Supplementary Tables and Figure Legends

Supplementary Figure S1. Shown are 134 gene models for *Drosophila* transcripts with abundant (>1000) genic piRNAs (see also Supplementary Table S2). The number of unique reads per 10nt window were plotted. Most of these transcripts exhibit 3' UTR-

directed piRNA production from the sense strand (red piRNA bars), with little or not piRNAs mapping to the antisense strand (blue). Gene models at the bottom of each page depict 5' UTR in blue, coding exons in black, and 3' UTR in red.

Supplementary Figure S2. (A) Overlap of piRNA-generating mRNAs in OSS cells, ovaries and embryos. The 50 piRNA cutoff for mRNAs in the OSS data represented 3.6 piRNAs per million reads. We identified mRNAs in ovary and early embryo small RNA data [14] that passed this threshold. Most of the piRNA-generating mRNAs in the animal overlap with those captured in OSS cells. (B) The primary piRNA pathway is preferentially active on highly expressed transcripts. Genes whose expression was called as present in OSS cells were divided into three categories: those with >1000 3' UTR piRNAs, those with 50-999 3' UTR piRNAs, and those with 0-49 3' UTR piRNAs. These were plotted against their $\log_2$ expression level (rounded to nearest integer) using the box-whisker convention, where the bounds of the box represent the 1st and 3rd quartiles, the line represents the median, and the dashed lines represent 1.5x the interquartile range. The outliers are represented as dots, with the number of transcripts in each range designated. The most abundant production of piRNAs comes from relatively highly-expressed genes (p value assessed by Kolmogorov-Smirnov test); still, a population of highly expressed transcripts do not generate piRNAs.

Supplementary Figure S3. Shown are 292 gene models for murine transcripts with abundant (>1000) genic piRNAs (see also Supplementary Table S5A). The number of unique reads per 10nt window were plotted. Most of these transcripts exhibit 3' UTR-directed piRNA production from the sense strand (red piRNA bars), with little or not piRNAs mapping to the antisense strand (blue). Gene models at the bottom of each page depict 5' UTR in blue, coding exons in black, and 3' UTR in red. To focus on

piRNAs that were unequivocally generated from these genes, we plotted only uniquely-mapping reads.

Supplementary Figure S5. *FTH1* is an exceptional transcript that generates piRNAs from across the transcript, excepting the 3' UTR. Gene model depicts 5' UTR in blue, coding exons in black, and 3' UTR in red. Following FTH1 is a list of gene maps for additional genes with abundant piRNAs arising from the CDS.

Supplementary Table S1: OSS cis-NAT-siRNA loci. The first worksheet (S1a) compiles information on 3' cis-NATs, including the genes in each cis-NAT pair, the location and length of the overlap region and the overall gene regions, the numbers of small RNAs recorded in overlap and non-overlap regions, the number of 21nt reads recorded in overlap and non-overlap regions, and enrichments for small RNAs and 21nt RNAs in the overlap regions. The second worksheet (S1b) compares the top 40 cis-NAT-siRNA-generating loci in S2 cells and OSS cells; the 17 loci in common are highlighted in green.

Supplementary Table S2: Coding transcripts that generate piRNAs in OSS cells. Three worksheets in this table compile information on piRNA-generating mRNAs in OSS cells, including the gene, its genomic location, and the length and number of sense-strand piRNAs in its 5' UTR, CDS and 3' UTR. For each transcript, the 5' UTR, CDS and 3' UTR were divided into 10 equal ranges and the number of piRNAs in each window were tabulated. These were used to generate the schematics in Figure 2A and 2B. The first worksheet (S2a) summarizes all 2356 mRNAs with >50 piRNAs, the second worksheet (S2b) contains the 125 transcripts with >1000 piRNAs defined as " 3' UTR"-enriched, and the third worksheet (S2c) compiles 55 additional transcripts with >1000 piRNAs across the body of the message.

Supplementary Table S3. Reexamination of published lists of piRNA clusters from of the top 200 fetal, 100 pre-pachytene, and 94 post-pachytene clusters [4, 10, 36]. The clusters were manually examined for piRNA distributions, noting the proximity and concordance of piRNA distributions with annotated Refseq gene models.

Supplementary Table S4. Summary of genic piRNAs determined in various stages of mouse spermatogenesis. (A) Proportion of 5’ UTR, CDS, and 3’ UTR piRNAs in mouse testis libraries, with representation as pie charts to illustrate the proportions. (B) Summary statistics of small RNA library sequencing.

Supplementary Table S5: Transcripts that generate piRNAs in murine testis. (A) Small RNA reads were compiled from data published by Hannon and colleagues[30]; shown are transcripts with >1000 piRNAs. (B) Small RNA reads compiled by this study of mouse adult testes samples. The table compiles information on the genomic location of each locus, and the length and number of sense-strand piRNAs in its 5' UTR, CDS and 3' UTR.

Supplementary Table S6: (A) GO-term enrichment analysis of the 125 3' UTR-enriched transcripts with >1000 piRNAs in OSS cells. The first worksheet includes "biochemical process" (BP) terms, the second worksheet includes "molecular function" (MF) terms, and the third worksheet includes "cell compartment" (CC) terms. (B) GO-term enrichment analysis of transcripts with <10 piRNAs in OSS cells. The first worksheet includes "biochemical process" (BP) terms, the second worksheet includes "molecular function" (MF) terms, and the third worksheet includes "cell compartment" (CC) terms.

Cells colored in orange are GO term categories shared between the datasets, while cells colored in green are exclusive GO term categories.

Supplementary Table S7: GO-term enrichment analysis of 3' UTR-enriched transcripts with >10 piRNAs in mouse adult testes or >50 piRNAs in 10dpp testes versus highly expressed gene that lack piRNAs. The first worksheet includes analysis from adult mouse testes total RNA piRNAs, while the second worksheet includes analysis from 10dpp testes. The first group of rows list the "biochemical process" (BP) terms, the "molecular function" (MF) terms and the "cell compartment" (CC) terms for genes containing piRNAs, while the following group of rows concern the dataset of genes lacking piRNAs. Rows colored in orange are GO term categories shared between the datasets, while rows colored in green are exclusive GO term categories.

Supplementary Table S8: Gene expression of abundant (>1000) piRNA-generating transcripts in *Mili* +/- and -/- testis. No overall coherent changes in transcript abundance were observed overall.