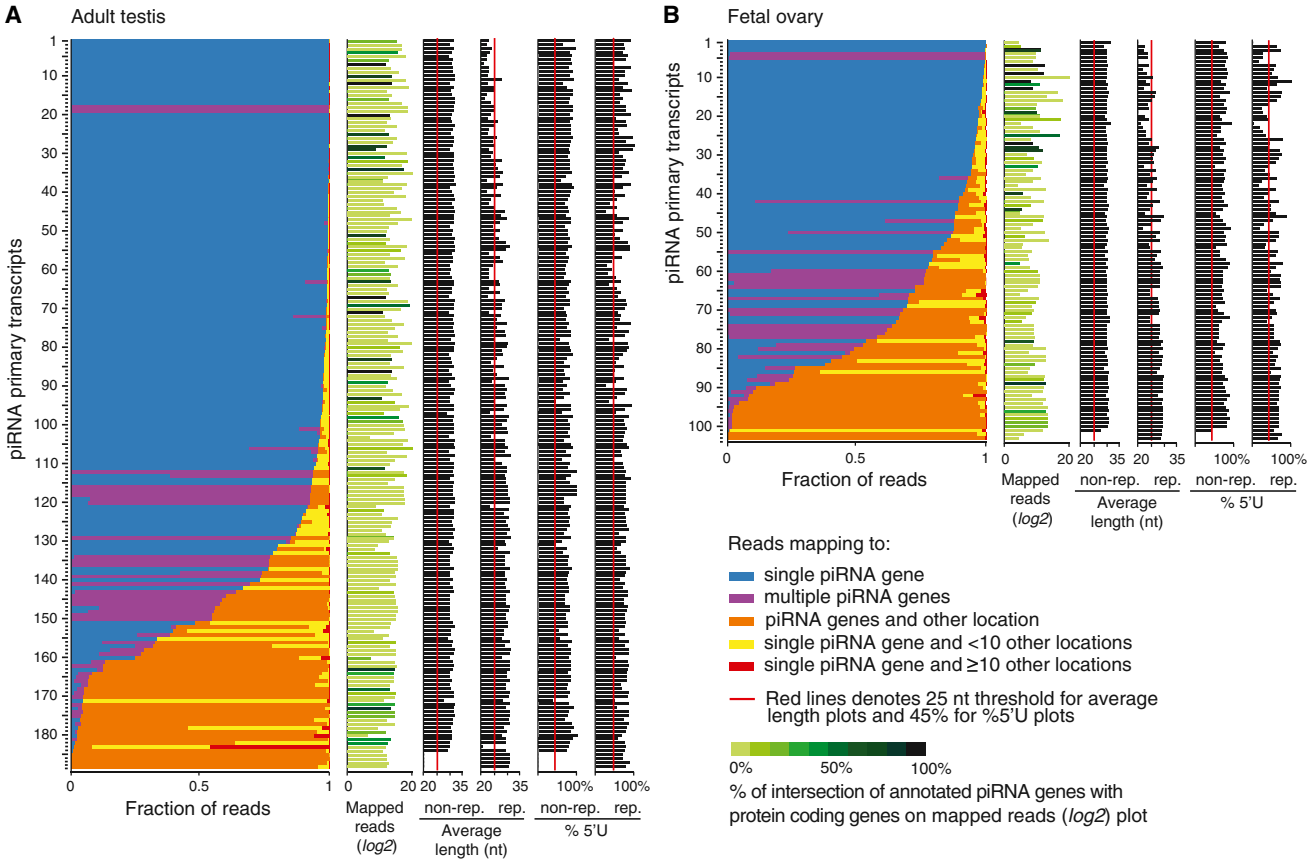


Figure 3. piRNA Genes Can Be Classified as Single Mappers or Multi-mappers

Reads from each piRNA gene within adult testis (A) and fetal ovary (B) were mapped against the human genome (Hg19). Reads were then categorized as either mapping to the particular piRNA gene (blue), multiple piRNA genes (purple), multiple piRNA genes and non-piRNA locations (orange), the particular piRNA gene and <10 non-piRNA locations (yellow), and the particular piRNA gene and >10 non-piRNA locations (red). The percentage of each category is shown. piRNA genes are ordered vertically according to increasing percentages of multi-mappers. An expanded version of this figure with the individual gene names is provided (Figure S2). For each piRNA gene, the total read coverage (in log2) scale is shown. The percentage of these reads mapping to known genes is indicated by increasing green shading. There was a much wider range in the proportion of piRNA reads aligning to individual piRNA genes in ovary compared to testis (56.9% versus 10.9%, respectively), suggesting that in ovary, a few genes constituted the primary source of piRNAs, whereas piRNAs were more equally mapped across the piRNA genes in testis. The average read length for reads from adult testis and fetal ovary mapping each piRNA gene is shown, further classified as either non-repetitive or repetitive reads. A vertical red line indicated a threshold for piRNA-like characteristics (length > 25 nt or %5′ U > 55). Non-repetitive reads mapping to predominantly repetitive and non-repetitive piRNA genes display piRNA-like characteristics (average length > 25 nt, %5′ U). In contrast, among the predominantly repetitive reads, only those mapping to repetitive, but not non-repetitive, piRNA genes display the same piRNA-like characteristics.

of both (Figures 3 and S2). Sequence analysis indicated that 88 piRNA genes were defined by sequence reads that predominantly mapped to a single piRNA gene while 62 piRNA genes were characterized by sequence reads that predominantly mapped to multiple genomic locations, including piRNA and non-piRNA genes. Non-repetitive piRNA genes account for 94.2% and 90.5% piRNAs in adult testis and fetal ovary, respectively (Figure 3, blue). The remaining repetitive-sequence piRNAs fall into three classes: those arising from piRNA gene expansion (Figure 3, purple), those resulting from duplication and insertion of retrotransposed and low-complexity repetitive sequences (Figure 3, yellow and red), and those arising from a combination of both (Figure 3, orange).

cDNA libraries prepared without periodate treatment contained many reads mapping to repeat regions displaying an average read length of 23.8 nt for adult testis and 22.6 nt for fetal ovary and lacked enrichment of 5′ U, indicating that most repetitive sequence reads were not bona fide piRNAs and presumably originated from other abundant transcripts also comprising these repeats. To minimize the impact of repeat sequence reads in defining piRNA expression, we required piRNAs to be ≥26 nt and having a 5′ U. Alternatively, if no filter is applied, the impact of repeat-sequence reads can be minimized by dividing the read count by the number of matching genomic locations for that read.

Transposon-annotated repetitive sequences are present throughout the genome but are not enriched in piRNA genes. Among those repetitive sequences present in piRNA genes are endogenous retroviral elements (ERVL), Alu family elements (Alu), mammalian-wide interspersed repeats (MIR) and long interspersed elements (LINEs, L1) (Table S1). Although piRNAs are often discussed in the context of controlling transposons,

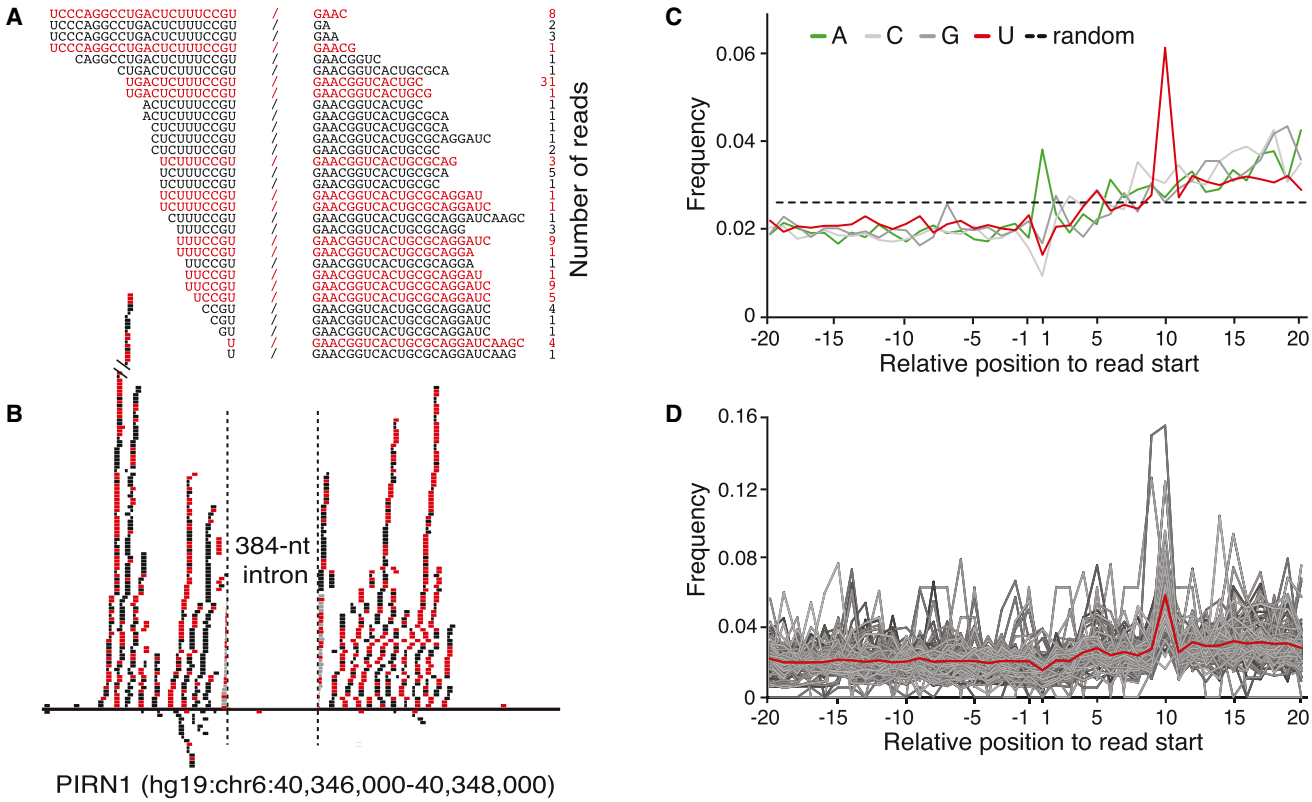


Figure 4. Alignment of Sequences Reveals Evidence of piRNA Biogenesis Post-splicing of Primary Transcript and Complimentary Strand Processing

(A) Alignment of individual reads to PIRN1 containing a 384 nt intron in reveals perfect matches spanning the intron. Reads that beginning with a 5′ Uridine and are ≥26 nt in length are shown in red. Reads that either do not begin with a 5′ U and/or are <26 nt in length are shown in gray. Sequence alignment of piRNAs crossing the exon-exon boundaries contain conventional U2-type donor site consensus sequence GU/GUGG and canonical acceptor consensus sequence AG/G within the piRNA sequence.

(B) Alignment of reads to the full PIRN1 region shows dense coverage of the region flanking the intronic sequence but no reads aligning to the intron itself. Reads that are ≥26 nt in length and begin with a 5′ U are shown in red. Reads that are <26 nt in length or do not begin with a 5′ U are shown in gray.

(C) Relative frequency of specific nucleotide (adenine, green; cytosine, light gray; guanine, dark gray; and uridine, red) at 5′ position of read on opposite strand for each read within piRNA regions from adult testis.

(D) The relative frequency of the starting position for reads derived from the antisense strand relative to the 5′ end of the piRNA were compared for all antisense strands beginning −20 to +20 nt relative to the start of the piRNA. Specific values for individual piRNA regions are shown in gray shaded lines. The red line represents the average value.

the ratios of actual repeat numbers in piRNA genes compared to genomic average were 0.95 for ERVL, 0.39 for Alu, 0.4 for MIR, 0.25 for L1, and 0.16 for CR1-like, indicating an underrepresentation of repeats across piRNA genes, as previously observed (Aravin et al., 2003, 2006; Landgraf et al., 2007; Saito et al., 2006).

piRNA Biogenesis and piRNA-Guided Target RNA Cleavage Mechanisms

Human piRNAs are processed from longer primary transcripts. Some of these primary transcripts have also been annotated as spliced ESTs. For those 19 piRNA genes with high read coverage annotated as spliced ESTs, mapping of piRNAs demonstrated that their piRNAs solely originated from exonic regions of their respective spliced primary piRNA transcripts. This included piRNAs mapping across exon-exon junctions (Figures 4A and 4B), consistent with earlier observations in *D. melanogaster* and mice (Aravin et al., 2001, 2003; Girard et al., 2006; Ishizu et al., 2012; Li et al., 2013; Vagin et al., 2006).

Alignment of reads to hg19 revealed evidence of piRNA-guided antisense RNA cleavage for approximately one-third of piRNA genes in adult testis. Antisense-orientation sequence reads showed a strong bias for an adenosine at the tenth nucleotide position (A10), and they represented up to 27% of all reads mapping to these piRNA genes. Since the majority of piRNAs started with a 5′ U and piRNA-guided target RNA cleavage always occurs between antisense nucleotides across piRNA residues 10 and 11, such bias is anticipated (e.g., Figures 4C and 4D). The length of these piRNA-guided antisense RNA cleavage fragments was >25 nt. In contrast to primary piRNAs, these antisense RNAs were depleted ~100-fold upon periodate treatment, indicating that they were not 2′-*O* methylated.

Similarly, in fetal ovary, 17 out of 53 piRNA genes demonstrated antisense cleavage, with only 4 genes revealing antisense

RNAs > 25 nt average read lengths. These antisense RNAs were periodate-treatment sensitive analogous to adult testis RNA. Collectively, this supports a model of piRNA-guided cleavage of the complimentary strand but does not support a ping-pong piRNA model involving every piRNA gene as described for *D. melanogaster* (Gunawardane et al., 2007; Brennecke et al., 2007). Somatic tissues, which lack piRNAs, also lacked short antisense RNAs of length > 25 nt, indicating that their processing was indeed piRNA guided.

DISCUSSION

We characterized piRNA features and gene expression in the human male and female gonads. 150 piRNA genes were identified, 99 of which were differentially expressed between testis and ovary. piRNA expression was largely limited to the adult testis and fetal ovary. While adult testis and fetal ovary represent widely differing periods relative to chronological age, germ cells in adult testis and fetal ovaries are at analogous stages of gametogenesis. The low abundance of piRNAs in the adult ovary may be because the oocytes, all of which have now completed meiosis I, are arrested and awaiting post-meiotic differentiation following hormonal stimulation. Alternatively, it may reflect the low number of germ cells relative to surrounding stromal cells. The paucity of piRNAs in fetal testis is likely attributable to the arrest of spermatogenesis at the spermatogonia stage (Beyret and Lin, 2011). While piRNAs have been observed in fetal testis from mice, these fetal testis were obtained from embryonic day 15.5 (E15.5) to E18.5, at which point the fetal testis would be in the mitotic stage of fetal testis development, whereas our testis were obtained from a 20 weeks, 3 days gestation fetus, at which point the fetal testis would be in the quiescent phase of fetal testis development.

Prior attempts at cataloging piRNAs have relied on deep sequencing of cDNA libraries prepared from short RNAs that co-immunoprecipitated with PIWIL proteins (Aravin et al., 2006; Girard et al., 2006). The individual sequences were then cataloged, for example in NCBI as "piRNA," in a fashion akin to what had been previously done with other small non-coding RNAs such as miRNAs (Kozomara and Griffiths-Jones, 2014). However, experimental background RNA arising from fragmentation of highly abundant RNAs, e.g., those derived from repetitive elements or incompletely annotated but abundant RNAs (e.g., tRNAs, rRNAs, Y RNAs, and small nuclear RNAs) was also erroneously classified as piRNAs. In addition, the individual piRNA sequence submission approach did not account for the unique biogenesis mechanism of piRNAs where a multi-kilobase-size piRNA primary transcript can yield thousands of unique piRNAs following cleavage at any single uridine residue along the transcript (90,154 in the case of PIRN6; Table S2). Depletion of 3′ unmodified RNAs by periodate treatment led to piRNA enrichment, and concomitant depletion of repetitive RNAs further improved annotation of piRNA genes and their borders.

To date, the NCBI database contains 32,045 sequences annotated as human piRNAs. However, of these, only 66% (20,295) aligned to our piRNA genes. 260 of the unaligned sequences represented fragments of rRNA, miRNA, tRNA, and small nucleolar RNAs (snoRNAs) (Figures S1C–S1G). The remaining 11,490 sequences were distributed across the genome without evidence of neighboring piRNA reads within 100 bp and are also likely to represent experimental background. The misannotation of NCBI-deposited piRNAs is apparent in small RNA sequence libraries prepared from tissues not expressing PIWIL proteins, such as fetal liver, where 57% of total reads annotated by NCBI as piRNAs also mapped to rRNAs (Figures S1C and S1D). Periodate treatment virtually eliminated reads mapping to falsely annotated piRNA genes, confirming that the majority of reads with repeat character lack 2′-*O*-methyl groups otherwise present on 3′ ends of piRNAs and that only tissues shown to express piRNA genes had enrichment of piRNAs with periodate treatment (Figure S3). Finally, falsely annotated piRNA genes frequently show read evidence in libraries prepared from tissues that otherwise do not contain bona fide piRNAs, such as adult liver (Figure S1G), further substantiating the misannotation of some repetitive sequence as piRNA or piRNA gene.

We observed piRNA-guided cleavage of transcripts with antisense orientation to a subset of piRNA genes through the overrepresentation of adenosine at the tenth nucleotide (Figure 4D). However, the majority of piRNA-producing genes are not associated with antisense transcripts, short antisense RNAs were depleted by periodate treatment (Figure S4), and the ratio of antisense A10 reads to sense U1 reads decreased following periodate treatment or PIWIL-IP (Figure S3C). Our findings indicate piRNA-guided cleavage activity of fully complimentary target RNAs (Reuter et al., 2011), but they does not support an amplification loop for piRNAs in human cells as predicted by the "ping-pong" model discovered in *D. melanogaster* (Gunawardane et al., 2007; Brennecke et al., 2007).

The presence of piRNAs in the female gonads, also observed by Roovers and colleagues (Roovers et al., 2015), provides insight into the possible roles of mammalian piRNAs and supports the potential for a conserved molecular function of piRNAs in gametogenesis between sexes. piRNAs have been implicated in silencing of transposable element placement of long-term methylation marks on DNA and guiding epigenetic reprogramming and polymerase II profiles (Rajasethupathy et al., 2012; Weick and Miska, 2014). In contrast to the population of piRNAs found in lower organism such as *D. melanogaster* and zebrafish, in which most piRNA reads map to repetitive elements (Aravin et al., 2003; Houwing et al., 2007; Saito et al., 2006), we found that only ~2% of piRNA sequences in human map to transposable elements. Instead, the majority of piRNAs (76%) map to intergenic regions. This is consistent with recent data from mouse and rooster (Girard et al., 2006; Li et al., 2013) and suggests that human piRNAs do not primarily function to suppress transposable elements.

For 82.4% of unique piRNA sequences, no fully complimentary sequence was found in the genome except for the piRNA genomic locus from where they originated. It is conceivable that, analogous to miRNAs, piRNAs may predominantly depend on their 5′ ends for target RNA recognition (Bartel, 2009). Considering the vast and unique sequence repertoire of piRNAs and their predominant localization to the nucleus, piRNAs may detect ubiquitous nascent primary RNA transcripts and possibly their transcribed genomic region, potentially to provide global gene

silencing or other nucleic-acid regulatory or localization functions during germline cell differentiation.

In summary, we provide characterization and annotation of human piRNA genes. Our data demonstrate that piRNAs are expressed in the fetal ovary, in addition to adult testis, and that certain piRNA genes are expressed exclusively in the testis or ovary. Moreover, even piRNA genes expressed in both the ovary and testis can express different regions of the same piRNA gene. This expands on earlier models that limited the roles for piRNAs to males. In contrast to *Drosophila*, predominant targeting of transposable elements and a ping-pong mechanisms of biogenesis are not observed for human piRNAs. Thus, the function of piRNAs in humans and their initial biogenesis remain elusive.

EXPERIMENTAL PROCEDURES

Fetal Tissue Samples

Fetal tissues were collected at the time of elective termination for non-medical indications at Columbia Presbyterian Hospital. All samples were de-identified at the time of collection. Exclusion criteria included any known or suspected fetal anomaly or maternal genetic diseases or fetal demise prior to the start of the procedure. Fetal tissue was collected and flash-frozen in liquid nitrogen within 5 min of delivery. All aspects of this study were reviewed and approved by the institutional review boards at Columbia Presbyterian Hospital and the Rockefeller University. RNA samples from adult tissue were obtained from Ambion (FirstChoice Total RNA).

RNA Isolation

1 ml TRIzol (Invitrogen) was added to 50 mg flash-frozen tissue, and tissue was mechanically disrupted using a Tissuelyser (QIAGEN). Total RNA was then purified using phenol-chloroform extraction. Total RNA was quantified using 260-nm spectrometry (Nanodrop) with 1 optical density (260 nm) unit corresponding to 40 μg total RNA.

Small RNA cDNA Library Preparation, Sequencing, and Data Analysis

Barcoded small RNA cDNA libraries were prepared from tissue-extracted total RNA as previously described using 2 μg per sample. Size fractionating during adaptor ligation was used to recover 19- to 35-nt fractions (Farazi et al., 2008; Ghildiyal and Zamore, 2009; Hafner et al., 2012). Amplified cDNA libraries were then sequenced with 50 cycles of single-end sequencing by synthesis (Illumina HiSeq 2000). We used an in-house computer pipeline to extract barcodes and align reads to the genome (hg19) and perform small RNA annotation (Farazi et al., 2011; Hafner et al., 2012). miRNA abundance was determined as the sum of all reads with two or fewer annotation mismatches as previously described (Bartel, 2009; Brown et al., 2013).

piRNA Discovery and Gene Border Definition

Small RNA cDNA sequence reads (19–35 nt in length) were filtered to exclude annotated snoRNA, rRNA, tRNA, and miRNA. The remaining reads were aligned to the human genome (Hg19). Group of reads located within 50 nt of each other were defined as a nucleating region. To identify potential piRNA regions, the properties of the nucleating regions were then graphed in three planes for (1) percentage of reads within the seed region beginning with uridine, (2) average length of reads within the region, and (3) number of reads within the seed region. Seed regions that clustered as outliers in all three qualities represented possible piRNA regions. These regions (n = 1,009) were then merged if they were located within 10 kb of each other and the merged regions were considered potential piRNA clusters. This pipeline identified 314 piRNA candidate genes. 209 piRNA genes had been previously cited in the literature, 195 of which overlapped with regions discovered using our seed-building approach (Rosenkranz and Zischler, 2012). We accepted those regions covered by even read distribution, enrichment with periodate treatment, and with support from immunoprecipitation of PIWIL ribonucleotide complexes as piRNA-producing genes. Due to limitations in available tissue, immunoprecipitation from fetal ovaries was not possible. Manual inspection was also used to precisely define borders of piRNA regions and clusters.

Periodate and β-Elimination of Total RNA to Enrich for piRNAs during Small RNA cDNA Library Preparation

Periodate treatment and β-elimination of total RNA was performed as described previously (Akbergenov et al., 2006; Alefelder et al., 1998). 20 μg total RNA was dried and resuspended in 17.5 μl borate buffer (4.38 mM $Na_2B_4O_7 \cdot 10H_2O$ and 50 mM H_3BO_3 [pH 8.6]). $NaIO_4$ was added (28.6 mM final), and the reaction was incubated for 10 min in the dark at 24°C. The reaction was quenched by the addition of glycerol (5% final) and further incubated for 10 min in the dark at 24°C. The RNA samples were concentrated to 5 μl in a SpeedVac (Eppendorf). For β-elimination, 50 μl of borate buffer (33.75 mM $Na_2B_4O_7 \cdot 10H_2O$ and 50 mM H_3BO_3 [pH adjusted by NaOH to pH 9.5]) was added and the samples were incubated for 90 min at 45°C. RNAs were ethanol precipitated and resuspended in H_2O. RNAs containing vicinal 2′ and 3′ diol groups react with NaIO4. β-Elimination shortens NaIO4-reacted RNA by 1 nt, leaving a 3′ monophosphate that prevents 3′ adaptor ligation during small RNA cDNA library preparation. Periodate-treatment-resistant 2′-*O*-methylated RNAs, on the other hand, remain susceptible to 3′ adaptor ligation. Periodate treatment therefore selectively enriches for piRNAs since they are 2′-*O*-methylated and thus resistant to periodate treatment and β-elimination, while miRNAs contain 2′ and 3′ terminal hydroxyl group and are reactive and altered to prevent 3′ adaptor ligation during library preparation.

Preparation of Adult Testis Lysate

20 g frozen adult testicles tissue samples from human specimens was obtained from a tissue cooperative network. Tissue junks were pulverized to obtain a tissue powder using a TissueLyser homogenizer (QIAGEN) in the presence of continuous liquid nitrogen. Powdered tissue samples from testicles were resuspended into two volumes of NP-40 buffer (20 mM Tris-HCl [pH 7.5], 150 mM KCl, 2 mM EDTA, 0.5% NP40, 1 mM NaF, 1 mM DTT, 100 U/ml RNasin ribonuclease inhibitor [Progmega], Complete EDTA-Free Protease Inhibitor) and incubated on ice for 10 min. All the subsequent steps were performed in the cold room. Testis lysates were prepared by Dounce homogenizers with 20 times initial strokes of pestle A followed by pestle B. Obtained tissue lysates were cleared by ultracentrifugation at 20,000 × *g* for 30 min at 4°C. Cleared testis lysates were then transferred to clean Eppendorf tubes and immediately used for immunoprecipitations or RNA purification.

Antibody Preparation

N-terminal human PIWIL proteins were cloned into glutathione S-transferase (GST)-tagged vectors, and proteins were expressed in *E. coli* bacteria and purified as GST-tagged proteins. Recombinant proteins were used as antigens to generate individual anti-PIWIL-specific rabbit polyclonal antibodies. Specificity and cross-reactivity of antibodies were tested using cell lysates prepared from stable FLAG- and hemagglutinin (HA)-tagged PIWIL protein-expressing HEK293 FlpIn cells and human testis tissue lysate. Furthermore, antigen-affinity-purified rabbit polyclonal antibodies were used to immunoprecipitate endogenous PIWIL protein from human testis lysate.

Immunoprecipitation of PIWIL Ribonucleoprotein Complexes and Isolation and Labeling of Bead-Bound Nucleic Acid

1.2 ml cleared human testis lysate (3 mg/ml total protein) in NT2 buffer (50 mM Tris-HCl [pH 7.4], 150 mM NaCl, 1 mM $MgCl_2$, and 0.05% NP40) supplemented with 1 mM DTT, 2 mM EDTA, and 100 U/ml RNasin was used. PIWIL antibodies were cross-linked to Dynal Protein G beads per the manufacturers' protocol. Antibody cross-linked beads were incubated with lysate for 4 hr at 4°C with gentle agitation. Using a magnetic rack, beads were washed twice with ice-cold NT2 and twice with NT2 supplemented with 1.0% Triton X-100. RNA that co-immunoprecipitated with PIWIL proteins was isolated by treatment of the beads with 0.6 mg/ml proteinase K in 0.3 ml proteinase K buffer (150 mM NaCl, 12.5 mM EDTA, 1% SDS, and 100 mM Tris-HCl [pH 7.5]), followed by phenol (at neutral pH)/chloroform extraction and ethanol precipitation.

ACCESSION NUMBERS

Raw sequencing data are available in NCBI SRA repository under BioProject: PRJNA292450 (http://www.ncbi.nlm.nih.gov/bioproject/PRJNA292450).

SUPPLEMENTAL INFORMATION

Supplemental information includes four figures and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.09.030 and at https://rnaworld.rockefeller.edu/piRNA_supplement.

AUTHOR CONTRIBUTIONS

Z.W., A.M., C.L., P.K.P., and S.J. performed the experiments. P.M. performed bioinformatics analysis. Z.W., P.K.P., Z.R., and T.T. designed the study. Z.W., P.M., and T.T. wrote the manuscript.

ACKNOWLEDGMENTS

We wish to thank members of the Tuschl lab for helpful discussions and advice. We thank Dr. Carolyn Westhoff and the faculty of the Division of Family Planning at Columbia University Medical Center for facilitating sample collection and Markus Landthaler (Berlin Institute for Medical System Biology, Max Delbrueck Center for Molecular Medicine of the Berlin Institute) for generating the PIWIL-protein-expressing HEK293 cell lines. This research was supported by the Howard Hughes Medical Institute (T.T.), National Institutes of Health grant HD068546, and the Center for Reproductive Medicine and Infertility (Z.W.).

Received: December 21, 2014
Revised: July 10, 2015
Accepted: September 10, 2015
Published: October 15, 2015

REFERENCES

Akbergenov, R., Si-Ammour, A., Blevins, T., Amin, I., Kutter, C., Vanderschuren, H., Zhang, P., Gruissem, W., Meins, F., Jr., Hohn, T., and Pooggin, M.M. (2006). Molecular characterization of geminivirus-derived small RNAs in different plant species. Nucleic Acids Res. *34*, 462–471.

Alefelder, S., Patel, B.K., and Eckstein, F. (1998). Incorporation of terminal phosphorothioates into oligonucleotides. Nucleic Acids Res. *26*, 4983–4988.

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.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., and Tuschl, T. (2003). The small RNA profile during Drosophila melanogaster development. Dev. Cell *5*, 337–350.

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., Tóth, 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.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215–233.

Beyret, E., and Lin, H. (2011). Pinpointing the expression of piRNAs and function of the PIWI protein subfamily during spermatogenesis in the mouse. Dev. Biol. *355*, 215–226.

Beyret, E., Liu, N., and Lin, H. (2012). piRNA biogenesis during adult spermatogenesis in mice is independent of the ping-pong mechanism. Cell Res. *22*, 1429–1439.

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.

Brown, M., Suryawanshi, H., Hafner, M., Farazi, T.A., and Tuschl, T. (2013). Mammalian miRNA curation through next-generation sequencing. Front. Genet. *4*, 145.

Carmell, M.A., Girard, A., van de Kant, H.J.G., 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.

Castañeda, J., Genzor, P., van der Heijden, G.W., Sarkeshik, A., Yates, J.R., 3rd, Ingolia, N.T., and Bortvin, A. (2014). Reduced pachytene piRNAs and translation underlie spermiogenic arrest in Maelstrom mutant mice. EMBO J. *33*, 1999–2019.

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

De Fazio, S., Bartonicek, N., Di Giacomo, M., Abreu-Goodger, C., Sankar, A., Funaya, C., Antony, C., Moreira, P.N., Enright, A.J., and O'Carroll, D. (2011). The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. Nature *480*, 259–263.

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

Di Giacomo, M., Comazzetto, S., Saini, H., De Fazio, S., Carrieri, C., Morgan, M., Vasiliauskaite, L., Benes, V., Enright, A.J., and O'Carroll, D. (2013). Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. Mol. Cell *50*, 601–608.

Di Giacomo, M., Comazzetto, S., Sampath, S.C., Sampath, S.S., and O'Carroll, D. (2015). G9a co-suppresses LINE1 elements in spermatogonia. Epigenetics Chromatin *7*, 24.

Ebhardt, H.A., Thi, E.P., Wang, M.-B., and Unrau, P.J. (2005). Extensive 3′ modification of plant small RNAs is modulated by helper component-proteinase expression. Proc. Natl. Acad. Sci. USA *102*, 13398–13403.

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.

Farazi, T.A., Horlings, H.M., Ten Hoeve, J.J., Mihailovic, A., Halfwerk, H., Morozov, P., Brown, M., Hafner, M., Reyal, F., van Kouwenhove, M., et al. (2011). MicroRNA sequence and expression analysis in breast tumors by deep sequencing. Cancer Res. *71*, 4443–4453.

Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. Nat. Rev. Genet. *15*, 829–845.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nat. Rev. Genet. *10*, 94–108.

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.

Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in Drosophila. Science *315*, 1587–1590.

Ha, H., Song, J., Wang, S., Kapusta, A., Feschotte, C., Chen, K.C., and Xing, J. (2014). A comprehensive analysis of piRNAs from adult human testis and their relationship with genes and mobile elements. BMC Genomics *15*, 545–561.

Hafner, M., Renwick, N., Farazi, T.A., Mihailović, A., Pena, J.T.G., and Tuschl, T. (2012). Barcoded cDNA library preparation for small RNA profiling by next-generation sequencing. Methods *58*, 164–170.

Hartshorne, G.M., Lyrakou, S., Hamoda, H., Oloto, E., and Ghafari, F. (2009). Oogenesis and cell death in human prenatal ovaries: what are the criteria for oocyte selection? Mol. Hum. Reprod. *15*, 805–819.

Hirano, T., Iwasaki, Y.W., Lin, Z.Y.-C., Imamura, M., Seki, N.M., Sasaki, E., Saito, K., Okano, H., Siomi, M.C., and Siomi, H. (2014). Small RNA profiling and characterization of piRNA clusters in the adult testes of the common marmoset, a model primate. RNA *20*, 1223–1237.

Horwich, M.D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., and Zamore, P.D. (2007). The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr. Biol. *17*, 1265–1272.

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.

Houwing, S., Berezikov, E., and Ketting, R.F. (2008). Zili is required for germ cell differentiation and meiosis in zebrafish. EMBO J. *27*, 2702–2711.

Ishizu, H., Siomi, H., and Siomi, M.C. (2012). Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. Genes Dev. *26*, 2361–2373.

Juliano, C., Wang, J., and Lin, H. (2011). Uniting germline and stem cells: the function of Piwi proteins and the piRNA pathway in diverse organisms. Annu. Rev. Genet. *45*, 447–469.

Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., Okada, T.N., Siomi, M.C., and Siomi, H. (2008). Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature *453*, 793–797.

Kirino, Y., and Mourelatos, Z. (2007a). Mouse Piwi-interacting RNAs are 2′-O-methylated at their 3′ termini. Nat. Struct. Mol. Biol. *14*, 347–348.

Kirino, Y., and Mourelatos, Z. (2007b). 2′-O-methyl modification in mouse piRNAs and its methylase. Nucleic Acids Symp. Ser. (Oxf). *51*, 417–418.

Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. *42*, D68–D73.

Kuramochi-Miyagawa, S., Kimura, T., Yomogida, K., Kuroiwa, A., Tadokoro, Y., Fujita, Y., Sato, M., Matsuda, Y., and Nakano, T. (2001). Two mouse piwi-related genes: miwi and mili. Mech. Dev. *108*, 121–133.

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.

Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A.O., Landthaler, M., et al. (2007). A mammalian microRNA expression atlas based on small RNA library sequencing. Cell *129*, 1401–1414.

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.

Lau, N.C., Ohsumi, T., Borowsky, M., Kingston, R.E., and Blower, M.D. (2009). Systematic and single cell analysis of Xenopus Piwi-interacting RNAs and Xiwi. EMBO J. *28*, 2945–2958.

Le Thomas, A., Tóth, K.F., and Aravin, A.A. (2014). To be or not to be a piRNA: genomic origin and processing of piRNAs. Genome Biol. *15*, 204.

Li, X.Z., Roy, C.K., Dong, X., Bolcun-Filas, E., Wang, J., Han, B.W., Xu, J., Moore, M.J., Schimenti, J.C., Weng, Z., and Zamore, P.D. (2013). An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. Mol. Cell *50*, 67–81.

Malki, S., van der Heijden, G.W., O'Donnell, K.A., Martin, S.L., and Bortvin, A. (2014). A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. Dev. Cell *29*, 521–533.

Malone, C.D., and Hannon, G.J. (2009). Small RNAs as guardians of the genome. Cell *136*, 656–668.

Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. Nat. Rev. Genet. *14*, 447–459.

Pillai, R.S., and Chuma, S. (2012). piRNAs and their involvement in male germline development in mice. Dev. Growth Differ. *54*, 78–92.

Rajasethupathy, P., Antonov, I., Sheridan, R., Frey, S., Sander, C., Tuschl, T., and Kandel, E.R. (2012). A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. Cell *149*, 693–707.

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.

Roovers, E.F., Rosenkranz, D., Mahdipour, M., Han, C.T., He, N., Chuva de Sousa Lopes, S.M., van der Westerlaken, L.A., Zischler, H., Butter, F., Roelen, B.A., and Ketting, R.F. (2015). Piwi proteins and piRNAs in mammalian oocytes and early embryos. Cell Rep. *10*, 2069–2082.

Ross, R.J., Weiner, M.M., and Lin, H. (2014). PIWI proteins and PIWI-interacting RNAs in the soma. Nature *505*, 353–359.

Rosenkranz, D., and Zischler, H. (2012). proTRAC–a software for probabilistic piRNA cluster detection, visualization and analysis. BMC Bioinformatics *13*, 5.

Saito, K., Nishida, K.M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., and Siomi, M.C. (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev. *20*, 2214–2222.

Sasaki, T., Shiohama, A., Minoshima, S., and Shimizu, N. (2003). Identification of eight members of the Argonaute family in the human genome. Genomics *82*, 323–330.

Shoji, M., Tanaka, T., Hosokawa, M., Reuter, M., Stark, A., Kato, Y., Kondoh, G., Okawa, K., Chujo, T., Suzuki, T., et al. (2009). The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. Dev. Cell *17*, 775–787.

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.

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.

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.

Weick, E.-M., and Miska, E.A. (2014). piRNAs: from biogenesis to function. Development *141*, 3458–3471.

Wilczynska, A., Minshall, N., Armisen, J., Miska, E.A., and Standart, N. (2009). Two Piwi proteins, Xiwi and Xili, are expressed in the Xenopus female germline. RNA *15*, 337–345.

Yang, Q., Hua, J., Wang, L., Xu, B., Zhang, H., Ye, N., Zhang, Z., Yu, D., Cooke, H.J., Zhang, Y., and Shi, Q. (2013). MicroRNA and piRNA profiles in normal human testis detected by next generation sequencing. PLoS ONE *8*, e66809.

Zheng, K., and Wang, P.J. (2012). Blockade of pachytene piRNA biogenesis reveals a novel requirement for maintaining post-meiotic germline genome integrity. PLoS Genet. *8*, e1003038.