

**Cell, Volume 137**

## Collapse of Germline piRNAs

in the Absence of Argonaute3

Reveals Somatic piRNAs in Flies

Chengjian Li, Vasily V. Vagin, Soohyun Lee, Jia Xu, Shengmei Ma, Hualin Xi, Hervé Seitz, Michael D. Horwich, Monika Syrzycka, Barry M. Honda, Ellen L.W. Kittler, Maria L. Zapp, Carla Klattenhoff, Nadine Schulz, William E. Theurkauf, Zhiping Weng, and Phillip D. Zamore

### **SUPPLEMENTAL DATA**

**Supplemental Discussion**

**Supplemental Experimental and Computational Procedures.**

**Supplemental References**

**Six Supplemental Tables.**

**Table S1** Female fertility.

**Table S2.** Sequencing Statistics.

**Table S3.** The fraction of sense and antisense piRNAs that were uniquely bound to each PIWI protein.

**Table S4.** Probes for Northern hybridization

**Table S5.** Primers for quantitative PCR.

**Table S6.** Fly stocks used in this study.

**Thirteen Supplemental Figures.**

**Figure S1.** *ago3* alleles.

**Figure S2.** Fertility analysis.

**Figure S3.** Resolution of piRNAs into three distinct classes of transposon families by comparing their sense fraction in *ago3*/TM6B and *ago3<sup>t2</sup>*/*ago3<sup>t3</sup>* ovaries with

the change in the abundance of antisense piRNAs between the two genotypes.

**Figure S4.** Detailed analysis of piRNAs in Oregon R versus *ago3*/TM6B ovaries.

**Figure S5.** Detailed analysis of piRNAs in *ago3*/TM6B versus *ago3<sup>t2</sup>*/*ago3<sup>t3</sup>* ovaries.

**Figure S6.** Detailed analysis of piRNAs bound to Ago3, Aub, and Piwi in *ago3*/TM6B and *ago3<sup>t2</sup>*/*ago3<sup>t3</sup>* ovaries.

**Figure S7.** Sequence bias of Aub-bound piRNAs in *ago3*/TM6B and *ago3<sup>t2</sup>*/*ago3<sup>t3</sup>* ovaries.

**Figure S8.** Analysis by class of piRNAs uniquely bound to Aub or Piwi in *ago3*/TM6B and *ago3<sup>t2</sup>*/*ago3<sup>t3</sup>* for 95 transposon families.

**Figure S9.** Overview of piRNA abundance in *aub*/CyO and *aub<sup>HN2</sup>*/*aub<sup>QC42</sup>* for 95 transposon families

**Figure S10.** Detailed analysis of piRNAs in *aub*/CyO versus *aub<sup>HN2</sup>*/*aub<sup>QC42</sup>* ovaries.

**Figure S11.** Tiling microarray analysis of transposon and mRNA expression in wild-type (*w<sup>1118</sup>*) and *ago3* mutant ovaries.

**Figure S12.** Tiling microarray analysis of transposon and mRNA expression in wild-type (*w<sup>1118</sup>*) and *aub* mutant ovaries.

**Figure S13.** Quantitative RT-PCR analysis of transposon expression in *aub<sup>HN2</sup>*/*aub<sup>QC42</sup>* mutant ovaries.

## Supplemental Discussion

In addition to Ago3:Aub and Aub:Ago3 ping-pong pairs, we also detected Aub:Aub ping-pong pairs in *ago3*/TM6B ovaries. These Aub:Aub ping-pong pairs have a strand bias similar to that observed for Ago3:Aub and Aub:Ago3 pairs. That is, group I, Aub-bound antisense piRNAs were 2.1 times more abundant than their Ago3-bound sense piRNA partners, and group II, Aub-bound sense piRNAs were 2.1 times more abundant than their Ago3-bound antisense piRNA partners. Similarly, Aub:Aub antisense piRNAs were twice as abundant than sense for group I and 2.3 times less abundant for group II.

These observations are hard to reconcile with the idea that a difference in the concentrations of Ago3 and Aub biases Aub-bound piRNAs towards antisense, as the protein partners in an Aub:Aub ping-pong cycle are, of course, at equal concentration. Moreover, piRNAs participating in Aub:Aub ping pong are antisense-biased in *ago3*/TM6B ovaries, but become sense biased or unbiased in *ago3<sup>12</sup>/ago3<sup>13</sup>* ovaries: group I piRNAs participating in an Aub:Aub ping-pong cycle were sense biased (antisense/sense = 0.78), and the group II piRNAs had essentially no bias (antisense/sense = 1.06).

The simplest resolution for this paradox is that our nearly 5 million piRNA reads (derived from more than 16 million genome-matching reads) from immunoprecipitation experiments significantly underestimate the number of Ago3:Aub ping-pong pairs, likely because Ago3-bound piRNAs are so much rarer than those bound to Aub. Perhaps in vivo many Aub-bound piRNAs have a corresponding Ago3-bound partner that we failed to sequence. Clearly, much deeper sequencing of piRNAs will be required to test this idea.

## Supplemental Experimental Procedures

### General Methods

Preparation of 2–4 day ovary lysate (Tuschl et al., 1999; Tomari et al., 2004) and Northern hybridization and quantitative RT-PCR analysis were as described (Vagin et al., 2006; Ghildiyal et al., 2008). Tables S4 and S5 report probe and primer sequences. Table S6 describes fly stocks. Figures were prepared using Excel (Microsoft, Redmond, WA, USA), IgorPro (WaveMetrics, Lake Oswego, OR, USA), and Illustrator (Adobe Systems, San Jose CA, USA). Microarray data are available via the NCBI gene expression omnibus web site (<http://www.ncbi.nlm.nih.gov/geo/>) using accession number GSE14370.

### Isolation of *ago3*<sup>t1</sup>, *ago3*<sup>t2</sup>, and *ago3*<sup>t3</sup> alleles

6,000 EMS-mutagenized fly lines (Koundakjian et al., 2004) were screened by tilling (Colbert et al., 2001; Till et al., 2003; Winkler et al., 2005) by the Seattle TILLING project (<http://tilling.fhcrc.org/>) according to Cooper et al. (2008). Candidate lines that contained mutations that induced premature stop codons in the *ago3* coding sequence were further characterized by sequencing of genomic PCR amplicons. A 1,532 bp region of *ago3* was analyzed using the primers 5'-AAC GAC GGA TGA ATC CAA GGG AGT TTT-3' and 5'-GAA ATA CCA TTG GTT TGC CGA ATT TGA-3'. PCR resequencing of candidate mutations employed the primers 5'-ATG AAT CCA AGG GAG T-3' and 5'-AGA GCA TTA CCA AGA ATC-3' or 5'-GGA CGT TAA TCA ACC A-3' and 5'-TGA CGA ATA CAA AGG T-3'.

### Isolation and characterization of *Df(3L)TTT*

In a screen for deficiency (*Df*) mutations in 3L heterochromatin,

KG03264/KG03264 males (carrying a P(SUPor-P) insert near the *nrm* locus) were treated with 4,000 rads of X-ray radiation, and mated en masse to *w/w*; TM3 *Sb*/TM6 *Hu Tb* virgin females. Single F1 K3264\*/TM3 *Sb* males (where \* indicates a mutagenized third chromosome) were then crossed to *Df(3L)FX3 e*/TM3 *Sb* females (Schulze et al., 2005), and stocks of putative Df mutations established from K3264\*/TM3 *Sb* siblings. To map the newly isolated Dfs relative to individual lethal complementation groups, these putative Dfs were tested for lethality against smaller heterochromatic deficiencies on 3L, as well as against lethal mutations in individual essential genes. The molecular extent of *Df(3L)TTT* was further characterized by isolating genomic DNA from homozygous *Df(3L)TTT/Df(3L)TTT* embryos, and testing it by PCR: primers were used to look for the presence of exon regions of several 3L heterochromatic gene sequences (including *ago3*) which are distal to previously defined genetic loci. This analysis revealed that *Df(3L)TTT* is a large deficiency deleting a number of genes, from *ago3* through to *vtd* as the most proximal gene removed (data not shown).

### **Generation of transgenic flies**

Ago3 was cloned from Oregon R ovary cDNA using 3' RACE. Ovary RNA was reverse transcribed using Superscript RT II (Invitrogen, Carlsbad, CA, USA) and a reverse transcription primer provided in the Gene Racer Kit (Invitrogen) (5'-GCT GTC AAC GAT ACG CTA CGT AAC GGC ATG ACA GTG (T)<sub>18</sub>-3'). First strand cDNA was treated with RNase H. Full length Ago3 coding sequence was amplified by PCR using KOD polymerase (Takara Bio, Otsu, Japan) and the primers 5'-CAC CAT GTC TGG AAG AGG AAA TTT GTT GAG C-3', which spans the predicted translational initiation site (Williams and Rubin, 2002) and 5'-GCT GTC AAC GAT ACG CTA CGT AAC G-3' (from the 3' RACE kit). The ~2,600 nt product was cloned into pEntr/D-Topo (Invitrogen). Sequencing of the

full length insert revealed an 867 amino acid open reading frame and a 70 nt 3' UTR. The open reading frame contained a single base change (A2143G) yielding I715V. pEntr/D-Topo-Ago3 was recombined into the pPFMW vector (UASp-Flag<sub>3</sub>-Myc<sub>6</sub>-ORF) from the Carnegie *Drosophila* Gateway Collection using Clonase (Invitrogen). Insert junctions were confirmed by sequencing. Genetic Services Inc. (Sudbury, MA, USA) injected plasmid DNA and identified transgenic flies. Transgenic insertions were mapped by inverse PCR (Huang et al., 2000).

### **Male fertility testing**

To test male fertility, one virgin male was mated to five virgin Oregon R females; five individual males were tested for each genotype. After three days, the male was removed and the five females were transferred to a fresh vial, and then transferred to new vials every other day until 4 vials were obtained. The females were removed from the last vial after two additional days. The number of the adult progeny from each vial was counted.

### **Female fertility testing**

Ten female virgins were collected and mated to five Oregon R virgin males on the day of collection in a small cage with a 60 mm diameter grape juice agar plate dabbed with yeast paste plate at 25°C. For wild-type controls, six Oregon R female virgins were mated to three Oregon R virgin males. After two days, the first plate was discarded and replaced with a fresh plate, and then the plate was changed and scored every subsequent day. The numbers of total eggs, eggs per female per day, and the dorsal appendage phenotype of embryos were scored every 24 h, and the numbers of eggs hatching were scored 48 h after the plate was changed. Twelve plates were scored in total for each genotype.

## **Immunohistochemistry and microscopy**

Egg chamber fixation and whole-mount antibody labeling were performed as previously described (Theurkauf, 1994). Vasa protein was detected with rabbit polyclonal anti-Vasa antibody (Liang et al., 1994) diluted 1:1000. Piwi, Aub and Ago3 were detected with rabbit polyclonal anti-Piwi, anti-Aub and anti-Ago3 antibodies (Brennecke et al., 2007) and diluted 1:1000. Stellate antiserum (Klattenhoff et al., 2007) diluted 1:1000. Rhodamine-conjugated phalloidin (Invitrogen) was diluted 1:100 to stain F-actin; TOTO3 (Invitrogen) was used at 0.2  $\mu$ M to visualize DNA. All tissues were mounted in 90% glycerol/PBS, with 1 mg/ml *p*-Phenylenediamine (Sigma-Aldrich, St. Louis, MO, USA). Samples were analyzed using a Leica TCS-SP inverted laser-scanning microscope with 63x (NA 1.32) PlanApo and 40x (NA 1.25) PlanApo oil-immersion objectives. Identical imaging conditions were used for each set of wild-type and mutant samples. Images were processed using Image J software.

## **Western blotting**

Fifty  $\mu$ g total protein was mixed with an equal volume of 100 mM Tris-Cl, pH 6.8, 4% [w/v] SDS, 200 mM DTT, 20% [v/v] glycerol, and 0.2% (w/v) bromophenol blue, boiled for 10 min, and then resolved by 8% polyacrylamide/SDS gel electrophoresis. After electrophoresis, proteins were transferred to a PVDF membrane (Immobilon-P, Millipore, Billerica, MA, USA). The membrane was blocked in TBST-milk (25 mM Tris-Cl, pH 7.4, 3.0 mM KCl, 140 mM NaCl, 0.05% [v/v] Tween-20, 5% [w/v] non-fat dry milk) at room temperature for 2 h. After blocking, the membrane was cut according the molecular weight markers and incubated overnight at 4°C in TBST-milk containing primary antibody (anti-Ago3 (Brennecke et al., 2007) diluted 1:500, anti-Aub (Brennecke et al., 2007) diluted

1:1000, anti-Piwi (Saito et al., 2006) diluted 1:50, anti-Vasa (Liang et al., 1994) diluted 1:1000, anti-Ago1 (Miyoshi et al., 2005) diluted 1:1000, anti-Ago2 (Miyoshi et al., 2005) without dilution, anti-Armi at 1:1000, and anti- $\alpha$ -Tubulin (Sigma-Aldrich) diluted 1:10,000. Next, the membrane was washed three times with TBST at room temperature for 30 min and incubated 2 h at room temperature with either sheep anti-mouse IgG-HRP (GE Healthcare, Piscataway, NJ, USA) at 1:10,000 or goat anti-rabbit IgG-HRP (GE Healthcare) at 1:10,000 in TBST-milk. Then the membrane was washed three times with TBST at room temperature for 30 min and developed with SuperSignal West Dura Extended Duration Substrate (Pierce, Rockford, IL, USA). Image data were captured with an LAS-3000 image reader (Fujifilm, Tokyo, Japan). Quantitative analysis was performed using ImageGauge V4.22 (Fujifilm, Tokyo, Japan).

### **Immunoprecipitation**

Twenty  $\mu$ l GammaBind G Sepharose (GE Healthcare) were incubated with 10  $\mu$ l anti-Ago3, anti-Aub or anti-Piwi antibody for 2 h at 4°C. Next, the beads were washed five times with lysis buffer (30 mM HEPES-KOH, pH 7.4, 100 mM  $\text{KCH}_3\text{CO}_2$ , 2 mM  $\text{Mg}(\text{CH}_3\text{CO}_2)_2$ ) containing 5 mM DTT, 0.5% [v/v] NP-40 and 1 tablet/50 ml complete-EDTA-free protease inhibitor cocktail tablet (Roche). Subsequently, 200  $\mu$ l of ovary lysate (10  $\mu$ g/ $\mu$ l total protein) was added and the mixture agitated gently at 4°C overnight. The supernatant (200  $\mu$ l) and antibody-bound material were separated by centrifugation at 3,000 rpm for 1 min at 4°C, and the antibody-bound beads washed five times with RIPA buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 1.0% [v/v] NP-40, 0.5% [w/v] DOC, 0.1% [w/v] SDS) containing 1 tablet/50 ml complete-EDTA-free protease inhibitor cocktail tablet, then resuspended in 200  $\mu$ l lysis buffer. Five  $\mu$ l of input, supernatant, or bound samples were subject to western blotting analysis to confirm



immunoprecipitation. The remaining samples were treated with 1 mg/ml proteinase K in 100 mM Tris-Cl, pH 7.5, 12.5 mM EDTA pH 8.0, 150 mM NaCl, 1% [w/v] SDS for 1 h at 65°C, extracted with an equal volume of phenol/chloroform, and then precipitated with three volumes 100% ethanol in the presence of 20 µg glycogen carrier. The precipitate was washed once with 80% ethanol and then dissolved in nuclease-free water.

### **Small RNA cloning and sequencing**

Total RNA was isolated from manually dissected ovaries from 2–4 day flies using mirVana (Ambion, Austin, TX, USA). The RNA was quantified by absorbance at 260 nm. To deplete 2S rRNA, 100 µg total RNA was annealed with 200 pmole of a complementary DNA oligonucleotide (5'-AGT CTT ACA ACC CTC AAC CAT ATG TAG TCC AAG CAG CAC T-3') in 20 µl at 95°C for 2 min, then the temperature was gradually decreased over 1 h to room temperature. 2S rRNA was digested with 2 units of RNase H (Invitrogen) in a 30 µl reaction containing 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM MgCl<sub>2</sub>, and 10 mM DTT. After 2S rRNA depletion, 18–29 nt small RNA was purified from a 15% denaturing urea-polyacrylamide gel (National Diagnostics, Atlanta, GA, USA). Half of the purified RNA was oxidized with 25 mM NaIO<sub>4</sub> in 60 mM borax, 60 mM boric acid, pH 8.6, for 30 min at room temperature followed by ethanol precipitation. 100 pmole of 3' pre-adenylated adapter (5'-rAppTCG TAT GCC GTC TTC TGC TTG T/ddC/-3') was ligated to oxidized or un-oxidized small RNAs using mutant Rnl2 (amino acids 1-249, K227Q) (17-23B) (Addgene, Cambridge, MA, USA) at 4°C for 12 h in 20 µl in 50 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 60 µg/ml BSA, 10% (v/v) DMSO, and 40 U RNasin (Promega, Madison, WI, USA). 3' ligated product was purified from a 15% denaturing urea-polyacrylamide gel (National Diagnostics), and then ligated to 100 pmole of

5' RNA adapter (5'-GUU CAG AGU UCU ACA GUC CGA CGA UC-3') using T4 RNA ligase (Ambion, Austin, TX, USA) at room temperature for 6 h in 20  $\mu$ l of 50 mM Tris-HCl, pH 7.8, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 1 mM ATP, and 10% DMSO. The ligated product was purified from a 10% denaturing urea-polyacrylamide gel (National Diagnostics). Half the ligated product was used to synthesize cDNA using Superscript III (Invitrogen) and the reverse transcription primer (5'-CAA GCA GAA GAC GGC ATA CGA-3'), and half was used as a minus RT control. The small RNA library was amplified using forward (5'-AAT GAT ACG GCG ACC ACC GAC AGG TTC AGA GTT CTA CAG TCC GA-3') and reverse (5'-CAA GCA GAA GAC GGC ATA CGA-3') primers, and then purified from a NuSieve GTG agarose gel (Lonza, Basel, Switzerland). Purified libraries were sequenced using a Solexa Genome Analyzer (Illumina, San Diego, CA, USA). Sequence and abundance data are available via the NCBI gene expression omnibus web site using accession number SRA007727.

## **Bioinformatics**

### **Sequence extraction and annotation**

For each sequence read, the first occurrence of the 6-mer perfectly matching the 5' end of the 3' linker was identified. Sequences without a match were discarded. The extracted inserts for sequences that contained the 3' linker were then mapped to the female *Drosophila melanogaster* genome (Release R5.5, excluding chromosome YHet). Inserts that matched fully to a genomic sequence were collected using either Bowtie (Langmead et al., 2009) or in-house suffix tree-based software (Gusfield, 1997; Delcher et al., 1999; Delcher et al., 2002), and the corresponding genomic coordinates were determined for downstream functional analysis. Sequences corresponding to pre-miRNAs or non-coding

RNAs (ncRNAs) were identified using the same suffix tree-based software. Genes and transposon annotations were retrieved from FlyBase (R5.5). In addition, unannotated transposons were identified using BLAST (Altschul et al., 1990) to query each transposon consensus sequence against the female genome, using an *e*-value cutoff of  $10^{-10}$ . The number of reads for each small RNA was normalized by the number of times that RNA mapped to the genome and by the abundance of ncRNAs in the sample. For all transposon analyses, if a piRNA mapped to a location where there were two or more blast-based transposon annotations, the number of reads were normalized by the number of annotations.

### **Background-corrected sequence bias detection**

For a set of sequences of various lengths (e.g. all the piRNAs bound to Piwi protein), the frequency of each nucleotide at each position was computed as a foreground count matrix. The background frequencies at each position were computed by averaging over all possible *k*-mers ( $k = 23-29$ ) in the consensus sequence of each transposon, strand-specifically (either sense or antisense), weighted by the length distribution of each foreground set. Binomial testing was performed for each nucleotide at each position in the foreground matrix, against the corresponding position in the background matrix. The significance level of testing used was  $\alpha = 2.2 \times 10^{-6}$ . This stringency ( $p = 0.001$ ) corrects for multiple testing (four nucleotides at 29 positions, in two orientations (sense and antisense) for two classes (total versus ping-pong pairs)). Only the nucleotides significantly above background were displayed, e.g. in Figure 5B, with *y*-axis values corresponding to the relative frequency of foreground minus background. This background correction helps distinguish an A or U bias that results from the inherent AU-richness of transposon sequences from an A or U bias reflecting the mechanism of piRNA biogenesis. At each significant position, the maximum

possible y-value is indicated as a grey bar. Sequences were analyzed as species and not weighted by their abundance (number of reads).

### **Consensus mapping of total piRNA**

The 23–29 nt small RNAs from the total RNA sequence reads in Oregon R, *ago3*/TM6B, *ago3*<sup>12</sup>/*ago3*<sup>13</sup> *aub*/CyO, and *aub*<sup>HN2</sup>/*aub*<sup>QC42</sup> were mapped to the consensus sequence of each transposon to generate Figures 5A, S4C and S4D, S5C and S5D, and S10C and S10D. To account for the sequence variation among individual copies of the same transposon in the genome, we first establish coordinate conversion for individual copies and the consensus. For each BLAST-hit (the sequence of a copy), all the ungapped segments of its alignment with the consensus sequence were collected and used for coordinate conversion. The genomic coordinate of any piRNA that mapped to an ungapped segment of a BLAST hit was converted to the corresponding position on the consensus sequence. Overlaps of a piRNA to multiple segments were corrected for multiple counting, as well as multiple hits on the same genomic position caused by repetitive sequences. An advantage of coordinate conversion over direct sequence matching to the consensus sequence is that it allows consistency in the small RNAs included in different analyses, i.e., the same set of the small RNAs mapped to transposon regions in the genome are included in the set mapped on the consensus sequence. For each piRNA, all the bases that cover that piRNA were counted as having the number of reads of that piRNA, normalized by the number of times they mapped to the genome and by the total abundance (number of reads) of ncRNAs in the sample.

### **Overlap analysis of total RNAs**

For each genotype, all the piRNAs whose 5' ends overlapped with another

piRNA on the opposite strand were collected in order to generate Figures 5A, S4G, S4H, S5G, S5H, S10G, and S10H.

### **Filtering and Normalization of immunoprecipitation data**

To avoid potential mis-assignment of piRNAs to a specific PIWI protein, we analyzed only those piRNAs that associated uniquely with Ago3, Aub, or Piwi. As 58% (corresponding to 58,265 reads) of Ago3-, 63% (301,656 reads) of Aub-, or 81% (544,102 reads) of Piwi-bound piRNAs were associated with just one of the three proteins, this restriction still permitted analysis of the majority of piRNA species.

The normalizing factor was set to the ratio of the sum of reads in each immunoprecipitated data set to that in the total RNA sample, for the small RNAs that mapped to the fly genome only once and were also sequenced at least once in both the immunoprecipitation sample and the genotype-matched total RNA sample. All piRNA reads in each immunoprecipitation data set was divided by this normalizing factor. This normalization strategy allows direct comparison of the relative abundance of piRNAs uniquely bound to one PIWI protein with those bound to another, as well as comparison between *ago3* heterozygotes and mutants.

### **Consensus mapping of immunoprecipitation data**

piRNAs from *ago3*/TM6B or *ago3*<sup>t2</sup>/*ago3*<sup>t3</sup> ovaries that were uniquely bound to Ago3, Aub, or Piwi were each mapped to the consensus sequence of each transposon, as described for “Consensus mapping of total piRNAs,” above. The abundance was normalized as described above.

### **Overlap graph by ping-pong types (Figures 4E, S6C, and S6D)**

For *ago3*/TM6B, there are nine possible ping-pong protein pairs: Ago3 (sense):Ago3 (antisense), Ago3 (sense):Aub (antisense), Ago3 (sense):Piwi (antisense), Aub (sense):Ago3 (antisense), Aub (sense):Aub (antisense), Aub (sense):Piwi (antisense), Piwi (sense):Ago3 (antisense), Piwi (sense):Aub (antisense), and Piwi (sense):Piwi (antisense). For *ago3<sup>12</sup>*/*ago3<sup>13</sup>*, there are four possible ping-pong protein pairs: Aub (sense):Aub (antisense), Aub (sense):Piwi (antisense), Piwi (sense):Aub (antisense), and Piwi (sense):Piwi (antisense). The sequences that were shared by two or more proteins were excluded. For each transposon, all the piRNA sequences that overlapped with piRNAs from the opposite strand were tallied.

### **Statistical testing of ping-pong pairs (Figures 4E and 6CD)**

To test the role of Ago3-mediated ping-pong amplification of Piwi- and Aub-bound piRNAs, we tallied the number of ping-pong pairs for each of the nine potential protein pairs for transposon-derived piRNAs (Figure 4E). We also tallied the number of pairs that could form in the same data set for overlaps of 8, 9, 11, and 12 nt, taking the average value as a measure of the number of ping-pong pairs that occur simply by chance, explained as follows.

Given two sets of piRNAs (e.g. Ago3-bound sense piRNAs and Aub-bound antisense piRNAs for group II transposons), the observed abundance of ping-pong pairs between the two sets was defined as the sum of all the piRNA reads in the first set and those the second set that have a 10-bp overlap at their 5' ends. Note that all the reads were normalized as described for "Normalization of immunoprecipitation data," above. The same computation was applied to 8, 9, 11, and 12-nt overlaps, and the expected value was computed by averaging over these four numbers. Pooled sample standard deviations from the two sets were used for background standard deviation. Estimating the expectation for the 10-bp

overlap is the problem of randomly throwing  $n$  number of the 5' end of antisense piRNAs and  $m$  number of the 10<sup>th</sup> position of sense piRNAs within some genomic scope (e.g. all transposon copies in the genome) and counting how many fall on the same position. Using overlaps of 8, 9, 11 and 12 nt would result in nearly identical distributions of random noise, given that there is no ping-pong for these overlaps, and thus can be used to estimate the expected value of the 10-nt overlap by chance.

The  $p$ -value of each 10-nt overlap was computed from Z-score = (Observed - Expected) / (standard deviation), assuming that it follows a Gaussian distribution.

### **Expression Analysis (Figures 6A, S11, and S12)**

Coordinates and raw signal values for *Drosophila* tiling 2.0R Array probes were extracted from Affymetrix (Santa Clara, CA, USA) Tiling Array Software. Because transposons account for ~20% of the probes on the tiling array, a special normalization work flow was designed to avoid over-normalization of signals. All probes that mapped to transposons were identified. The remaining probes (which include intergenic and genic regions, with majority of probes corresponding to intergenic regions) were quantile-normalized across wild-type ( $w^{1118}$ ) and mutant replicates based on the assumption that the overall distribution of these probes should remain the same across the different strains. Then the signals for transposon probes were calculated using the normalized values of non-transposon probes at the same raw signal level. To summarize the signal for each transposon element, probes that mapped to any copy of the element were grouped together and the Hodges-Lehmann estimator (Hollander and Wolfe, 1999) was used to calculate the pseudo-median of their normalized signals; pseudo-median is less sensitive to large numbers of outlier probes

(probes whose values were 1, which indicate the values could not be reliably measured) in tiling array experiments. Differentially expressed transposon families were identified by contrasting their pseudo-median values in mutant replicates against the wild-type replicates. To correct for multiple testing, False Discovery Rates (FDRs) were calculated from t-test  $p$ -values. FDR < 0.02 was used to call significantly changed transposons. Transposons with average expression value less < 149 (the 50<sup>th</sup> percentile expression value of all RefSeq mRNAs) in both wild type ( $w^{1118}$ ) and mutants were flagged, as they are at the detection limit of the tiling array. Similarly, expression values of mRNAs were summarized by calculating the pseudo-median of probes mapped to each of the RefSeq mRNA transcripts. Differentially expressed transcripts were identified using FDR < 0.02 cutoff.

### **Supplemental References**

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* 215, 403-410.
- 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.
- Colbert, T., Till, B. J., Tompa, R., Reynolds, S., Steine, M. N., Yeung, A. T., McCallum, C. M., Comai, L., and Henikoff, S. (2001). High-throughput screening for induced point mutations. *Plant Physiol* 126, 480-484.
- Cooper, J. L., Till, B. J., and Henikoff, S. (2008). Fly-TILL: Reverse genetics using a living point mutation resource. *Fly (Austin)* 2, 300-302.
- Delcher, A. L., Kasif, S., Fleischmann, R. D., Peterson, J., White, O., and Salzberg,



- S. L. (1999). Alignment of whole genomes. *Nucleic Acids Res* 27, 2369-2376.
- Delcher, A. L., Phillippy, A., Carlton, J., and Salzberg, S. L. (2002). Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res* 30, 2478-2483.
- 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.
- Gusfield, D. (1997). *Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology* (Cambridge, UK: Cambridge University Press).
- Huang, A. M., Rehm, E. J., and Rubin, G. M. (2000). Recovery of DNA Sequences Flanking P-element Insertions: Inverse PCR and Plasmid Rescue. In *Drosophila Protocols*, Sullivan, W., M. Ashburner, and R. S. Hawley, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 429-238.
- Johnstone, O., and Lasko, P. (2004). Interaction with eIF5B is essential for Vasa function during development. *Development* 131, 4167-4178.
- Klattenhoff, C., Bratu, D. P., McGinnis-Schultz, N., Koppetsch, B. S., Cook, H. A., and Theurkauf, W. E. (2007). *Drosophila* rasiRNA pathway mutations disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response. *Dev Cell* 12, 45-55.
- Koundakjian, E. J., Cowan, D. M., Hardy, R. W., and Becker, A. H. (2004). The Zuker collection: a resource for the analysis of autosomal gene function in

- Drosophila melanogaster*. *Genetics* 167, 203-206.
- 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.
- Liang, L., Diehl-Jones, W., and Lasko, P. (1994). Localization of vasa protein to the *Drosophila* pole plasm is independent of its RNA-binding and helicase activities. *Development* 120, 1201-1211.
- Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H., and Siomi, M. C. (2005). Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. *Genes Dev* 19, 2837-2848.
- 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.
- Schulze, S. R., Sinclair, D. A., Fitzpatrick, K. A., and Honda, B. M. (2005). A genetic and molecular characterization of two proximal heterochromatic genes on chromosome 3 of *Drosophila melanogaster*. *Genetics* 169, 2165-2177.
- Schupbach, T., and Wieschaus, E. (1991). Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology. *Genetics* 129, 1119-1136.
- Theurkauf, W. E. (1994). Immunofluorescence analysis of the cytoskeleton during oogenesis and early embryogenesis. *Methods Cell Biol* 44, 489-505.
- Till, B. J., Colbert, T., Tompa, R., Enns, L. C., Codomo, C. A., Johnson, J. E., Reynolds, S. H., Henikoff, J. G., Greene, E. A., Steine, M. N., Comai, L., and

- Henikoff, S. (2003). High-throughput TILLING for functional genomics. *Methods Mol Biol* 236, 205-220.
- Tomari, Y., Du, T., Haley, B., Schwarz, D. S., Bennett, R., Cook, H. A., Koppetsch, B. S., Theurkauf, W. E., and Zamore, P. D. (2004). RISC assembly defects in the *Drosophila* RNAi mutant armitage. *Cell* 116, 831-841.
- Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P., and Sharp, P. A. (1999). Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev* 13, 3191-3197.
- 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.
- Williams, R. W., and Rubin, G. M. (2002). ARGONAUTE1 is required for efficient RNA interference in *Drosophila* embryos. *Proc Natl Acad Sci U S A* 99, 6889-6894.
- Winkler, S., Schwabedissen, A., Backasch, D., Bokel, C., Seidel, C., Bonisch, S., Furthauer, M., Kuhrs, A., Cobreros, L., Brand, M., and Gonzalez-Gaitan, M. (2005). Target-selected mutant screen by TILLING in *Drosophila*. *Genome Res* 15, 718-723.

**Table S1.** Female fertility, hatch-rates, and patterning defects in *ago3* mutants and *ago3* mutants rescued with Ago3-expressing transgenes. FM, amino-terminal flag-myc tag. The rate of egg hatching was only partially rescued, perhaps because a fully wild-type distribution or concentration of Ago3 requires the native *ago3* promoter or untranslated mRNA sequences.

Maternal genotype	Total eggs	Hatch Rate (%)	Dorsal appendage phenotype (%)		
			Wild-type appendages	appendages fused	no appendages
Oregon R	3,880	96.1	100	0	0
<i>bw; st</i>	3,214	70.5	98.4	1.1	0.6
<i>ago3/TM6B, Tb</i>	3,545	76.1	94.8	2.9	2.3
<i>ago3<sup>t2</sup>/ago3<sup>t3</sup></i>	398	0	33.2	30.2	36.7
<i>ago3<sup>t1</sup>/Df(3L)TTT</i>	7,329	0	53.4	25.3	21.3
<i>ago3<sup>t2</sup>/Df(3L)TTT</i>	5,514	0	61.1	28.6	10.2
<i>ago3<sup>t3</sup>/Df(3L)TTT</i>	3,233	0	58.2	28.1	13.7
<i>actin5c-Gal4-VP16/UASp-FM-Ago3-5; +/+</i>	4,418	94.6	97.1	1.4	1.5
<i>actin5c-Gal4-VP16/UASp-FM-Ago3-5; ago3/TM3, Sb</i>	3,956	93.5	98.4	0.8	0.7
<i>actin5c-Gal4-VP16/UASp-FM-Ago3-5; ago3<sup>t3</sup>/ago3<sup>t3</sup></i>	2,982	15.3	96.3	1.7	2.0
<i>nanos-Gal4-VP16/UASp-FM-Ago3-5; +/+</i>	5,915	86.2	96.5	2.0	1.5
<i>nanos-Gal4-VP16/UASp-FM-Ago3-5; ago3/TM3, Sb</i>	7,804	88.2	95.9	1.8	2.4
<i>nanos-Gal4-VP16/UASp-FM-Ago3-5; ago3<sup>t2</sup>/ago3<sup>t3</sup></i>	3,922	4.0	95.9	2.1	2.0

**Table S2A.** Sequencing statistics: reads. miRNAs were depleted from total RNA by oxidation or piRNAs were enriched by immunoprecipitation (I.P.) using antibodies specific for Ago3, Aub, or Piwi protein. "Total small RNA reads" correspond to genome-matching reads after excluding annotated non-coding RNAs (ncRNAs), such as rRNA, snRNA, snoRNA, or tRNA. "Transposon-matching reads" correspond to small RNAs mapped to *Drosophila melanogaster* transposons.

Ovary genotype	Oxidized or immunoprecipitated	Total reads	Reads perfectly matching genome	Reads matching annotated ncRNAs	Small RNA reads (excluding ncRNAs)	Pre-miRNA-matching reads	Reads excluding ncRNA and pre-miRNA-matching	23–29 nt small RNA reads			
								Total	Transposon-matching		
									Total	Sense	Antisense
Oregon R	oxidized	2,632,312	2,001,043	293,464	1,707,579	16,592	1,690,987	1,324,566	1,049,766	287,306	773,261
ago3/TM6B	oxidized	1,863,176	1,394,761	327,474	1,067,287	6,667	1,060,620	876,455	693,724	137,578	563,200
ago3 <sup>12</sup> /ago3 <sup>13</sup>	oxidized	856,039	699,107	400,133	298,974	7,005	291,969	202,490	146,307	50,772	97,096
ago3/TM6B	Ago3 I.P.	5,268,201	1,536,789	768,350	768,439	89,537	678,902	395,070	146,959	103,438	44,645
	Aub I.P.	2,839,217	2,147,087	451,786	1,695,301	50,603	1,644,698	1,409,665	981,950	205,576	786,673
	Piwi I.P.	2,420,222	1,862,025	243,907	1,618,118	56,438	1,561,680	1,382,576	1,002,912	219,507	789,947
ago3 <sup>12</sup> /ago3 <sup>13</sup>	Aub I.P.	3,003,999	2,306,992	938,452	1,368,540	50,054	1,318,486	929,370	301,476	177,452	127,218
	Piwi I.P.	2,601,550	2,000,239	674,739	1,325,500	43,861	1,281,639	858,076	248,859	101,033	149,412
aub/CyO	oxidized	5,433,382	3,689,928	2,393,685	1,296,243	44,138	1,252,105	569,032	386,491	116,403	273,632
aub <sup>HN2</sup> /aub <sup>QC42</sup>	oxidized	4,621,929	3,110,675	2,283,336	827,339	52,258	775,081	309,848	171,667	51,081	122,162

**Table S2B.** Sequencing statistics: species. miRNAs were depleted from total RNA by oxidation, and piRNAs were enriched by immunoprecipitation (I.P.) using antibodies specific for Ago3, Aub, or Piwi protein. "Total small RNA reads" correspond to genome-matching reads after excluding annotated non-coding RNAs (ncRNAs), such as rRNA, snRNA, snoRNA, or tRNA. "Transposon-matching reads" correspond to small RNAs mapped to *Drosophila melanogaster* transposons.

Ovary genotype	Oxidized or immunoprecipitated	Total species	Species perfectly matching genome	Species matching annotated ncRNAs	Small RNA species (excluding ncRNAs)	Pre-miRNA-matching species	Small RNA species, excluding ncRNA and pre-miRNA-matching	23–29 nt small RNA species			
								Total	Transposon-matching		
									Total	Sense	Antisense
Oregon R	oxidized	1,050,474	650,595	20,599	629,996	552	629,444	449,802	316,692	104,618	214,589
<i>ago3</i> /TM6B	oxidized	739,202	480,591	16,707	463,884	411	463,473	361,275	257,129	67,205	191,992
<i>ago3<sup>ts2</sup></i> / <i>ago3<sup>ts3</sup></i>	oxidized	256,053	172,904	17,707	155,197	461	154,736	110,154	72,355	25,014	47,919
<i>ago3</i> /TM6B	Ago3 I.P.	1,142,088	207,949	32,401	175,548	656	174,892	100,597	32,859	19,555	13,599
	Aub I.P.	782,728	421,529	20,848	400,681	510	400,171	322,760	190,743	58,944	133,398
	Piwi I.P.	1,047,007	665,608	25,746	639,862	518	639,344	546,271	343,820	102,279	244,077
<i>ago3<sup>ts2</sup></i> / <i>ago3<sup>ts3</sup></i>	Aub I.P.	1,169,524	716,569	43,355	673,214	665	672,549	457,111	84,835	48,706	36,882
	Piwi I.P.	1,228,783	801,931	39,446	762,485	601	761,884	508,121	129,715	56,252	74,368
<i>aub</i> /CyO	oxidized	1,797,796	610,612	43,306	567,306	756	566,550	263,243	185,700	63,357	123,801
<i>aub<sup>HN2</sup></i> / <i>aub<sup>OC42</sup></i>	oxidized	1,372,955	412,058	39,926	372,132	781	371,351	160,716	107,017	34,570	73,348

**Table S3.** The fractions of sense and antisense piRNAs that are uniquely bound to each PIWI protein.

**Overall**

genotype	Ago3-bound		Aub-bound		Piwi	
	sense	antisense	sense	antisense	sense	antisense
<i>ago3</i> /TM6B	0.76	0.24	0.29	0.71	0.27	0.73
<i>ago3<sup>t2</sup>/ago3<sup>t3</sup></i>			0.59	0.41	0.38	0.62

**Group I piRNAs**

genotype	Ago3-bound		Aub-bound		Piwi-bound	
	sense	antisense	sense	antisense	sense	antisense
<i>ago3</i> /TM6B	0.78	0.22	0.29	0.71	0.31	0.69
<i>ago3<sup>t2</sup>/ago3<sup>t3</sup></i>			0.59	0.41	0.52	0.48

**Group II piRNAs**

genotype	Ago3-bound		Aub-bound		Piwi-bound	
	sense	antisense	sense	antisense	sense	antisense
<i>ago3</i> /TM6B	0.26	0.74	0.73	0.27	0.70	0.30
<i>ago3<sup>t2</sup>/ago3<sup>t3</sup></i>			0.47	0.53	0.54	0.46

**Group III piRNAs**

genotype	Ago3-bound		Aub-bound		Piwi-bound	
	sense	antisense	sense	antisense	sense	antisense
<i>ago3</i> /TM6B	0.75	0.25	0.19	0.81	0.12	0.88
<i>ago3<sup>t2</sup>/ago3<sup>t3</sup></i>			0.64	0.36	0.10	0.90

**Table S4.** Probes for Northern hybridization

Small RNA detected	Probe sequence (DNA or RNA)
Aub-bound <i>roo</i> piRNA	5'-AAGAAATCAGTAATGCACTCTAGTA-3'
Piwi-bound <i>roo</i> piRNA	5'-GACTATTTACTTAGGCCTCTGCGTA-3'
Aub-bound <i>roo</i> piRNA	5'-AAGAAAUCAGUAAUGCACUCUAGUA-3'
Piwi-bound <i>roo</i> piRNA	5'-GACUAUUUACUUAGGCCUCUGCGUA-3'
<i>Su(Ste)</i> piRNA	5'-TCGGGCTTGTTCTACGACGATGAGA-3'
AT-chX-1 piRNA	5'-GCCCGAGCCGTCTAACGATGAAACA-3'
miR-8	5'-GACATCTTTACCTGACAGTATTA-3'
2S rRNA	5'-TACAACCCTCAACCATATGTAGTCCAAGCA-3'



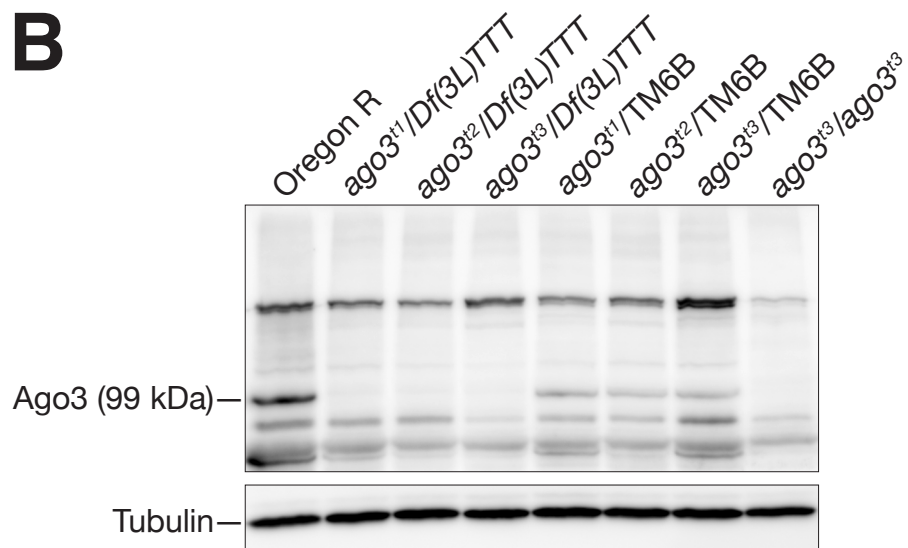
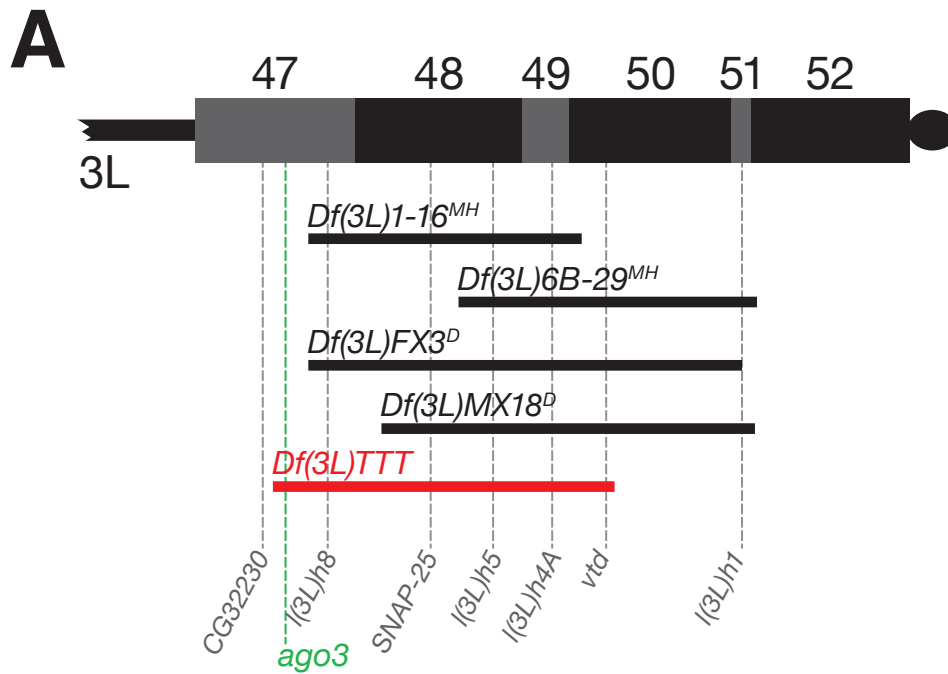
**Table S5.** Primers for quantitative RT-PCR.

RNA detected	Forward primer sequence	Reverse primer sequence
<i>412</i>	5'-CACCGGTTTGGTCGAAAG-3'	5'-GGACATGCCTGGTATTTTGG-3'
<i>Accord</i>	5'-ACAATCCACCAACAGCAACA-3'	5'-AAAAGCCAAAATGTCGGTTG-3'
<i>Accord2</i>	5'-TTGCTTTCGGACTTCGTCTT-3'	5'-TTCCACAACGAAAAACAACCA-3'
<i>Actin5c</i>	5'-AAGTTGCTGCTCTGGTTGTCG-3'	5'-GCCACACGCAGCTCATTGTAG-3'
<i>Blood</i>	5'-TGCCACAGTACCTGATTTTCG-3'	5'-GATTCGCCTTTTACGTTTGC-3'
<i>Diver</i>	5'-GGCACCACATAGACACATCG-3'	5'-GTGGTTTGCATAGCCAGGAT-3'
<i>Diver2</i>	5'-CTTCAGCCAGCAAGGAAAAC-3'	5'-CTGGCAGTCGGGTGTAATTT-3'
<i>gtwin</i>	5'-TTCGCACAAGCGATGATAAG-3'	5'-GATTGTTGTACGGCGACCTT-3'
<i>gypsy</i>	5'-GTTTCATACCCTTGGTAGTAGC-3'	5'-CAACTTACGCATATGTGAGT-3'
<i>gypsy6</i>	5'-GACAAGGGCATAACCGATACTGTGGA-3'	5'-AATGATTCTGTTCGGACTTCCGTCT-3'
<i>HeT-A</i>	5'-CGCGCGGAACCCATCTTCAGA-3'	5'-CGCCGCAGTCGTTTGGTGAGT-3'
<i>Hopper</i>	5'-GGCTGGCTTCAACAAAAGAA-3'	5'-GGACTCCCAGAAAACGTCATA-3'
<i>I-element</i>	5'-GACCAAATAAAAATAATACGACTTC-3'	5'-AACTAATTGCTGGCTTGTATG-3'
<i>Invader1</i>	5'-GTACCGTTTTTGTAGCCCGTA-3'	5'-AACTACGTTGCCCATTTCTGG-3'
<i>Max</i>	5'-TCTAGCCAGTCGAGGCGTAT-3'	5'-TGGAAGAGTGTGCGCTTTGTG-3'
<i>mdg1</i>	5'-AACAGAAACGCCAGCAACAGC-3'	5'-CGTTCCCATGTCCGTTGTGAT-3'
<i>R1A1</i>	5'-AATTCCCAGCTGTGCTAGA-3'	5'-GTCTCAAGGCACCTTTCAGC-3'
<i>rp49</i>	5'-CCGCTTCAAGGGACAGTATCTG-3'	5'-ATCTCGCCGAGTAAACGC-3'
<i>Rt1a</i>	5'-CCACACAGACTGAGGCAGAA-3'	5'-ACGCATAACTTTCCGGTTTG-3'
<i>ZAM</i>	5'-ACTTGACCTGGATACACTCACAAC-3'	5'-GAGTATTACGGCGACTAGGGATAC-3'

**Table S6.** Fly stocks used in this study.

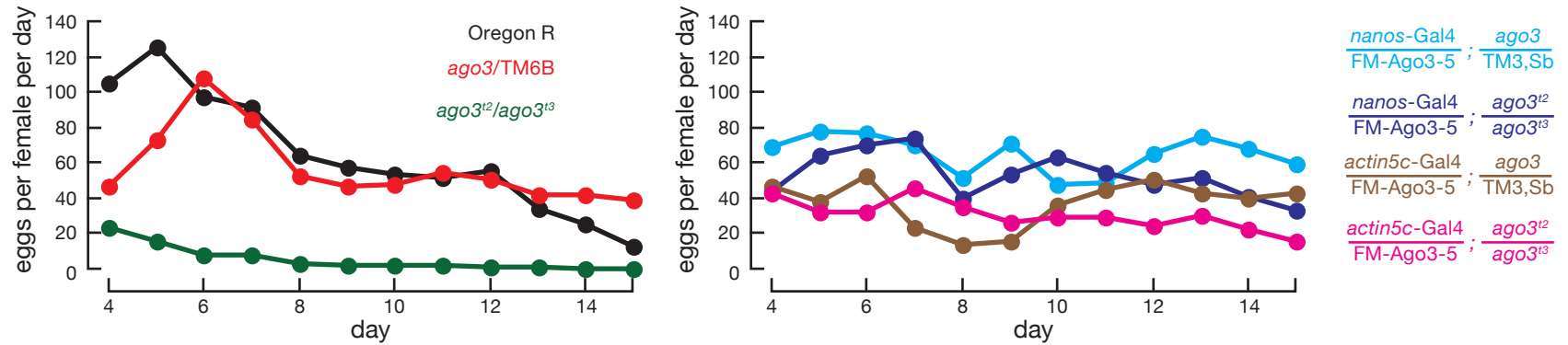
Genotype	Notes
<i>aub</i> <sup>HN2</sup> <i>cn bw</i> /CyO	Schupbach and Wieschaus (1991)
<i>aub</i> <sup>QC42</sup> <i>cn bw</i> /CyO	Schupbach and Wieschaus (1991)
<i>bw</i> ; <i>st ago3</i> <sup>t1</sup> /TM6B <i>Tb</i>	<i>ago3</i> <sup>t1</sup> allele
<i>bw</i> ; <i>st ago3</i> <sup>t2</sup> /TM6B <i>Tb</i>	<i>ago3</i> <sup>t2</sup> allele
<i>bw</i> ; <i>st ago3</i> <sup>t3</sup> /TM6B <i>Tb</i>	<i>ago3</i> <sup>t3</sup> allele
<i>Df(3L)TTT</i> /TM3 <i>Sb</i>	<i>ago3</i> deficiency
<i>w</i> ; <i>P[w</i> <sup>+</sup> , <i>actin5c-Gal4-VP16</i> ]/CyO; <i>st ago3</i> <sup>t2</sup> /TM3 <i>Sb</i>	to generate rescue flies
<i>w</i> ; <i>P[w</i> <sup>+</sup> , <i>nos-Gal4-VP16</i> ]/CyO; <i>st ago3</i> <sup>t2</sup> /TM3 <i>Sb</i>	to generate rescue flies
<i>w</i> ; <i>P[w</i> <sup>+</sup> , <i>UASp-FM-Ago3-5</i> ]/CyO; <i>st ago3</i> <sup>t3</sup> /TM, <i>Sb</i>	to generate rescue flies
Oregon R	wild-type control for sequencing
<i>w</i> <sup>1118</sup>	wild-type control for microarrays
<i>w</i> ; <i>Sp</i> /CyO; <i>GFP-Vasa</i> /TM3 <i>Sb</i>	Johnstone and Lasko (2004)

**Figure S1.** *ago3* alleles. (A) *Df(3L)TTT* deletes multiple genes (schulze et al., 2005), including the *ago3* locus, creating the *ago3*<sup>ΔTTT</sup> allele. (B) No full-length Ago3 protein was detected by Western blotting with anti-Ago3 antibody in any of the trans-heterozygous combinations of *Df(3L)TTT* with the three *ago3* EMS alleles isolated by tilling.

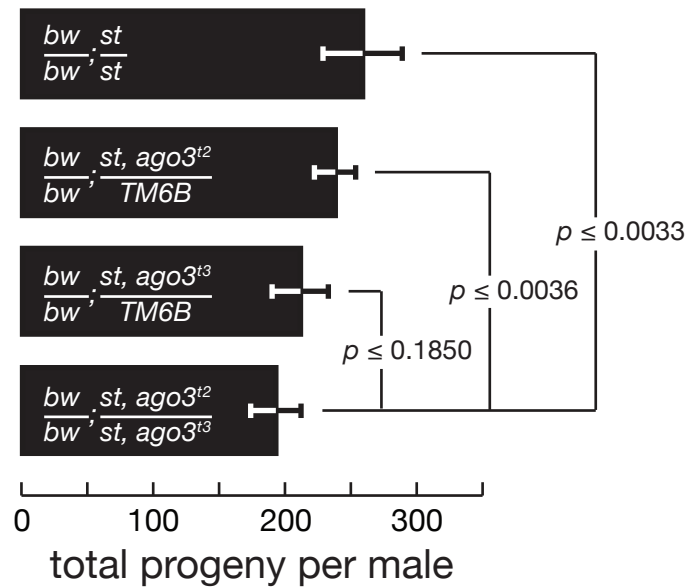


**Figure S2.** Fertility analysis. (A) *ago3* mutant females lay few eggs. The low egg production is rescued by an Ago3-expressing transgene, establishing that loss of Ago3 is the cause of the defect. (B) *ago3* mutant males have reduced fertility.

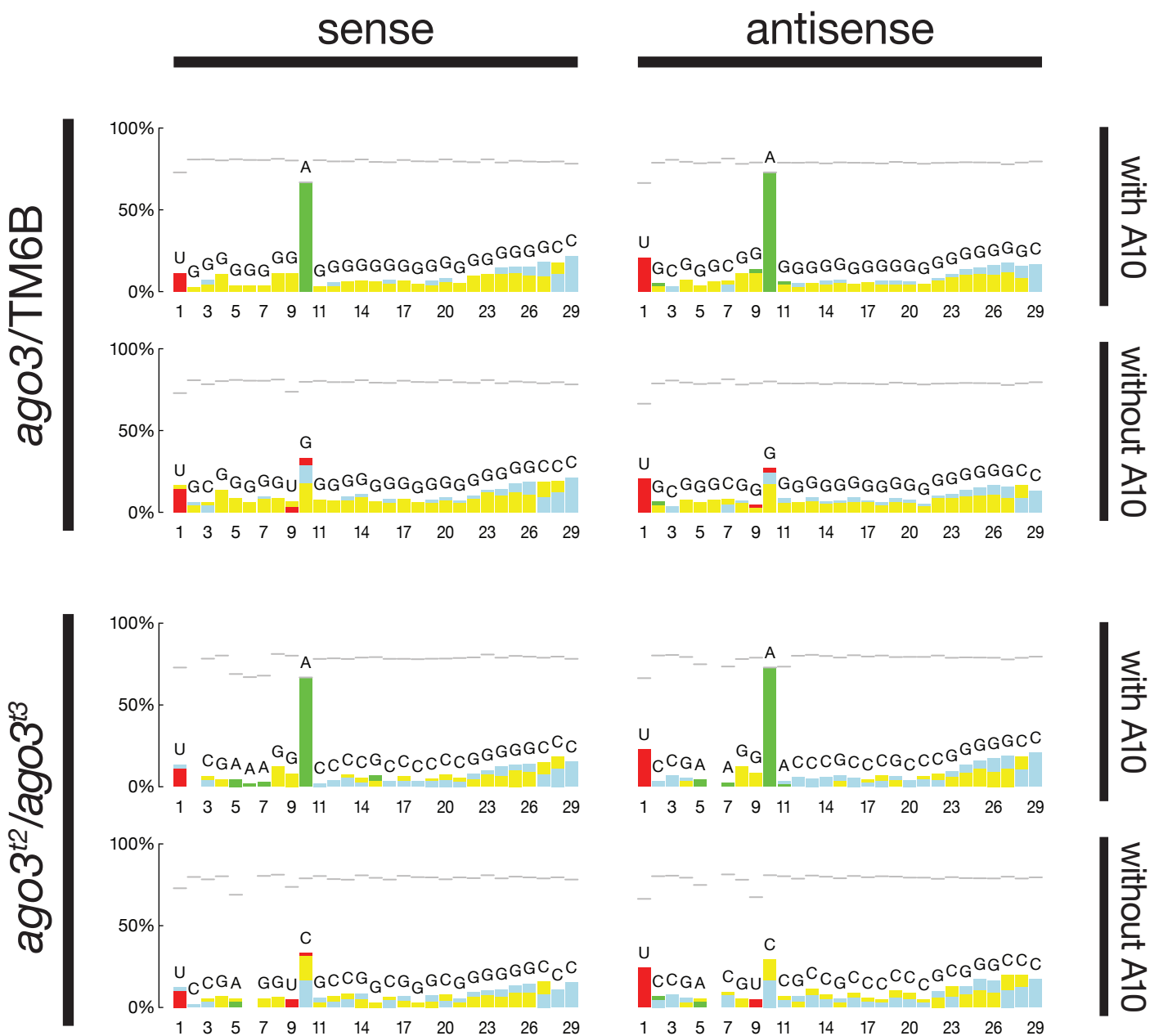
**A**



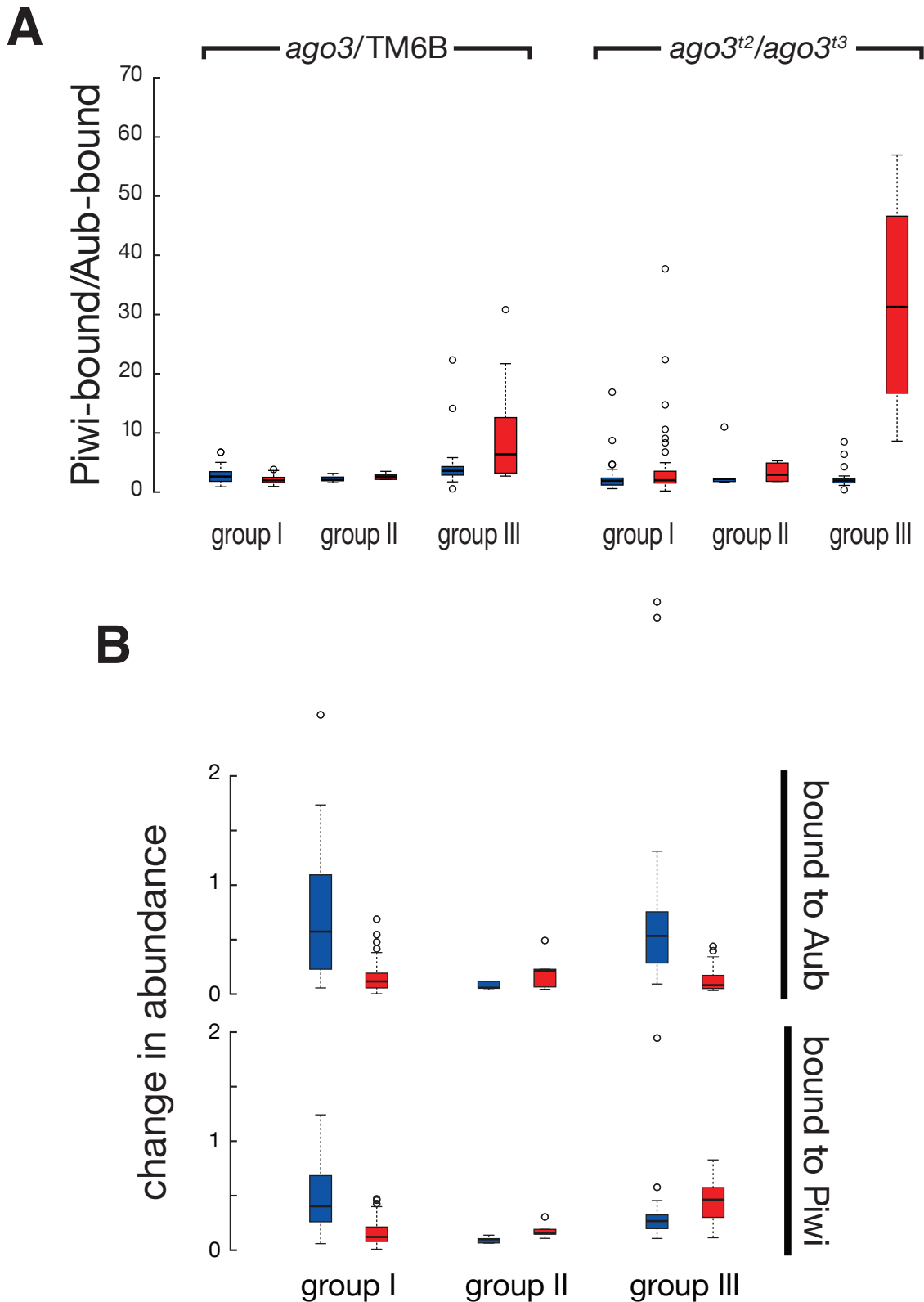
**B**



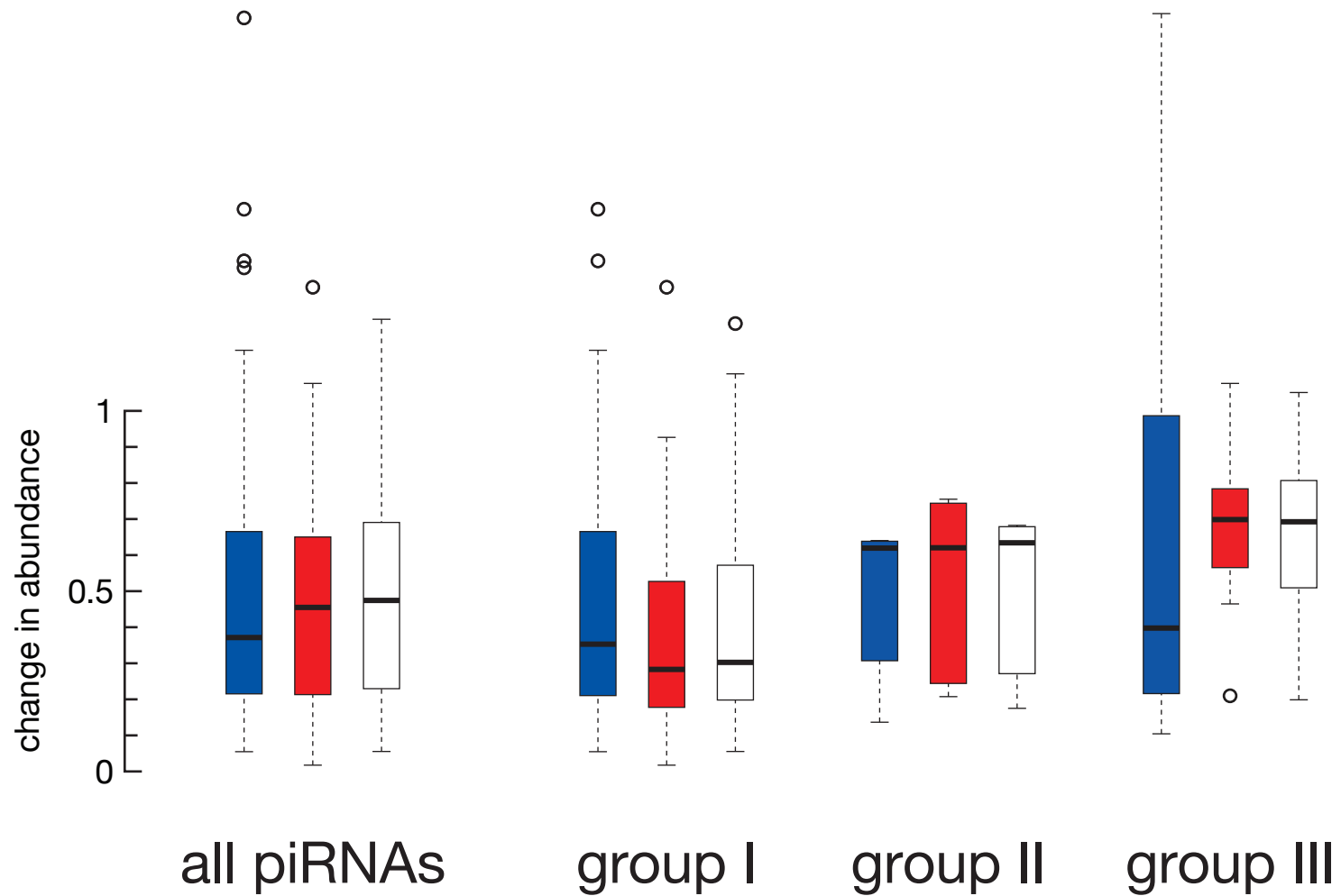
**Figure S7.** Sequence bias of piRNAs uniquely associated with Aub in *ago3/TM6B* and *ago3<sup>t2</sup>/ago3<sup>t3</sup>* ovaries. piRNAs were separated by orientation—sense or antisense—and by the identity of their tenth nucleotide—A or not A. All four classes of piRNAs in both genotypes were more likely to begin with U than expected from the bias inherent in the underlying bias among all annotated transposon families in *Drosophila melanogaster*, irrespective of the identity of their tenth nucleotide. The analysis reports the percentage of each nucleotide in excess of the inherent bias of all transposon sequences for each piRNA position (see Supplemental Experimental Procedures). Significance testing ( $p = 0.001$ ) was Bonferroni corrected for four nucleotides at 29 positions, in two orientations (sense and antisense), among either three (heterozygotes) or two proteins (mutant). Gray bars indicate the maximum possible value for each position.



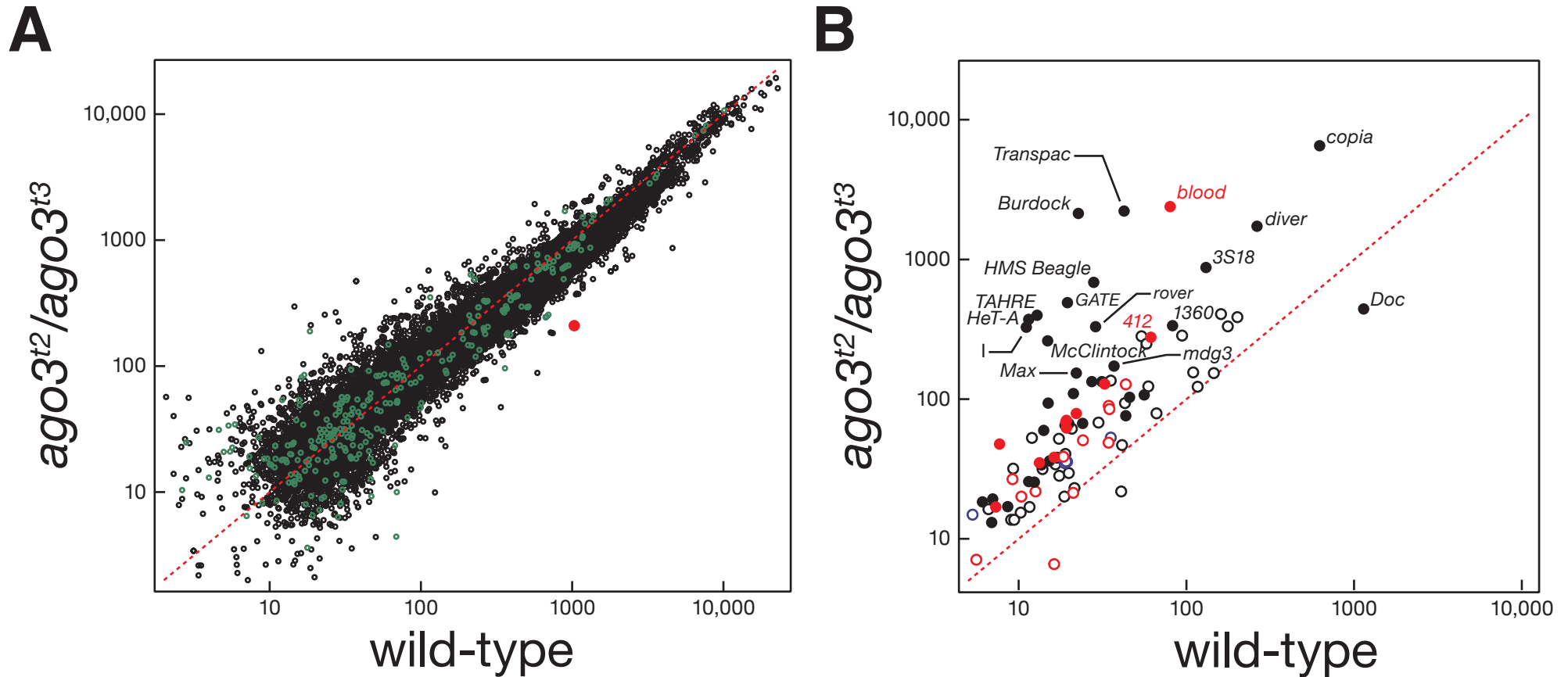
**Figure S8.** (A) The ratio of uniquely Piwi-bound to uniquely Aub-bound piRNAs in ovaries. Median values for *ago3*/TM6B ovaries were 2.64 (sense) and 2.00 (antisense) for group I; 2.06 (sense) and 2.63 (antisense) for group II; and 3.59 (sense) and 6.39 (antisense) for group III. Median values for *ago3<sup>t2</sup>/ago3<sup>t3</sup>* ovaries were 1.92 (sense) and 2.00 (antisense) for group I; 2.21 (sense) and 2.94 (antisense) for group II; and 1.95 (sense) and 31.3 (antisense) for group III. (B) Box plots illustrating the change in abundance for piRNAs, separated by group and uniquely bound to Aub or Piwi in *ago3<sup>t2</sup>/ago3<sup>t3</sup>* versus *ago3*/TM6B for the 95 transposon families for which we tallied  $\geq 500$  piRNA reads in *ago3*/TM6B. Blue, sense; red, antisense.



**Figure S9.** Box plots illustrating the change in abundance for piRNAs, separated by group, in *aub<sup>HN2</sup>/aub<sup>QC42</sup>* versus *aub/CyO* for the 95 transposon families for which we tallied  $\geq 500$  piRNA reads in *ago3/TM6B*. The median change for sense, antisense, and total for all piRNAs was 0.37, 0.45, 0.47; for group I, 0.35, 0.28, 0.30; for group II, 0.62, 0.62, 0.63; and for group III, 0.40, 0.70, 0.69. Sense, antisense, total.

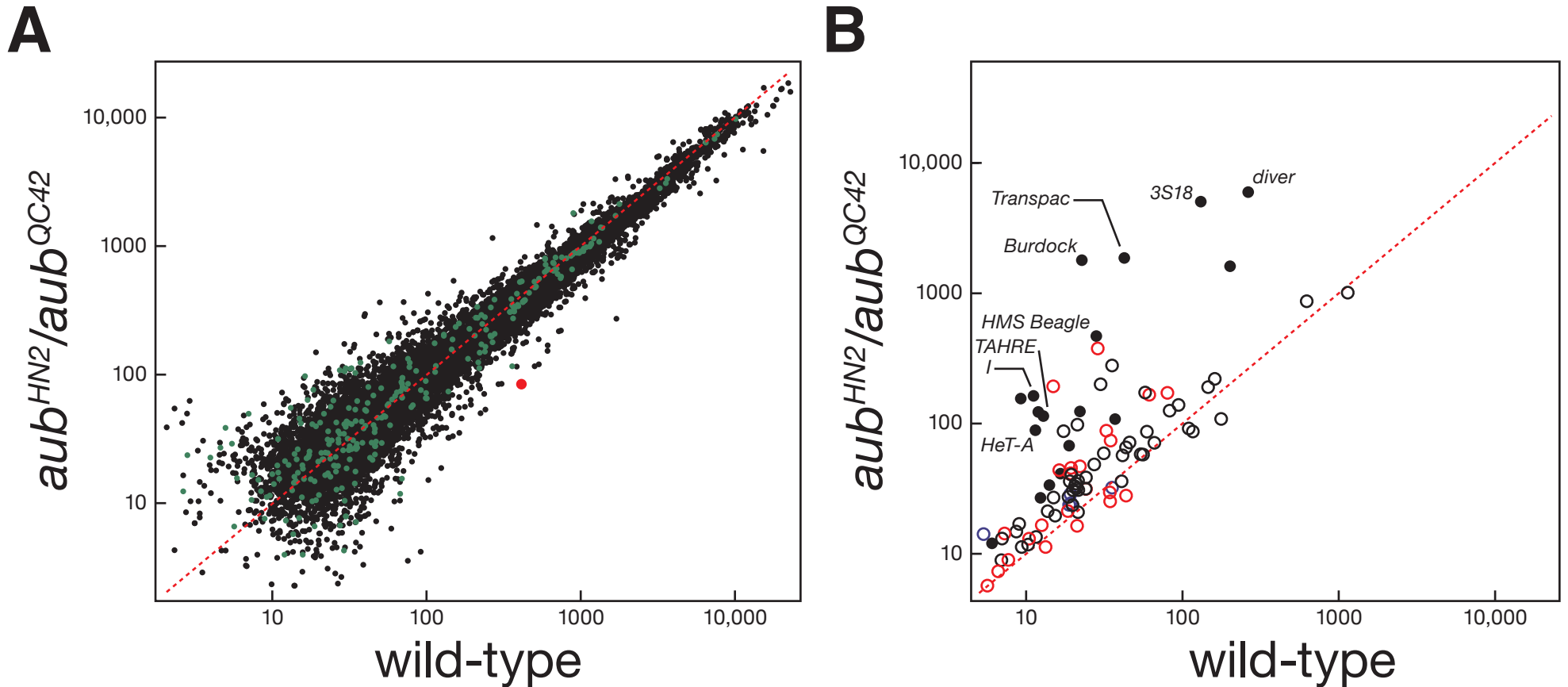


**Figure S11.** Tiling microarray analysis of mRNA and transposon expression in wild-type ( $w^{1118}$ ) and *ago3* mutant ovaries. (A) mRNA expression. Black, euchromatic genes; green, heterochromatic genes; Red, the *ago3* mRNA. All of the changes detected for mRNA corresponded to  $FDR > 0.02$ , suggesting that loss of Ago3 has no significant effect on mRNA expression. The  $FDR$  for *ago3* mRNA was 0.089. (B) Transposon expression. Group I transposons, black circles; group II transposons, blue circles; group III transposons, red circles. Open circles,  $FDR > 0.02$ ; filled circles,  $FDR < 0.02$ .





**Figure S12.** Tiling microarray analysis of mRNA and transposon expression in wild-type ( $w^{1118}$ ) and *aub* mutant ovaries. (A) mRNA expression. Black, euchromatic genes; green, heterochromatic genes; Red, the *aub* mRNA. All of the changes detected for mRNA corresponded to FDR > 0.02, suggesting that loss of Aub has no significant effect on mRNA expression. (B) Transposon expression. Group I transposons, black circles; group II transposons, blue circles; group III transposons, red circles. Open circles, FDR > 0.02; filled circles, FDR < 0.02.



**Figure S13.** Transposon expression in *aub<sup>HN2</sup>/aub<sup>QC42</sup>* mutants. Quantitative RT-PCR was used to assess the change in expression, relative to *actin*, between *aub* heterozygous and homozygous ovaries. The average  $\pm$  standard deviation for at least three independent biological samples are shown.

