Heterochromatin-dependent transcription of satellite DNAs in the *Drosophila melanogaster* female germline

Xiaolu Wei[1]*, Danna G. Eickbush[2], Iain Speece[2], Amanda M. Larracuente[2]*

*corresponding authors: xiaolu_wei@urmc.rochester.edu, alarracu@bio.rochester.edu

Author affiliations:

1. Department of Biomedical Genetics, University of Rochester Medical Center, Rochester, NY 14642

2. Department of Biology, University of Rochester, Rochester, NY 14627

ORCID AML 0000-0001-5944-5686
ORCID XW 0000-0001-9952-3757

Running head: Drosophila satellite DNA regulation

Keywords: satellite DNA, piRNAs, heterochromatin, ovary, testes, *Drosophila melanogaster*

ABSTRACT

Large blocks of tandemly repeated DNAs—satellite DNAs (satDNAs)—play important roles in heterochromatin formation and chromosome segregation. We know little about how satDNAs are regulated, however their misregulation is associated with genomic instability and human diseases. We use the *Drosophila melanogaster* germline as a model to study the regulation of satDNA transcription and chromatin. Here we show that complex satDNAs (>100-bp repeat units) are transcribed into long noncoding RNAs and processed into piRNAs (PIWI interacting RNAs). This satDNA piRNA production depends on the Rhino-Deadlock-Cutoff complex and the transcription factor Moonshiner—a previously-described non-canonical pathway that licenses heterochromatin-dependent transcription of dual-strand piRNA clusters. We show that this pathway is important for establishing heterochromatin at satDNAs. Therefore, satDNAs are regulated by piRNAs originating from their own genomic loci. This novel mechanism of satDNA regulation provides insight into the role of piRNA pathways in heterochromatin formation and genome stability.

INTRODUCTION

Repetitive DNA makes up a large fraction of eukaryotic genomes (Britten and Kohne 1968). Most repeat-dense genomic regions are gene-poor and tightly packed into heterochromatin (reviewed in (Allshire and Madhani 2018; Janssen, et al. 2018)). Tandem arrays of repeated sequences called satellite DNAs (satDNAs) are abundant in the heterochromatin of the pericentromeres, subtelomeres, and on sex chromosomes (Charlesworth, et al. 1994; Schmidt 1998; Schueler, et al. 2001). SatDNAs are typically viewed as selfish genetic elements that can spread rapidly in genomes and are generally repressed (Doolittle and Sapienza 1980; Orgel and Crick 1980). However, satDNAs play roles in chromatin structure, chromosome segregation, and genome stability across species (Weiler and Wakimoto 1995; Dernburg, et al. 1996; Lippman, et al. 2004; Bouzinba-Segard, et al. 2006; Zhu, et al. 2011; Swanson, et al. 2013; Plohl, et al. 2014; Rosic, et al. 2014). SatDNA-derived transcripts have been detected in many species (Ugarkovic 2005; Usakin, et al. 2007; Biscotti, et al. 2015; Ferreira, et al. 2015). In insects, these transcripts may have roles in early embryos (Pathak, et al. 2013; Halbach, et al. 2020) and spermatogenesis (Mills, et al. 2019). Although the regulation and function of repeat-derived transcripts remains poorly understood, across organisms they may be generally important for maintaining genome stability and integrity (reviewed in (Janssen, et al. 2018)).

Small RNA pathways protect genome integrity by silencing repeats like transposable elements (TEs)—genomic parasites that mobilize and can cause genome instability. In these pathways, small RNAs guide Argonaute proteins to cleave mRNA or silence genomic DNA by targeting complementary sequences (Hutvagner and Simard 2008). Among the most abundant types of repeat-derived small RNAs in animal germlines are the 23–32-nt PIWI-interacting RNAs (piRNAs) (Aravin, et al. 2006; Girard, et al.

2006; Grivna, et al. 2006; Lau, et al. 2006; Brennecke, et al. 2007; Houwing, et al. 2007). These piRNAs are particularly well-studied in Drosophila ovaries. The piRNA precursors are transcribed from discrete genomic loci called piRNA clusters that primarily contain truncated TE sequences. The piRNAs derived from these loci repress TE activity through both posttranscriptional (Brennecke, et al. 2007; Gunawardane, et al. 2007) and transcriptional gene silencing (Wang and Elgin 2011; Sienski, et al. 2012; Le Thomas, et al. 2013; Rozhkov, et al. 2013). There are two types of piRNA sources in Drosophila ovaries—uni-strand and dual-strand piRNA clusters. Uni-strand piRNA clusters require promoter sequences and are either expressed only in somatic tissues (*e.g. flamenco*), or in both somatic tissues and the germline (*e.g. 20A*) (Brennecke, et al. 2007; Malone, et al. 2009; Mohn, et al. 2014). However, most piRNA clusters are heterochromatic dual-strand clusters, which are bidirectionally transcribed and do not necessarily require promoters (*e.g. 42AB*, *80F*, and *38C1/2* (Brennecke, et al. 2007; Mohn, et al. 2014; Andersen, et al. 2017)). Dual-strand piRNA clusters are expressed primarily in the germline, where their transcription is licensed by a non-canonical pathway that depends on the heterochromatin protein-1 (HP1) variant called Rhino (Rhi) (Klattenhoff, et al. 2009; Zhang, et al. 2014). Rhi recruits Deadlock (Del), an unstructured linker protein (Wehr, et al. 2006; Czech, et al. 2013), and Cutoff (Cuff), a protein related to the yeast Rai1 decapping enzyme (Pane, et al. 2011), to H3K9me3 chromatin. This complex is referred to as the Rhino, Deadlock, and Cutoff (RDC) protein complex (Mohn, et al. 2014). Moonshiner (Moon)—a paralog of the transcription factor TFIIA-L—interacts with Del and recruits TBP-related factor 2 (TRF2) to initiate transcription of dual-strand piRNA clusters (Andersen, et al. 2017). Most piRNA studies in Drosophila focus on their important role in repressing TE activity to protect genome integrity (*e.g.* (Brennecke, et al. 2007)). Given that TEs and

satDNAs both are abundant repeats in heterochromatin whose activities are associated with genomic instability, we suspect that satDNAs may also be regulated by this piRNA pathway.

Consistent with our hypothesis that piRNA pathways regulate satDNAs, small RNAs derived from satDNAs exist in germlines (*e.g.* (Aravin, et al. 2003)). However, little is known about these satDNA-derived small RNAs. Here we leverage publicly-available RNA-seq and ChIP-seq datasets, and complement these data with cytological and molecular analyses of expression to study the regulation of satDNAs in the germline. SatDNAs are categorized based on their repeat unit size as simple (1-10bp) or complex (>100bp). We focus on two abundant families of complex satDNA in *D. melanogaster: Responder* (*Rsp*) and satellites in the *1.688* g/cm^3 family (*1.688*). We show that complex satDNAs are expressed and processed primarily into piRNAs in both testes and ovaries. In ovaries, this expression is dependent on the RDC complex and Moon—important proteins that regulate dual-strand piRNA clusters. Disruptions of the piRNA pathway lead to a loss of both satDNA-derived piRNAs and heterochromatin marks at satDNA loci. Our analyses suggest a model where the establishment of heterochromatin at satDNA is regulated by piRNAs originating from their own genomic loci. These results add insight into the role of piRNA pathways in heterochromatin formation and genome stability.

RESULTS AND DISCUSSION

Positive correlation between satDNA transcription and genomic copy number

To study satDNA expression patterns, we characterized transcripts from two representative complex satDNA families in *D. melanogaster*—*Rsp* and *1.688*—across tissues and developmental time points. *Rsp* consists of a dimer of two closely-related ~120-bp repeats in the pericentric heterochromatin on chromosome 2R of *D. melanogaster* (Wu, et al. 1988; Pimpinelli and Dimitri 1989). The *1.688* family of repeats are the most abundant complex satellite type in *D. melanogaster* (Lohe and Roberts 1988). They exist as discrete tandem arrays in the pericentric heterochromatin named after their repeat unit sizes on chromosome 2L (*260-bp*), chromosome 3L (*354-bp* and *356-bp*), and the X chromosome (*359-bp*) (Losada and Villasante 1996; Abad, et al. 2000).

We mined modENCODE datasets (Table S1 and (Graveley, et al. 2011; Brown, et al. 2014)) and found evidence for satDNA expression in total RNA-seq datasets from both sexes, and across different developmental stages (Figure 1-S1). Transcript abundance of both *Rsp* and *1.688* is low (RPM_{Rsp}<10 and $RPM_{1.688}$<250) and generally increases throughout development (Figure 1-S1A and C). We also observed an increase in transcript abundance with adult age—*e.g. Rsp* has ~1.5 and ~3.5-fold higher expression in 30-day-old adults compared to 1-day old adults in females and males, respectively (Figure 1-S1A and C). Both satDNA families are expressed in gonads, head, and other tissues (Figure 1-S1B and D). SatDNA-derived reads have very low abundance in the poly-A selected RNA-seq data (RPM_{Rsp}<0.2 and $RPM_{1.688}$<10; Table S2), indicating that the majority of satDNA transcripts are not polyadenylated.

We analyzed RNA-seq data to determine whether satellite transcription is uni- or bi-directional. The highly repetitive nature of satDNAs makes confidently assigning satellite-derived reads to a genomic location difficult, however, we can take advantage of our contiguous and validated contigs for the major *Rsp* locus on chromosome 2R and for the *260-bp* locus, a variant of *1.688*, on chromosome 2L (Khost, et al. 2017). Our analysis of these loci showed that reads mapped uniquely to both strands (Figure 1-S2A). We further analyzed all *Rsp* and *1.688* satellites in the genome, and revealed a nearly 1:1 ratio of reads from the plus and minus strands (Figure 1-S2B; all mapped and uniquely mapped reads). Together, these results suggest that satDNAs are transcribed from both strands.

We validated the presence of these transcripts using RNA fluorescence *in situ* hybridization (FISH). Both *Rsp* and *1.688* satellite transcripts are visible in testes and ovaries (Figure 1A; Figure 1-S3A). These signals are undetectable after treating with RNase A prior to probe hybridization (Figure 1-S3B), which degrades single stranded RNAs, or RNaseH post probe hybridization (Figure 1-S3C), which degrades the RNA in DNA-RNA hybrids, suggesting these signals are from RNA rather than DNA. We detected foci of satDNA transcript signal in ovarian nurse cells and in pre-meiotic testicular germ cells. Interestingly, in testes we detected *Rsp* signal in earlier stages of spermatogenesis (*i.e.*, germline stem cells/spermatogonia) than the *1.688* signals (*i.e.* primary spermatocytes; Figure 1A). The difference in timing is notable, as *Rsp* is the target of a well-known male meiotic drive system called *Segregation Distorter* (*SD*) (Sandler, et al. 1959) that causes a defect in post-meiotic germ cells (reviewed in (Larracuente and Presgraves 2012)). *Rsp* transcription may therefore play some specific role in the male germline distinct from other complex satDNA.

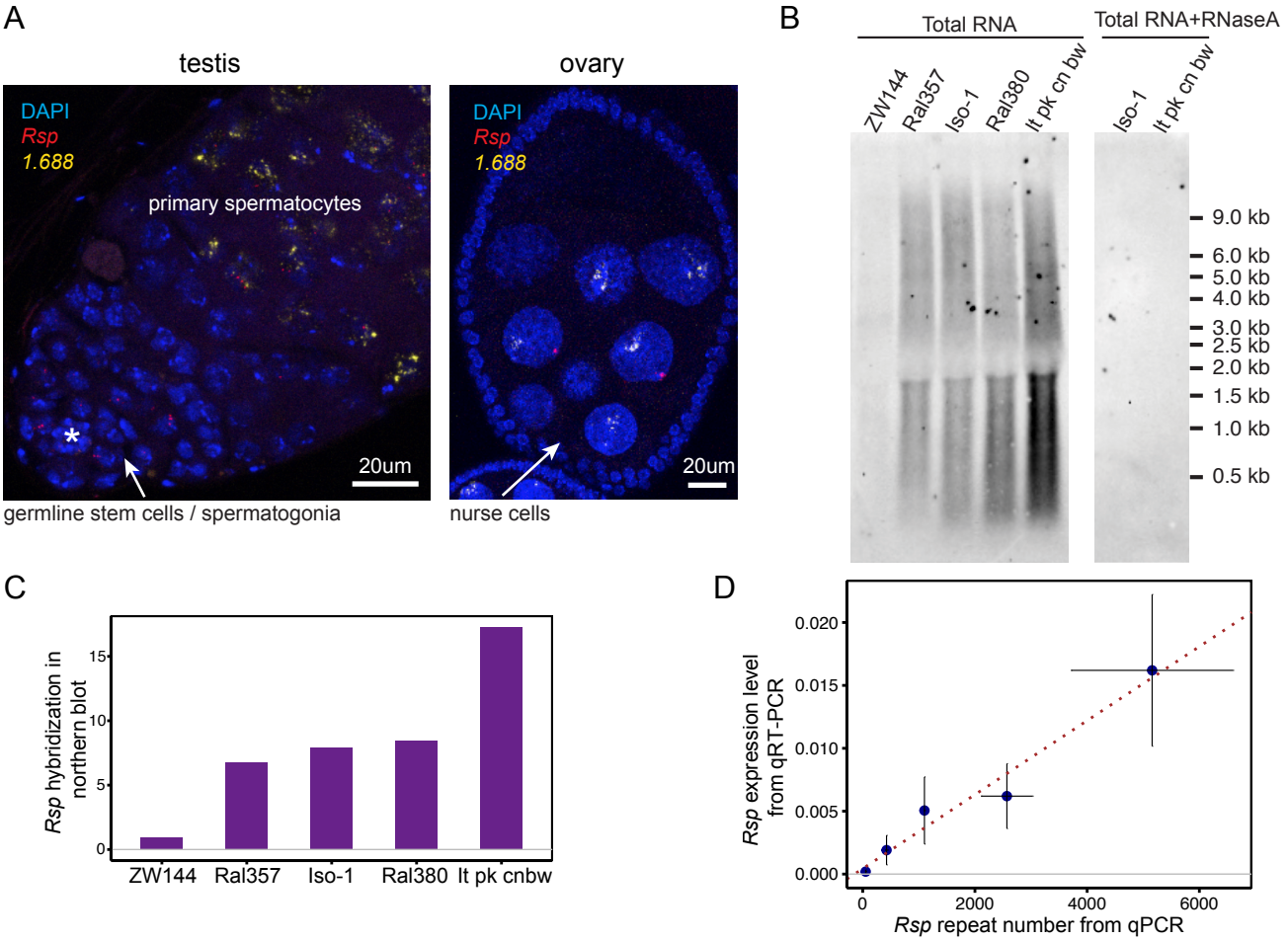


Figure 1. SatDNAs are expressed in ovaries and testes. (A) RNA FISH shows evidence for *Rsp* and *1.688*-derived transcripts in testes and ovaries, asterisk indicates the hub. (B) Northern blot probed with *Rsp.* Total RNA was extracted from ovaries of fly lines with varying copy numbers of *Rsp*: ZW144 (238 copies), Ral357 (640 copies), Iso-1 (1100 copies), Ral380 (2339), lt pk cn bw (4086). There is no signal after RNaseA treatment. (C) Northern signal quantification shows *Rsp* transcript abundance correlates with its genomic copy number (Pearson's correlation coefficient r^2=0.93, p-value=0.02); (D) qPCR and qRT-PCR quantification of *Rsp* copy number and expression level, respectively, of strains used in northern blot. A linear regression line is shown in the plot with red dotted line (Pearson's correlation coefficient r^2=0.98, p-value=0.003). Details for (C) and (D) in Table S4.

To determine whether the length and/or abundance of transcripts correlated with locus size, we performed a northern blot analysis using total RNA in ovaries from lines with different *Rsp* repeat copy numbers (Table S3; (Khost, et al. 2017)). We determined that while the range of transcript lengths was similar among these lines–from <300 nt to > 9000 nt (Figure 1B)–the abundance of *Rsp* transcripts correlated with the number of repeats (Figure 1C and Table S4, Pearson's correlation coefficient r^2=0.93, p-value=0.02). We validated the hybridization results using qPCR and qRT-PCR to quantify *Rsp* genomic DNA and RNA transcript abundance, respectively (Figure 1D; Table S4, Pearson's correlation coefficient r^2=0.98, p-value=0.003).

Satellite transcripts are processed into piRNAs in Drosophila germline

Many repeat-derived transcripts are processed into piRNAs (Aravin, et al. 2003; Brennecke, et al. 2007) and endo-siRNAs (Czech, et al. 2008; Ghildiyal, et al. 2008; Okamura, et al. 2008). To ask if complex satDNA-derived RNAs are processed into small RNAs, we re-analyzed published small RNA-seq data (Table S1; (Ghildiyal, et al. 2010; Rozhkov, et al. 2010; Fagegaltier, et al. 2014; Mohn, et al. 2014; Quenerch'du, et al. 2016; Andersen, et al. 2017; Parhad, et al. 2017)). We indeed detected satDNA-derived small RNAs in testes and ovaries (Figure 2-S1 A and B). The distribution of small RNA sizes is typical for piRNA populations: an average of 90% of the RNAs range from 23nt to 28nt, with a peak at 24-26nt in *D. melanogaster* (Brennecke, et al. 2007) (Figure 2A for *Rsp* and Figure 2-S1C for *1.688*). This size range and their abundance in the germline (Figure 2-S1 A and B) together suggest that the satDNA-derived small RNAs may primarily be piRNAs. In the Drosophila ovary, piRNAs are amplified through a piRNA-guided RNA cleavage process called the ping-pong cycle. piRNAs produced through this pathway show a signature of ping-pong: a 10nt-overlap

of antisense-sense piRNAs with a preference of uridine at the 5' end (1U) or adenosine at nucleotide position 10 (10A) (Brennecke, et al. 2007; Gunawardane, et al. 2007). Our analysis of the ovary small RNA-seq data (Mohn, et al. 2014; Andersen, et al. 2017; Parhad, et al. 2017) confirms a ping-pong signature for satDNA-derived small RNAs, Z-score=4.55 for *Rsp* and 6.85 for *1.688* satellite (Figure 2-S1E and G) and ~60-80% have 1U/10A (Figure 2-S1D and F).

The distribution of piRNA reads that pileup along the length of a single monomer (*359-bp* and *260-bp*) or dimer (*Rsp*) of these satDNA repeats, is not uniform (Figure 2-S2A and B). For example, a large fraction of piRNAs map to the 180-230-nt region of a *Rsp* dimer. This pattern could arise if these piRNAs came from repeat fragments overrepresented in the genome. However, when we extract and align all repeat variants from the genome, the alignment depth is more uniform across the monomers/dimers than the piRNA pileup (Figure 2-S2C). Moreover, we observe similar non-uniform patterns of piRNA pileup over germline-dominant TEs (*e.g. invader6*, *mdg3* and *Het-A* (Figure 2-S3)), suggesting that details of piRNA processing may drive the pattern. The piRNA read pileup pattern differs between ovaries and testes, especially for *Rsp* (Figure 2-S2A and B), suggesting that there may be differences in transcription machinery, precursor production, or precursor processing between ovaries and testes.

We then asked if satellite-derived small RNAs are bound by the PIWI proteins. We re-analyzed published Piwi, Aubergine (Aub), and Argonaute3 (Ago3) RIP-seq data from ovaries (Brennecke, et al. 2007; Mohn, et al. 2015; Sato, et al. 2015) and found evidence that *Rsp* and *1.688* RNAs interact with each of these proteins (Table S5). For example, ~0.4% and 0.1% of Piwi bound RNAs map to *359-bp* (a *1.688* variant) and *Rsp*, respectively. For comparison, ~2% and 17% Piwi bound RNAs mapped to

the dual-strand piRNA clusters *80F* and *42AB,* respectively. In contrast, only an average of 0.0005% of the reads from Piwi RIP-seq data mapped to miRNAs, which are abundant small RNAs not known to be bound by Piwi. This suggests that the abundance of satellite RNA in the RIP-seq data is not likely due to noise or contamination. Our results from Aub and Ago3 RIP data are similar to Piwi (Table S5; *e.g.* 1.6% and 0.1% of Aub-bound RNAs map to *359-bp* and *Rsp*, respectively; and 0.8% and 0.07% of Ago3-bound RNAs map to *359-bp* and *Rsp*, respectively). Together, these results further indicate that satellite-derived transcripts are processed into piRNAs in the female germline.

Our finding that satDNAs generate piRNAs from both strands, similar to dual-strand piRNA clusters, suggests that they may be regulated similarly. Dual-strand piRNA clusters are associated with the heterochromatin binding protein Rhi (Klattenhoff, et al. 2009; Zhang, et al. 2014). We therefore re-analyzed publicly available ChIP-seq datasets from ovaries (Zhang, et al. 2014; Parhad, et al. 2017) to test if satDNA regions are also Rhi associated. We found that the level of Rhi enrichment on satDNAs (*Rsp* and *1.688*) is similar to known dual-strand piRNA clusters (*42AB*, *38C1*/*2* and *80F*) (Figure 2B). Indeed, complex satDNAs are in the top 30% of all repeats enriched in Rhi (Figure 2B and Figure 2-S4A; full data in supplemental data files). In contrast, the uni-strand piRNA clusters (*20A* and *flamenco*) have lower Rhi enrichment compared to complex satDNAs (pairwise t-test with Benjamini Hochberg (Benjamini 1995) adjusted p-value=0.0081) and dual-strand piRNA clusters (pairwise t-test with Benjamini Hochberg (Benjamini 1995) adjusted p-value=0.0067). Our results for piRNA clusters are consistent with previous studies (Klattenhoff, et al. 2009; Mohn, et al. 2014; Andersen, et al. 2017). Unlike the uneven distribution of piRNAs along satellite monomers/dimers (Figure 2-S2A and B), the distribution of Rhi ChIP-seq

reads (Figure 2-S2D) is similar to the alignment depth of genomic repeats (Figure 2-S2C). This suggests that Rhi localizes to the large satDNA genomic loci, rather than a subset of smaller clusters or repeats across the genome (*e.g.* the 12 copies of *Rsp* inside an intron of *Ago3* on chromosome 3L; Figure 2-S2C) or in potentially unannotated piRNA clusters. We also find that satDNAs are enriched for H3K9me3 (datasets from (Klenov, et al. 2014; Le Thomas, et al. 2014; Mohn, et al. 2014)), which is consistent with their enrichment for Rhi (Figure 2-S4B).

While Rhi marks genomic piRNA-producing regions, how these piRNA clusters are established and maintained is an open question. Maternally deposited Piwi and Aub proteins, together with piRNAs, may play a role in defining regions for future piRNA production (Megosh, et al. 2006; Brennecke, et al. 2008; Rangan, et al. 2011; Le Thomas, et al. 2014). Depleting Piwi in the nucleus with *piwi* mutants lacking a nuclear localization signal (NLS) (Klenov, et al. 2014), or knocking down germline *piwi* (Le Thomas, et al. 2013; Mohn, et al. 2014) affects H3K9me3 level on a subset of active transposons, but not on piRNA clusters (Klenov, et al. 2014; Mohn, et al. 2014). However, transiently knocking down *piwi* expression earlier in the embryonic germline generally leads to a decrease of H3K9me3 level in heterochromatic regions, including dual-strand piRNA clusters in the adult ovary (Akkouche, et al. 2017), indicating that *piwi* is at least important for heterochromatin establishment. These *piwi* mutants also show an increase in spliced non-precursor transcripts, and a decrease in piRNA production from dual-strand piRNA clusters in the adult ovary (Akkouche, et al. 2017). Together, these studies suggest that Piwi and piRNAs play roles in establishing the chromatin state of piRNA clusters in early embryos but are dispensable for their maintenance (Czech, et al. 2018). To test if Piwi also plays a role in establishing the chromatin state of satDNAs, we analyzed H3K9me3 ChIP-seq data of embryonic *piwi*

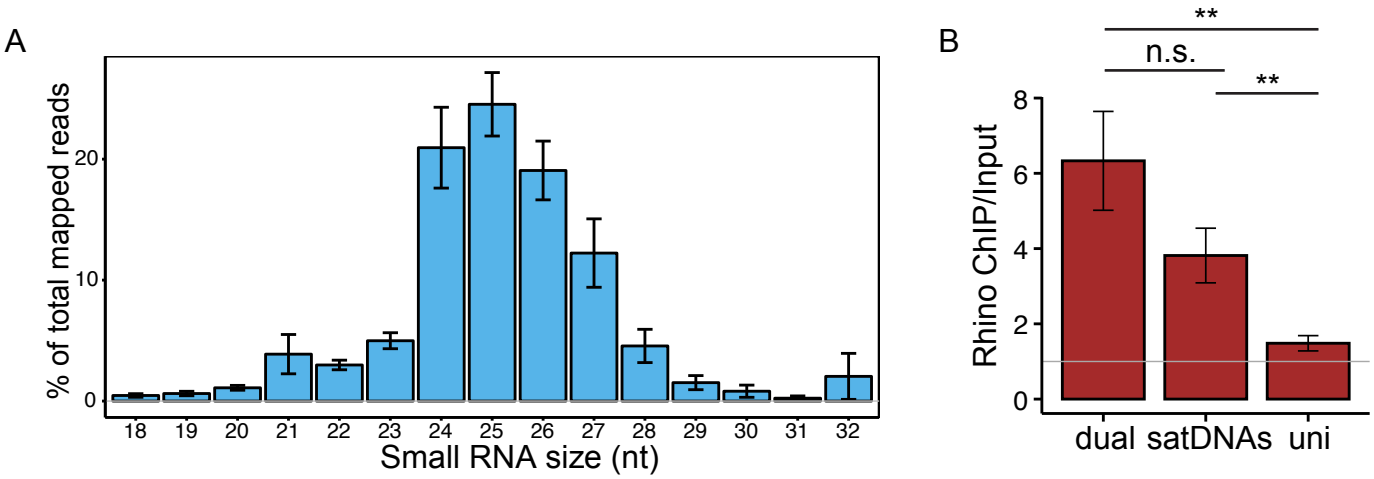


Figure 2. SatDNA produce small RNAs in *D. melanogaster* ovaries. (A) Size distribution of *Rsp* small RNAs in testes and ovaries (*1.688* distribution is in Figure 2-S1C). (B) Rhino ChIP-seq result from ovaries showing the enrichment scores for dual-strand (dual) and uni-strand (uni) piRNA clusters, and satDNAs. The enrichment scores for each satDNA and piRNA cluster are shown in Figure 2-S4A. P-values are estimated by pairwise t-tests with FDR correction (Benjamini 1995). ** adjusted p-value<0.01.

knock down ovaries (Akkouche, et al. 2017). We detected a decrease of H3K9me3 at satDNAs (Figure 2-S5A), suggesting that Piwi is also required for the establishment of heterochromatin at satDNAs. Consistent with the H3K9me3 level decrease, piRNA production from satDNAs is also reduced in embryonic *piwi* knock down ovaries (with some variation among replicates observed for *Rsp*) (Figure 2-S5B) and the total RNA levels for satDNAs are increased (Figure 2-S5C), similar to dual-strand piRNA clusters (Akkouche, et al. 2017). However, it is difficult to distinguish between satDNA precursor and non-precursor RNAs, as it is not clear if satDNA-derived transcripts are spliced.

SatDNA transcription regulation resembles dual-strand piRNA clusters

Because we find that satDNAs generate piRNAs in the female germline and their chromatin is associated with Rhi, we suspect that the same transcription and RNA processing machinery are used by both satDNAs and dual-strand piRNA clusters. To determine if satDNA expression is regulated like dual-strand piRNA clusters, we used publicly available small RNA-seq datasets generated from mutants of genes involved in the heterochromatin-dependent transcription initiation of dual-strand piRNA clusters: Rhi, Cuff, Del (RDC), and Moon (Klattenhoff, et al. 2009; Pane, et al. 2011; Czech, et al. 2013; Le Thomas, et al. 2014; Mohn, et al. 2014; Andersen, et al. 2017; Parhad, et al. 2017). We normalized piRNA abundance to the number of reads mapped to either miRNAs (Figure 3A) or the uni-strand *flamenco* cluster (Figure 3-S1), neither of which should be affected by mutations in the RDC pathway.

Our analysis of the known piRNA clusters agrees with published results. The dual-strand piRNA clusters, *42AB* and *80F,* are Rhi- and Moon-dependent, and that *38C1/2*

is Rhi-dependent but not Moon-dependent. The uni-strand piRNA clusters *20A* and *flamenco* are not dependent on either protein (Klattenhoff, et al. 2009; Pane, et al. 2011; Mohn, et al. 2014; Andersen, et al. 2017). Interestingly, the pools of complex satDNA-derived piRNAs are also reduced in RDC and Moon mutants (Figure 3A; Figure 3-S1). In *rhi* mutants, *Rsp* piRNA abundance is 0.2–6.3% their levels in wild type datasets. Similarly, piRNA abundance for *359-bp* and *260-bp* are 1.5–8.8% and 0.5–6.3% their levels in wild type datasets, respectively (complete list of log2 fold change for satDNAs in Table S6 and Table S7). The reduction in satDNA piRNA level is robust to normalization method (miRNA in Figure 3A; *flamenco* cluster in Figure 3-S1). Similar to complex satDNAs, we found that piRNAs corresponding to the subtelomeric 1872-nt tandem repeat, HETRP, are reduced in the mutants (*e.g*. *rhi* mutants are 0.4-3.1% of the levels in wild type datasets), suggesting that it is also regulated by the RDC complex and Moon (Table S6 and Table S7). While simple satellite repeats like AAGAG did not show decreased expression in these mutants (Table S6 and Table S7), the low abundance of AAGAG reads (the number of reads mapping to AAGAG are only ~0.5% of *Rsp* and ~0.03 of *359-bp*) and known sources of bias for simple repeats (*e.g.,* PCR bias in the RNA-seq library preparation step (Wei, et al. 2018)) points to the need for different approaches to verify this finding. Overall, our results indicate that piRNA production from complex satDNAs is regulated by the heterochromatin-dependent transcription machinery associated with dual-strand piRNA clusters.

To further examine how the RDC complex and Moon affect complex satDNA transcription, we re-analyzed total RNA-seq data of the corresponding mutants (Mohn, et al. 2014; Andersen, et al. 2017). RDC and Moon mutants affect piRNA precursor transcription at dual-strand piRNA clusters *42AB* and *80F* (Mohn, et al. 2014;

Andersen, et al. 2017). We detected similar decreases of steady state long RNA transcript levels at dual-strand piRNA clusters as reported (Figure 3-S2), however, we did not observe a significant decrease in steady state long RNA transcript levels for satDNAs (Table S8). We then tried qRT-PCR using total RNA from ovaries of *rhino* (*rhi-*) and *moonshiner* (*moon*$^{-}$) mutants (Andersen, et al. 2017). After controlling for genomic repeat copy number with qPCR (Figure 3B), *Rsp* expression level is lower, but not significantly so (p-value=0.048 and 0.077; Figure 3C) in both *rhi* and *moon* mutants compared to wildtype. Because complex satDNAs have a generally low expression level (*Rsp*, *260-bp,* and *359-bp* total RNA levels are ~3%, ~3% and ~20%, respectively, of *42AB* and *80F*), we may have insufficient power to detect the decreased level in the mutants or it is possible that only a subset of the satDNA loci have RDC-regulated transcription. Alternatively, the proportion of piRNA precursors-to-non-precursor transcripts in these mutants may have shifted (similar to embryonic *piwi* knockdown ovaries (Akkouche, et al. 2017)), such that the abundance of piRNA precursors decreases, but the total RNA level does not. We suspect that the ratio of precursor to non-precursor transcripts from dual-strand piRNA clusters is higher than satDNAs, given the higher level of piRNA production. If so, a decrease in piRNA precursor levels from satDNAs may be buffered by unchanged levels of non-precursors, making the effects difficult to detect.

We also asked if the satDNA-derived piRNA pool is affected in mutants of twelve genes involved in piRNA precursor export from the nucleus, primary piRNA biogenesis, and the ping-pong cycle (Figure 3-S3) (Czech, et al. 2018) (datasets from (Malone, et al. 2009; Handler, et al. 2011; Olivieri, et al. 2012; Preall, et al. 2012; Zhang, et al. 2012; Czech, et al. 2013; Sato, et al. 2015; Wang, et al. 2015); Table S1). For each of the datasets analyzed, we recapitulate results for all known piRNA clusters reported

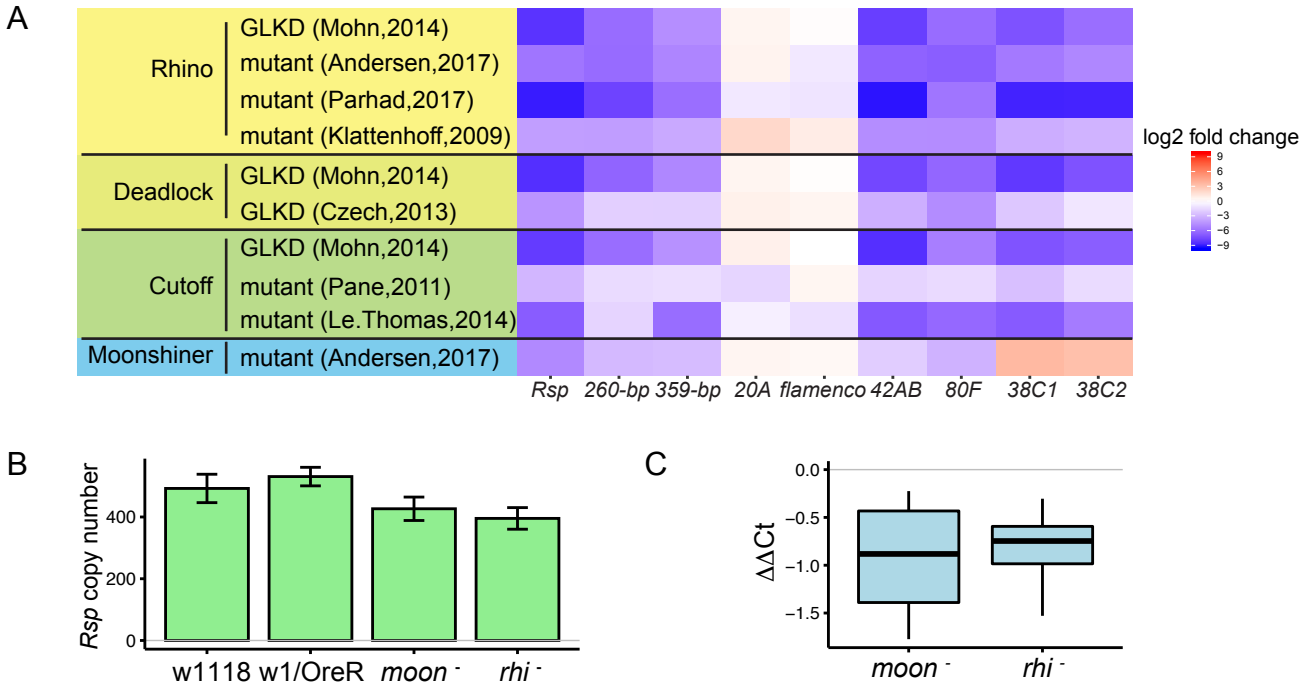


Figure 3. SatDNA loci are regulated by the heterochromatin-dependent transcription machinery in Drosophila ovaries. (A) Heatmap showing the quantification of transcript abundance in small RNA-seq data from mutants of *rhino*, *cutoff*, *deadlock*, and *moonshiner* for satDNAs (*Rsp, 260-bp* and *359-bp* satellite) and piRNA clusters, normalized by miRNA level. GLKD: germline knockdown. (B) qPCR estimate of *Rsp* copy number in wild types and mutants. (C) qRT-PCR estimate of *Rsp* transcript level in mutants compared to wild types. ΔΔCt = ΔCt(wild type) - ΔCt(mutant), a negative value indicates lower expression in mutant. Student's t-test, p-value=0.077, 0.048.

previously (Figure 3-S3) (Czech, et al. 2018). Our reanalysis of these data suggests that piRNA production for all complex satDNA is regulated by the primary piRNA pathway (Gasz, Vreteno, Shutdown), UAP56, and the ping-pong pathway (Ago3, Krimper), but further study is required to determine if additional proteins in each of these pathways are involved. Some of our re-analysis results varied between datasets from different studies for satDNAs. For example, satDNAs show decreased piRNA levels in one mutant Zucchini dataset (Olivieri, et al. 2012), but show increased piRNA levels in two independent Zucchini datasets (Malone, et al. 2009; Handler, et al. 2011). While further work is required to determine all of the components involved in processing satDNA transcripts, our results suggest that piRNA production at satDNA loci is regulated by the heterochromatin-dependent dual-strand piRNA pathway.

CONCLUSIONS

Here we show that complex satDNAs produce piRNAs in male and female *D. melanogaster* germlines. In ovaries, the satDNAs are transcribed into long non-coding RNAs by the heterochromatin-dependent machinery Rhino-Deadlock-Cutoff and Moonshiner and processed into piRNAs. Thus, complex satDNA transcription is regulated in a manner similar to dual-strand piRNA clusters in the female germline (Figure 4). This non-canonical transcription relies on the heterochromatic state of regions destined for piRNA production (Andersen, et al. 2017). The H3K9me3 chromatin at canonical piRNA clusters is established in embryos in a Piwi-dependent manner and is important for subsequent piRNA production (Akkouche, et al. 2017). Similar to piRNA clusters, we find that H3K9me3 chromatin at complex satDNA is reduced when transiently depleting Piwi in the embryos. We therefore hypothesize that satDNA-derived piRNAs are involved in establishing heterochromatin at satDNA

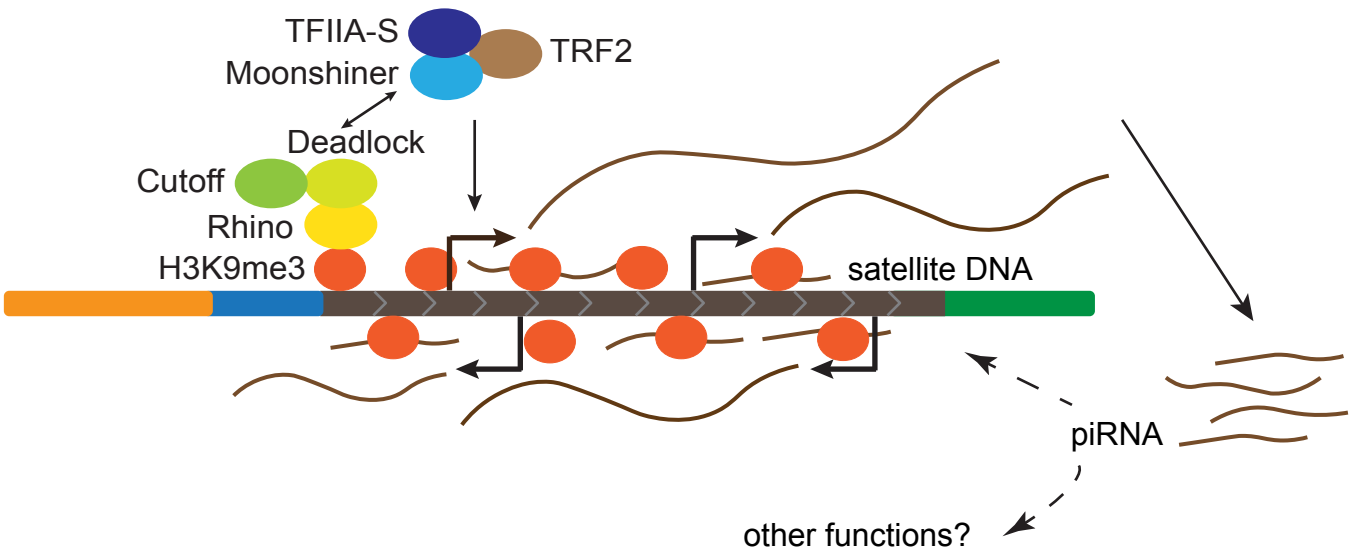


Figure 4. Model for maintenance of satDNA chromatin in female germline. Complex satDNA transcription is regulated by the heterochromatin-dependent Rhino-Deadlock-Cutoff and Moonshiner machinery, and the long RNA transcripts are processed into piRNAs. These piRNAs play roles in the establishment of heterochromatin at their own genomic loci in embryos. This pathway may be important for maintaining genome stability in pericentric heterochromatin, proper nuclear organization and other unexplored functions.

loci in early embryos. Maternally deposited Piwi and Aub may also play roles in defining and establishing piRNA production regions in embryos (Harris and Macdonald 2001; Megosh, et al. 2006; Brennecke, et al. 2008; Le Thomas, et al. 2014). However, the maintenance of the chromatin state at these regions (piRNA clusters) is not dependent on Piwi (Klenov, et al. 2014; Mohn, et al. 2014). Therefore, whether the chromatin state is actively maintained by other heterochromatin factors, or if it is instead stable throughout development into adulthood remains an open question.

piRNA pathways are primarily studied for their conserved role in protecting genome integrity by repressing TE activity in different organisms (Aravin, et al. 2006; Girard, et al. 2006; Grivna, et al. 2006; Lau, et al. 2006; Brennecke, et al. 2007; Houwing, et al. 2007) (reviewed in (Parhad and Theurkauf 2019)). However, these pathways may play a more general role in maintaining genome stability (Aravin, et al. 2003) through heterochromatin formation at pericentromeric satDNAs. The connection between satDNAs and the piRNA pathway has received little attention in previous studies, perhaps because tandem repeats are typically difficult to study. Here we discover that the piRNA pathway indeed affects heterochromatin formation at satDNAs. Our study suggests a simple, self-regulating model, where the piRNA pathway establishes H3K9me3-marked satDNA chromatin at an early stage of development. Establishing heterochromatin at satDNA loci may be a critical step for the larger-scale heterochromatinization of pericentric regions. Thus, our findings support a broader view of the piRNA pathway, not just as a mechanism for repressing potentially harmful TEs, but as one important for properly packing repeat-rich genomic regions, in general.

Disrupting the regulation of these satDNA regions may lead to severe consequences (reviewed in (Janssen, et al. 2018)). For example, the de-repression of satDNA is associated with cellular senescence and various cancers (*e.g.* (Ting, et al. 2011; Zhu,

et al. 2011)), however, untangling the causes, from the effects, of genome instability is difficult. In the ovaries of RDC or Moon mutants, loss of piRNA production at piRNA clusters leads to overexpression of TEs, however the consequences for satDNA are likely to be more complicated. In these mutants, the loss of satDNA-derived piRNA may lead to reduced heterochromatin, which has its own consequences. First, heterochromatic regions form a distinct phase-separated nuclear compartment that contributes to nuclear organization and gene regulation (Larson, et al. 2017; Strom, et al. 2017), and chromocenter formation (Jagannathan, et al. 2018). Un-regulated satDNA may disrupt this organization. Second, de-repressed satDNA may lead to structural rearrangements and genome instability. A loss of H3K9me3 in Drosophila results in aberrant recombination in repetitive genomic regions (Peng and Karpen 2007), presumably through elevated rates of double strand breaks (DSB). The ectopic recombination through DSB repair involving repeats may contribute to chromosomal aberrations associated with cancer (reviewed in (Janssen, et al. 2018)).

Lastly, rearrangements in satDNA regions may contribute to mitotic defects, as heterochromatic satDNAs play important roles in chromosome segregation (*e.g.* (Dernburg, et al. 1996; Bouzinba-Segard, et al. 2006; Rosic, et al. 2014)). Bouts of genome instability caused by the loss of heterochromatin at satDNA may not only contribute to variation between cells within an individual (*e.g.* cancer cells), but over long time scales may also contribute to variation in satDNA organization between individuals and species. SatDNAs are among the most dynamic and rapidly evolving regions of the genome, with species having their own species-specific profiles of satellite repeats (reviewed in (Ferree and Prasad 2012; Plohl, et al. 2012)). The Drosophila complex satDNAs that we study here are highly dynamic, with their location varying dramatically between species (Larracuente 2014; Jagannathan, et al. 2017;

Sproul, et al. 2020). The rapid turnover of satDNAs can have consequences for speciation—the *1.688* satellite is associated with genetic incompatibilities between species that results in mitotic defects and embryonic lethality (Ferree and Barbash 2009) for reasons that we still do not understand.

Genes in the piRNA pathway evolve rapidly and under positive selection (*e.g.* (Vermaak, et al. 2005; Obbard, et al. 2009; Lee and Langley 2012; Simkin, et al. 2013; Yi, et al. 2014)). This is typically discussed as a consequence of arms races with TEs (reviewed in (Parhad and Theurkauf 2019)), where TEs evolve rapidly to counteract or escape detection by the piRNA pathway, and the piRNA pathway evolves rapidly to keep up with TEs. However, there are few examples of specific TEs counteracting or evading the piRNA pathway ((Blumenstiel, et al. 2016; Cosby, et al. 2019); but see an interesting potential case in Drosophila TART non-LTR retroelements (Ellison, et al. 2020)). The piRNA pathway also affects genes and viruses, and these “off target” effects may contribute to the rapid evolution of the pathway (i.e. the autoimmunity hypothesis) (Blumenstiel, et al. 2016). An additional clue might come from the male germline, where proportionally fewer piRNAs are derived from TEs than in the female germline (Nishida, et al. 2007; Nagao, et al. 2010; Quenerch'du, et al. 2016). In testes, small RNAs derived from non-TE and non-satellite repeats (Quenerch'du, et al. 2016) may play roles in mitigating intragenomic conflicts over transmission through the germline (e.g. meiotic drive systems) (Meiklejohn and Tao 2010; Courret, et al. 2019). The clearest example of this conflict is in *D. simulans,* where endo-siRNAs derived from a duplication of an X-linked meiotic driver suppress drive in males (Tao, et al. 2007; Lin, et al. 2018). Similarly, the piRNA pathway is involved in intragenomic conflicts—in *D. melanogaster,* piRNAs derived from the tandemly repeated Y-linked *Suppressor of Stellate* (*Su(Ste)*) suppress X-linked *Stellate* transcripts (Aravin, et al.

2001; Aravin, et al. 2004; Kotelnikov, et al. 2009; Nagao, et al. 2010), which would otherwise result in male sterility (Palumbo, et al. 1994; Hurst 1996; Belloni, et al. 2002). The piRNA pathway is also implicated in a meiotic drive system that targets satDNA (*Rsp*) in the male germline (*Segregation Distorter*; (Meiklejohn and Tao 2010; Larracuente and Presgraves 2012; Gell and Reenan 2013)). An emerging theme in studies of meiotic drive systems in Drosophila species is that there is a link between meiotic drive, piRNAs, and heterochromatin (Courret, et al. 2019), with satDNAs often at the center of these conflicts.

Many mysteries remain surrounding the roles of the piRNA pathway and factors shaping its evolution, which may be in part driven by its functions outside of TE control. Our finding that satDNAs are regulated by the piRNA pathway highlight a potentially important factor that may shed light on some of these mysteries. Moving forward, it will be important for piRNA studies to continue to focus on satDNA and how these dynamic compartments of the genome contribute to genome function and stability.

MATERIALS AND METHODS

Drosophila stock

Iso-1 was used as the wild type strain, unless stated otherwise. In the qPCR validation experiment, *rhi* mutants (*rhi*$^{-}$) are transheterozygotes from the Vienna Drosophila Resource Center (VDRC 313487 and 313488) as are the *moonshiner* mutants (*moon*$^{-}$) (VDRC 313735 and 313738) as described in (Andersen, et al. 2017). Based on the origin and genetic background of these mutants, w^{1118} or the progeny from OregonR (Ore) crossed to w^{1} were used as the wild type controls for *rhi* $^{-}$ and *moon*$^{-}$. All flies were maintained at 23ºC on cornmeal medium.

Small RNA-seq

Six- to eight-day-old testes were dissected in RNase free PBS buffer. Total RNA was extracted using *mir*Vana™ miRNA Isolation Kit (Ambion) with procedures for isolating RNA fractions enriched for small RNAs (<200nt), then treated with RNase free DNase I (Promega) at 37ºC for 1 hour. Library preparation and sequencing were performed by Genomics Research Center at University of Rochester. Briefly, 2S rRNA was depleted (Wickersheim and Blumenstiel 2013), small RNA library was prepared with TruSeq Small RNA Library Prep Kit (Illumina) and sequenced by Illumina platform HiSeq2500 Single-end 50 bp.

Total RNA-seq

Six- to eight-day-old testes were dissected in RNase free PBS buffer. Total RNA was extracted using *mir*Vana™ miRNA Isolation Kit (Ambion) with procedures for isolating RNA fractions enriched for long RNAs (>200nt), then treated with RNase free DNase I (Promega) at 37ºC for 1 hour. Library preparation and sequencing were performed by Genomics Research Center at University of Rochester. Briefly, rRNA was removed

and total RNA library was prepared with TruSeq Stranded Total RNA Library Prep Human/Mouse/Rat (Illumina) and sequenced by Illumina platform HiSeq2500 Paired-end 125 bp.

SatDNA analysis

Reads were mapped to the heterochromatin-enriched genome assembly (Chang and Larracuente 2019) and counted based on their annotations (*e.g. Rsp* or *1.688*). Due to the highly repetitive nature of satDNAs, around 80% of total RNA-seq and 99% of small RNA-seq reads that are mapped to satDNA regions are not uniquely assigned; discarding these multiple mapped reads would result in loss of statistical power in the satDNA analysis. To deal with this, multiple mapped reads were randomly assigned to one of their multiple best mapping locations. Reads were then counted based on the annotations of their assigned mapping locations. A similar approach was used in our analysis of piRNA clusters, except that only uniquely mapped reads were counted so that the published results could serve as controls for our method. Additional details specific to small RNA-seq, RNA-seq, ChIP-seq, and RIP-seq analyses are given below.

RNA-seq analysis

All total RNA-seq datasets reanalyzed in our study are listed in Table S1. Total RNA-seq reads were trimmed for adaptors and then mapped to the genome using Bowtie2 (Langmead and Salzberg 2012). A customized python script was used to count reads that mapped to each repeat feature or piRNA cluster, and RPM values were reported by normalizing raw counts to 1,000,000 total mapped reads (https://github.com/LarracuenteLab/Dmelanogaster_satDNA_regulation (Wei 2020), htseq_bam_count_proportional.py). For results shown in Table S8, DESeq2 (Love, et

al. 2014) was used to perform differential expression analysis of the raw counts with combined data from different studies (Mohn, et al. 2014; Andersen, et al. 2017), with experimental condition and associated study as covariates. This analysis method is conservative and leads to smaller log2 fold changes than published results of piRNA clusters. For comparison with the published results, a similar approach was used to analyze piRNA clusters (Mohn, et al. 2014; Andersen, et al. 2017). Briefly, quantification of reads mapping to 1-kb windows inside each piRNA cluster was estimated using a customized python script (https://github.com/LarracuenteLab/Dmelanogaster_satDNA_regulation (Wei 2020), htseq_bam_count_proportional.py), and subsequent differential expression analysis between mutants and wildtype was done using DESeq2 (Love, et al. 2014) (results shown in Figure 3-S2).

Small RNA-seq analysis

All small RNA-seq datasets reanalyzed in our study are listed in Table S1. Small RNA-seq reads were trimmed for adaptors, then mapped to the genome using Bowtie (Langmead 2010). A customized python script (https://github.com/LarracuenteLab/Dmelanogaster_satDNA_regulation (Wei 2020), htseq_bam_count_proportional.py) was used to count reads that mapped to each repeat feature or piRNA cluster. To control for differences in small RNA abundance and compare across samples, raw counts were then normalized to the number of reads that mapped to either miRNAs or the *flamenco* piRNA cluster. The difference in expression was represented by the log2 fold changes of these normalized counts in mutants compared to wild type (*i.e.*, $\log2(\text{count}_{\text{mutant}}/\text{count}_{\text{WT}})$) for each repeat and piRNA cluster. We further analyzed the size distribution and relative nucleotide bias at positions along each satDNA by extracting reads mapped to the satDNA of interest

using a customized python script (https://github.com/LarracuenteLab/Dmelanogaster_satDNA_regulation (Wei 2020), extract_sequence_by_feature_gff.py). The 10nt overlap Z-score of piRNAs mapped to each satDNA was calculated using piPipes (Han, et al. 2015). To determine which parts of these repeats are represented in piRNA or ChIP datasets, the read pileup patterns along the consensus sequence of a satDNA were examined (*e.g.* Figure 2-S2). We BLAST-ed either reads (ChIP or piRNA) mapping to a particular satDNA or genomic satDNA variants (as a control) to the consensus dimer (for *1.688* satellite) or trimer (for *Rsp* because it has left and right consensus sequences) of this satDNA, and then convert coordinates along a dimer/trimer to coordinates along a monomer/dimer consensus sequence. All plots were made in R (R core team 2017).

ChIP/RIP-seq analysis

All total ChIP-seq and RIP-seq datasets reanalyzed in our study are listed in Table S1. ChIP-seq and RIP-seq reads were trimmed for adaptors and mapped to genome using Bowtie2 (Langmead and Salzberg 2012). A customized python script (https://github.com/LarracuenteLab/Dmelanogaster_satDNA_regulation (Wei 2020), htseq_bam_count_proportional.py) was used to count reads that mapped to each repeat feature or piRNA cluster. Raw counts were normalized to 1,000,000 total mapped reads. For the ChIP-seq results, enrichment scores of each repeat and piRNA cluster were reported by comparing the ChIP sample with the antibody of interest to its no-antibody input control sample. For the RIP-seq result, we report the percentage of reads mapped to each repeat and piRNA cluster with miRNAs as the negative control.

RNA FISH

A Cy5-labeled oligo probe (5'-Cy5TTTTCCAAATTTCGGTCATCAAATAATCAT-3') previously described in (Ferree and Barbash 2009) was used to detect *1.688* transcripts. Custom Stellaris FISH Probes were designed for *Rsp* (Table S9) and RNA FISH was performed following the manufacturer's instructions (Biosearch Technologies, Inc). Three- to six-day-old ovaries and testes were dissected in RNase free PBS buffer, fixed with 4% paraformaldehyde in PBS buffer at room temperature for 30 minutes, and then washed twice with PBS for 5 minutes. To permeabilize, tissues were kept in RNase free 70% ethanol at 4ºC overnight. The ethanol was aspirated, and samples washed with Stellaris wash buffer on a nutating mixer for 3 minutes and kept still for 2 minutes at room temperature. Hybridization was then performed with each probe in Stellaris hybridization buffer in a humidity chamber at 37ºC overnight. The working concentration was 100nM for the oligo probe and 125nM for the Stellaris probes. From this point, samples were kept in the dark. The samples were washed with Stellaris wash buffer twice at 37ºC for 30 minutes each. Samples were then transferred to mounting medium containing DAPI and imaged with Leica SP5 laser scanning confocal microscope.

For RNaseA controls, after fixation and permeation, tissues were treated with RNase A (100ug/ml) in RNase digestion buffer (5 mM EDTA, 300 mM NaCl, 10 mM Tris-HCl pH 7.5, cold spring harbor protocols (http://cshprotocols.cshlp.org/content/2013/3/pdb.rec074146.full)) at 37ºC for 1hr and washed three times with Stellaris wash buffer at room temperature for 10 minutes before hybridization.

For RNase H controls, after probe hybridization and washing, tissues were treated with 1.5uL RNase H (5,000 units/ml) (New England Biolabs) in 50uL final volume in

1X RNAse H buffer at 37ºC for 2 hours and washed three times with Stellaris wash buffer at room temperature for 10 minutes before mounting and imaging.

qPCR

For genomic DNA qPCR, three- to eight-day-old flies were mashed with pipette tips for 5-10 seconds and incubated in buffer (10 mM Tris-Cl pH 8.2, 1 mM EDTA, 25 mM NaCl, 200 ug/ml Proteinase K) at 37ºC for 30 minutes (Gloor and Engels 1992). To extract nucleic acids, a mixture of phenol/Sevag (1:1) of equal volume was added, and the samples vortexed for 45 to 60 seconds and then centrifuged for 3-5 minutes. The aqueous top layers were saved, an equal volume of Sevag added, and the samples vortexed for 30 seconds then centrifuged for 1 minute. The aqueous top layers were saved and a second Sevag extraction performed. Diluted nucleic acid samples (concentration of 0.04ng/uL) were used for qPCR to determine the repeat copy numbers in the genome. Repeat copy numbers are normalized to the tRNA:Lys-CTT copy numbers.

For RNA qRT-PCR, three- to six-day-old ovaries were dissected in RNase free PBS buffer, and total RNA was extracted using the *mir*Vana™ miRNA Isolation Kit (Ambion). RNA samples were treated with RNase free DNase I (Promega) at 37ºC for 1 hour. The RNA samples were reverse transcribed using random hexamer primers and M-MuLV Reverse Transcriptase (New England Biolabs) and the resulting cDNA subjected to qPCR. To exclude the possibility of DNA signal in qRT-PCR experiments, controls with no Reverse Transcriptase enzyme were used for all samples in the reverse transcription step. Expression levels were normalized to ribosomal protein S3 (RPS3) expression. To detect the transcript abundance difference between wild type and mutant, ΔΔCT was calculated (Livak and Schmittgen 2001).

The sequences of primers used are: *Rsp* (forward: GGAAAATCACCCATTTTGATCGC, reverse: CCGAATTCAAGTACCAGAC); tRNA (forward: CTAGCTCAGTCGGTAGAGCATGA, reverse: CCAACGTGGGGCTCGAAC); RPS3: (forward: AGTTGTACGCCGAGAAGGTG, reverse: TGTAGCGGAGCACACCATAG).

Northern Blot Analysis

Isolation of total RNA and RNase controls: stocks of *D. melanogaster* were chosen which represented a range of *Rsp* repeat copy numbers; flies were collected (0 to 20 hours old) and aged for six days. Ovaries were dissected from approximately 20 females (i.e., 6.0 to 6.8 days old) from each stock, and total nucleic acid isolated using a standard phenol/chloroform procedure (Khost, et al. 2017). Total nucleic acid was then treated with DNase I as recommended (20 units; Promega), re-extracted with phenol/chloroform, and ethanol precipitated. Total RNA was resuspended in distilled water. The integrity of the RNA was checked on 1% agarose gels, and the concentration estimated by an optical density at 260 nm.

For RNase controls, 10 µg of total RNA was resuspended in 50 mM NaCl, 5 mM EDTA, 10 mM Tris pH 7.5, 100 µg/ml RNaseA and incubated at 37°C for 30 minutes. Samples were phenol/chloroform extracted, 10 µg of ytRNA added as carrier, and ethanol precipitated.

Northern Blot Analysis: Total RNA (10 µg)/ RNase controls were suspended in 1×MOPS [0.04 M MOPS (morpholinepropanesulfonic acid) pH 7.0, 0.01 M Na acetate, 0.001 M EDTA], 2.2 M formaldehyde, 50% formamide. The RNA was then heated at 65°C for 15 min, placed on ice, and one-tenth volume loading buffer (1× MOPS, 50% formamide, 2.2 M formaldehyde, 4% Ficoll, 0.25% bromophenol blue) added. RNAs

were separated on a 1% agarose gel containing 0.5 M formaldehyde/1× MOPS at 40 V for 3 h. Standard RNA lanes were cut from the gel and stained with ethidium bromide to monitor electrophoresis. Gels were washed for 25 minutes in sterile water (with 4 changes). RNA was transferred to GeneScreen Plus nylon membrane (prewet in 10× SSC) by capillary action using 10× SSC. After transfer, the nylon membrane was rinsed in 2× SSC, UV crosslinked, and then baked for 2 h under vacuum at 80°C. The membrane was prehybridized in 2× SSC, 5× Denhardt's solution, 1% sodium dodecyl sulfate (SDS), 10% polyethylene glycol (PEG- molecular weight, 8,000), 25 mM sodium phosphate (pH 7.2), 0.1% sodium pyrophosphate, and 50% formamide for 3 h at 55°C. Hybridizations were done overnight at 55°C in the same buffer containing a biotinylated RNA probe (see slot blot; primers: T7_rsp2 5′-TAATACGACTCACTATAGGGCCGAATTCAAGTACCAGAC-3′ and rsp1 5′-GGAAAATCACCCATTTTGATCGC-3′). The hybridized membranes were washed in 1M sodium phosphate pH 6.8, 0.5M EDTA, 5% SDS (2X, 10 minutes each) at 60°C and then at 1M sodium phosphate pH 6.8, 0.5M EDTA, 1% SDS (3X, 10 minutes each) at 65°C. The washed membranes were then processed as recommended for the Chemiluminescent Nucleic Acid Detection Module (ThermoScientific), and the signal recorded on a ChemiDoc XR+ (Bio-Rad).

Slot blot

Five female flies were mashed and the total nucleic acid phenol/Sevag extracted as described above for qPCR. Approximately 200 ng of the nucleic acid was denatured (final concentration 0.25 M NaOH, 0.5 M NaCl) for 10 minutes at room temperature, the sample transferred to a tube with an equal volume of ice-cold loading buffer (0.1X SSC, 0.125 M NaOH) and left on ice. The slot blotter was then prepared and samples loaded as recommended for the 48-well BioDot SF microfiltration apparatus (Bio-Rad).

After loading, the wells were washed with 200 µl of loading buffer. The nylon membrane (GeneScreen Plus) was then rinsed for 2 minutes with 2X SSC before being UV crosslinked (Stratalinker). The membrane was first hybridized with a biotinylated rp49 RNA probe in North2South hybridization solution (ThermoScientific) at 65°C overnight. The membrane was processed as recommended for the Chemiluminescent Nucleic Acid Detection Module (ThermoScientific), and the signal recorded on a ChemiDoc XR+ (Bio-Rad). The membrane was then stripped with a 100°C solution of 0.1X SSC/ 0.5% SDS (3 times for ~20 minutes each) and re-hybridized with a *Rsp* probe (60°C overnight) and processed as above. Signals were quantitated using the ImageLab software (Bio-Rad).

To make the biotinylated RNA probes, gel extracted PCR amplicons (primers: *Rsp* 5'-TAATACGACTCACTATAGGGGAAAATCACCCATTTTGATCGC-3' and 5'-CCGAATTCAAGTACCAGAC-3'; rp49 5'-GTAATACGACTCACTATAGGGCAGTAAACGCGGTTCTGCATG-3' and 5'-CAGCATACAGGCCCAAGATC-3') were transcribed using the Biotin RNA Labeling Mix (Roche) and T7 polymerase (Promega).

Acknowledgements

This work was supported by the National Institutes of Health General Medical Sciences (R35 GM119515 to AML), a Steven Biggar and Elisabeth Asaro fellowship in Data Science to AML, and a University of Rochester Messersmith Fellowship to XW. We thank Drs. Ching-Ho Chang, John Sproul, Cécile Courret, and Lucas Hemmer for providing feedback on the manuscript. We also thank the University of Rochester Center for Integrated Research Computing for access to computing facilities and the University of Rochester Genomics Research Center for sequencing services.

Data availability

Sequencing data generated in this paper are available in the NCBI Sequence Read Archive under project accession PRJNA647441. All data files and code to recreate analyses and figures, and supplemental data files are deposited in GitHub (https://github.com/LarracuenteLab/Dmelanogaster_satDNA_regulation)(Wei 2020) and at the Dryad Digital Repository (link forthcoming).

Competing interests: The authors acknowledge that they have no competing interests

References

Abad JP, Agudo M, Molina I, Losada A, Ripoll P, Villasante A. 2000. Pericentromeric regions containing 1.688 satellite DNA sequences show anti-kinetochore antibody staining in prometaphase chromosomes of *Drosophila melanogaster*. Mol Gen Genet 264:371-377.

Akkouche A, Mugat B, Barckmann B, Varela-Chavez C, Li B, Raffel R, Pelisson A, Chambeyron S. 2017. Piwi Is Required during Drosophila Embryogenesis to License Dual-Strand piRNA Clusters for Transposon Repression in Adult Ovaries. Mol Cell 66:411-419 e414.

Allshire RC, Madhani HD. 2018. Ten principles of heterochromatin formation and function. Nat Rev Mol Cell Biol 19:229-244.

Andersen PR, Tirian L, Vunjak M, Brennecke J. 2017. A heterochromatin-dependent transcription machinery drives piRNA expression. Nature 549:54-59.

Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, 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 AA, Klenov MS, Vagin VV, Bantignies F, Cavalli G, Gvozdev VA. 2004. Dissection of a natural RNA silencing process in the *Drosophila melanogaster* germ line. Molecular and cellular biology 24:6742-6750.

Aravin AA, Lagos-Quintana M, Yalcin A, Zavolan M, Marks D, Snyder B, Gaasterland T, Meyer J, Tuschl T. 2003. The Small RNA Profile during *Drosophila melanogaster* Development. Developmental Cell 5:337-350.

Aravin AA, Naumova NM, Tulin AV, Vagin VV, Rozovsky YM, Gvozdev VA. 2001. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. Current Biology 11:1017-1027.

Belloni M, Tritto P, Bozzetti MP, Palumbo G, Robbins LG. 2002. Does Stellate cause meiotic drive in *Drosophila melanogaster*? Genetics 161:1551-1559.

Benjamini Y, Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 57:289-300.

Biscotti MA, Canapa A, Forconi M, Olmo E, Barucca M. 2015. Transcription of tandemly repetitive DNA: functional roles. Chromosome Res 23:463-477.

Blumenstiel JP, Erwin AA, Hemmer LW. 2016. What Drives Positive Selection in the Drosophila piRNA Machinery? The Genomic Autoimmunity Hypothesis. Yale J Biol Med 89:499-512.

Bouzinba-Segard H, Guais A, Francastel C. 2006. Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. Proc Natl Acad Sci U S A 103:8709-8714.

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128:1089-1103.

Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. 2008. An epigenetic role for maternally inherited piRNAs in transposon silencing. Science 322:1387-1392.

Britten RJ, Kohne DE. 1968. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. Science 161:529-540.

Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, Booth BW, Wen J, Park S, Suzuki AM, et al. 2014. Diversity and dynamics of the Drosophila transcriptome. Nature 512:393-399.

Chang CH, Larracuente AM. 2019. Heterochromatin-Enriched Assemblies Reveal the Sequence and Organization of the *Drosophila melanogaster* Y Chromosome. Genetics 211:333-348.

Charlesworth B, Sniegowski P, Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371:215-220.

Cosby RL, Chang NC, Feschotte C. 2019. Host-transposon interactions: conflict, cooperation, and cooption. Genes Dev 33:1098-1116.

Courret C, Chang CH, Wei KH, Montchamp-Moreau C, Larracuente AM. 2019. Meiotic drive mechanisms: lessons from Drosophila. Proc Biol Sci 286:20191430.

Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, Perrimon N, Kellis M, Wohlschlegel JA, Sachidanandam R, et al. 2008. An endogenous small interfering RNA pathway in Drosophila. Nature 453:798-802.

Czech B, Munafo M, Ciabrelli F, Eastwood EL, Fabry MH, Kneuss E, Hannon GJ. 2018. piRNA-Guided Genome Defense: From Biogenesis to Silencing. Annu Rev Genet 52:131-157.

Czech B, Preall JB, McGinn J, Hannon GJ. 2013. A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway. Mol Cell 50:749-761.

Dernburg AF, Sedat JW, Hawley RS. 1996. Direct evidence of a role for heterochromatin in meiotic chromosome segregation. Cell 86:135-146.

Doolittle WF, Sapienza C. 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature 284:601-603.

Ellison CE, Kagda MS, Cao W. 2020. The Drosophila TART transposon manipulates the piRNA pathway as a counter-defense strategy to limit host silencing. bioRxiv:https://doi.org/10.1101/2020.1102.1120.957324.

Fagegaltier D, Konig A, Gordon A, Lai EC, Gingeras TR, Hannon GJ, Shcherbata HR. 2014. A genome-wide survey of sexually dimorphic expression of Drosophila miRNAs identifies the steroid hormone-induced miRNA let-7 as a regulator of sexual identity. Genetics 198:647-668.

Ferree PM, Barbash DA. 2009. Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in Drosophila. PLoS biology 7:e1000234.

Ferree PM, Prasad S. 2012. How can satellite DNA divergence cause reproductive isolation? Let us count the chromosomal ways. Genet Res Int 2012:430136.

Ferreira D, Meles S, Escudeiro A, Mendes-da-Silva A, Adega F, Chaves R. 2015. Satellite non-coding RNAs: the emerging players in cells, cellular pathways and cancer. Chromosome Res 23:479-493.

Gell SL, Reenan RA. 2013. Mutations to the piRNA pathway component aubergine enhance meiotic drive of segregation distorter in *Drosophila melanogaster.* Genetics 193:771-784.

Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, Lee S, Xu J, Kittler EL, Zapp ML, Weng Z, et al. 2008. Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science 320:1077-1081.

Ghildiyal M, Xu J, Seitz H, Weng Z, Zamore PD. 2010. Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway. RNA 16:43-56.

Girard A, Sachidanandam R, Hannon GJ, Carmell MA. 2006. A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature 442:199-202.

Gloor GB, Engels WR. 1992. Single Fly Preps for PCR. Drosophila Information Services 71:148-149.

Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, et al. 2011. The developmental transcriptome of *Drosophila melanogaster.* Nature 471:473-479.

Grivna ST, Beyret E, Wang Z, Lin H. 2006. A novel class of small RNAs in mouse spermatogenic cells. Genes Dev 20:1709-1714.

Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC. 2007. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. Science 315:1587-1590.

Halbach R, Miesen P, Joosten J, Taskopru E, Rondeel I, Pennings B, Vogels CBF, Merkling SH, Koenraadt CJ, Lambrechts L, et al. 2020. A satellite repeat-derived piRNA controls embryonic development of Aedes. Nature 580:274-277.

Han BW, Wang W, Zamore PD, Weng Z. 2015. piPipes: a set of pipelines for piRNA and transposon analysis via small RNA-seq, RNA-seq, degradome- and CAGE-seq, ChIP-seq and genomic DNA sequencing. Bioinformatics 31:593-595.

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

Harris AN, Macdonald PM. 2001. Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C. Development 128:2823-2832.

Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB, et al. 2007. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell 129:69-82.

Hurst LD. 1996. Further evidence consistent with Stellate's involvement in meiotic drive. Genetics 142:641-643.

Hutvagner G, Simard MJ. 2008. Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Biol 9:22-32.

Jagannathan M, Cummings R, Yamashita YM. 2018. A conserved function for pericentromeric satellite DNA. Elife 7.

Jagannathan M, Warsinger-Pepe N, Watase GJ, Yamashita YM. 2017. Comparative Analysis of Satellite DNA in the *Drosophila melanogaster* Species Complex. G3 (Bethesda) 7:693-704.

Janssen A, Colmenares SU, Karpen GH. 2018. Heterochromatin: Guardian of the Genome. Annu Rev Cell Dev Biol 34:265-288.

Khost DE, Eickbush DG, Larracuente AM. 2017. Single-molecule sequencing resolves the detailed structure of complex satellite DNA loci in *Drosophila melanogaster*. Genome Res 27:709-721.

Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A, et al. 2009. The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell 138:1137-1149.

Klenov MS, Lavrov SA, Korbut AP, Stolyarenko AD, Yakushev EY, Reuter M, Pillai RS, Gvozdev VA. 2014. Impact of nuclear Piwi elimination on chromatin state in *Drosophila melanogaster* ovaries. Nucleic Acids Res 42:6208-6218.

Kotelnikov RN, Klenov MS, Rozovsky YM, Olenina LV, Kibanov MV, Gvozdev VA. 2009. Peculiarities of piRNA-mediated post-transcriptional silencing of Stellate repeats in testes of *Drosophila melanogaster.* Nucleic acids research 37:3254-3263.

Langmead B. 2010. Aligning short sequencing reads with Bowtie. Curr Protoc Bioinformatics Chapter 11:Unit 11 17.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nature methods 9:357-359.

Larracuente AM. 2014. The organization and evolution of the Responder satellite in species of the *Drosophila melanogaster* group: dynamic evolution of a target of meiotic drive. BMC Evol Biol 14:233.

Larracuente AM, Presgraves DC. 2012. The Selfish Segregation Distorter Gene Complex of *Drosophila melanogaster.* Genetics 192:33-53.

Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, Agard DA, Redding S, Narlikar GJ. 2017. Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. Nature 547:236-240.

Lau NC, Seto AG, Kim J, Kuramochi-Miyagawa S, Nakano T, Bartel DP, Kingston RE. 2006. Characterization of the piRNA complex from rat testes. Science 313:363-367.

Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Toth KF. 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev 27:390-399.

Le Thomas A, Stuwe E, Li S, Du J, Marinov G, Rozhkov N, Chen YC, Luo Y, Sachidanandam R, Toth KF, et al. 2014. Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. Genes Dev 28:1667-1680.

Lee YC, Langley CH. 2012. Long-term and short-term evolutionary impacts of transposable elements on Drosophila. Genetics 192:1411-1432.

Lin CJ, Hu F, Dubruille R, Vedanayagam J, Wen J, Smibert P, Loppin B, Lai EC. 2018. The hpRNA/RNAi Pathway Is Essential to Resolve Intragenomic Conflict in the Drosophila Male Germline. Dev Cell 46:316-326 e315.

Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD, et al. 2004. Role of transposable elements in heterochromatin and epigenetic control. Nature 430:471-476.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402-408.

Lohe AR, Roberts PA. 1988. Evolution of satellite DNA sequences in *Drosophila*. In: Verma RS, editor. Heterochromatin: Molecular and Structural Aspects. Cambridge, UK: Cambridge University Press.

Losada A, Villasante A. 1996. Autosomal location of a new subtype of 1.688 satellite DNA of Drosophila melanogaster. Chromosome Res 4:372-383.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.

Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ. 2009. Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell 137:522-535.

Megosh HB, Cox DN, Campbell C, Lin H. 2006. The role of PIWI and the miRNA machinery in Drosophila germline determination. Curr Biol 16:1884-1894.

Meiklejohn CD, Tao Y. 2010. Genetic conflict and sex chromosome evolution. Trends in ecology & evolution 25:215-223.

Mills WK, Lee YCG, Kochendoerfer AM, Dunleavy EM, Karpen GH. 2019. RNA from a simple-tandem repeat is required for sperm maturation and male fertility in Drosophila melanogaster. Elife 8.

Mohn F, Handler D, Brennecke J. 2015. piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. Science 348:812-817.

Mohn F, Sienski G, Handler D, Brennecke J. 2014. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila. Cell 157:1364-1379.

Nagao A, Mituyama T, Huang H, Chen D, Siomi MC, Siomi H. 2010. Biogenesis pathways of piRNAs loaded onto AGO3 in the Drosophila testis. RNA 16:2503-2515.

Nishida KM, Saito K, Mori T, Kawamura Y, Nagami-Okada T, Inagaki S, Siomi H, Siomi MC. 2007. Gene silencing mechanisms mediated by Aubergine piRNA complexes in Drosophila male gonad. RNA 13:1911-1922.

Obbard DJ, Gordon KH, Buck AH, Jiggins FM. 2009. The evolution of RNAi as a defence against viruses and transposable elements. Philos Trans R Soc Lond B Biol Sci 364:99-115.

Okamura K, Chung WJ, Ruby JG, Guo H, Bartel DP, Lai EC. 2008. The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. Nature 453:803-806.

Olivieri D, Senti KA, Subramanian S, Sachidanandam R, Brennecke J. 2012. The cochaperone shutdown defines a group of biogenesis factors essential for all piRNA populations in Drosophila. Mol Cell 47:954-969.

Orgel LE, Crick FH. 1980. Selfish DNA: the ultimate parasite. Nature 284:604-607.

Palumbo G, Bonaccorsi S, Robbins LG, Pimpinelli S. 1994. Genetic analysis of Stellate elements of *Drosophila melanogaster*. Genetics 138:1181-1197.

Pane A, Jiang P, Zhao DY, Singh M, Schupbach T. 2011. The Cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline. EMBO J 30:4601-4615.

Parhad SS, Theurkauf WE. 2019. Rapid evolution and conserved function of the piRNA pathway. Open Biol 9:180181.

Parhad SS, Tu S, Weng Z, Theurkauf WE. 2017. Adaptive Evolution Leads to Cross-Species Incompatibility in the piRNA Transposon Silencing Machinery. Dev Cell 43:60-70 e65.

Pathak RU, Mamillapalli A, Rangaraj N, Kumar RP, Vasanthi D, Mishra K, Mishra RK. 2013. AAGAG repeat RNA is an essential component of nuclear matrix in Drosophila. RNA Biol 10:564-571.

Peng JC, Karpen GH. 2007. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. Nat Cell Biol 9:25-35.

Pimpinelli S, Dimitri P. 1989. Cytogenetic analysis of segregation distortion in *Drosophila melanogaster*: the cytological organization of the *Responder* (*Rsp*) locus. Genetics 121:765-772.

Plohl M, Mestrovic N, Mravinac B. 2014. Centromere identity from the DNA point of view. Chromosoma 123:313-325.

Plohl M, Mestrovic N, Mravinac B. 2012. Satellite DNA evolution. Genome Dyn 7:126-152.

Preall JB, Czech B, Guzzardo PM, Muerdter F, Hannon GJ. 2012. shutdown is a component of the Drosophila piRNA biogenesis machinery. RNA 18:1446-1457.

Quenerch'du E, Anand A, Kai T. 2016. The piRNA pathway is developmentally regulated during spermatogenesis in Drosophila. RNA 22:1044-1054.

R core team. 2017. R: A language and environment for statistical computing: R Foundation for Statistical Computing, Vienna, Austria

Rangan P, Malone CD, Navarro C, Newbold SP, Hayes PS, Sachidanandam R, Hannon GJ, Lehmann R. 2011. piRNA Production Requires Heterochromatin Formation in Drosophila. Current biology : CB 21:1373-1379.

Rosic S, Kohler F, Erhardt S. 2014. Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. J Cell Biol.

Rozhkov NV, Aravin AA, Zelentsova ES, Schostak NG, Sachidanandam R, McCombie WR, Hannon GJ, Evgen'ev MB. 2010. Small RNA-based silencing strategies for transposons in the process of invading Drosophila species. RNA 16:1634-1645.

Rozhkov NV, Hammell M, Hannon GJ. 2013. Multiple roles for Piwi in silencing Drosophila transposons. Genes Dev 27:400-412.

Sandler L, Hiraizumi Y, Sandler I. 1959. Meiotic Drive in Natural Populations of *Drosophila Melanogaster*. I. the Cytogenetic Basis of Segregation-Distortion. Genetics 44:233-250.

Sato K, Iwasaki YW, Shibuya A, Carninci P, Tsuchizawa Y, Ishizu H, Siomi MC, Siomi H. 2015. Krimper Enforces an Antisense Bias on piRNA Pools by Binding AGO3 in the Drosophila Germline. Mol Cell 59:553-563.

Schmidt T. 1998. Genomes, genes and junk: the large-scale organization of plant chromosomes. Trends in Plant Science 3:195-199.

Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard HF. 2001. Genomic and genetic definition of a functional human centromere. Science 294:109-115.

Sienski G, Donertas D, Brennecke J. 2012. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell 151:964-980.

Simkin A, Wong A, Poh YP, Theurkauf WE, Jensen JD. 2013. Recurrent and recent selective sweeps in the piRNA pathway. Evolution 67:1081-1090.

Sproul JS, Khost DE, Eickbush DG, Negm S, Wei X, Wong I, Larracuente AM. 2020. Dynamic Evolution of Euchromatic Satellites on the X Chromosome in *Drosophila melanogaster* and the simulans Clade. Mol Biol Evol 37:2241-2256.

Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. 2017. Phase separation drives heterochromatin domain formation. Nature 547:241-245.

Swanson EC, Manning B, Zhang H, Lawrence JB. 2013. Higher-order unfolding of satellite heterochromatin is a consistent and early event in cell senescence. J Cell Biol 203:929-942.

Tao Y, Masly JP, Araripe L, Ke Y, Hartl DL. 2007. A sex-ratio meiotic drive system in Drosophila simulans. I: an autosomal suppressor. PLoS Biol 5:e292.

Ting DT, Lipson D, Paul S, Brannigan BW, Akhavanfard S, Coffman EJ, Contino G, Deshpande V, Iafrate AJ, Letovsky S, et al. 2011. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. Science 331:593-596.

Ugarkovic D. 2005. Functional elements residing within satellite DNAs. EMBO reports 6:1035-1039.

Usakin L, Abad J, Vagin VV, de Pablos B, Villasante A, Gvozdev VA. 2007. Transcription of the 1.688 satellite DNA family is under the control of RNA interference machinery in Drosophila melanogaster ovaries. Genetics 176:1343-1349.

Vermaak D, Henikoff S, Malik HS. 2005. Positive selection drives the evolution of rhino, a member of the heterochromatin protein 1 family in Drosophila. PLoS Genet 1:96-108.

Wang SH, Elgin SCR. 2011. Drosophila Piwi functions downstream of piRNA production mediating a chromatin-based transposon silencing mechanism in female germ line. Proceedings of the National Academy of Sciences of the United States of America 108:21164-21169.

Wang W, Han BW, Tipping C, Ge DT, Zhang Z, Weng Z, Zamore PD. 2015. Slicing and Binding by Ago3 or Aub Trigger Piwi-Bound piRNA Production by Distinct Mechanisms. Mol Cell 59:819-830.

Wehr K, Swan A, Schupbach T. 2006. Deadlock, a novel protein of Drosophila, is required for germline maintenance, fusome morphogenesis and axial patterning in oogenesis and associates with centrosomes in the early embryo. Dev Biol 294:406-417.

Wei KH, Lower SE, Caldas IV, Sless TJS, Barbash DA, Clark AG. 2018. Variable Rates of Simple Satellite Gains across the Drosophila Phylogeny. Mol Biol Evol 35:925-941.

Wei, X. Dmelanogaster satDNA regulation repository [Internet]. 2020. Last accessed 8/20/20 Available from: https://github.com/LarracuenteLab/Dmelanogaster_satDNA_regulation

Weiler KS, Wakimoto BT. 1995. Heterochromatin and gene expression in Drosophila. Annu Rev Genet 29:577-605.

Wickersheim ML, Blumenstiel JP. 2013. Terminator oligo blocking efficiently eliminates rRNA from Drosophila small RNA sequencing libraries. Biotechniques 55:269-272.

Wu CI, Lyttle TW, Wu ML, Lin GF. 1988. Association between a satellite DNA sequence and the Responder of *Segregation Distorter* in *D. melanogaster*. Cell 54:179-189.

Yi M, Chen F, Luo M, Cheng Y, Zhao H, Cheng H, Zhou R. 2014. Rapid evolution of piRNA pathway in the teleost fish: implication for an adaptation to transposon diversity. Genome Biol Evol 6:1393-1407.

Zhang F, Wang J, Xu J, Zhang Z, Koppetsch BS, Schultz N, Vreven T, Meignin C, Davis I, Zamore PD, et al. 2012. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. Cell 151:871-884.

Zhang Z, Wang J, Schultz N, Zhang F, Parhad SS, Tu S, Vreven T, Zamore PD, Weng Z, Theurkauf WE. 2014. The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. Cell 157:1353-1363.

Zhu Q, Pao GM, Huynh AM, Suh H, Tonnu N, Nederlof PM, Gage FH, Verma IM. 2011. BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. Nature 477:179-184.

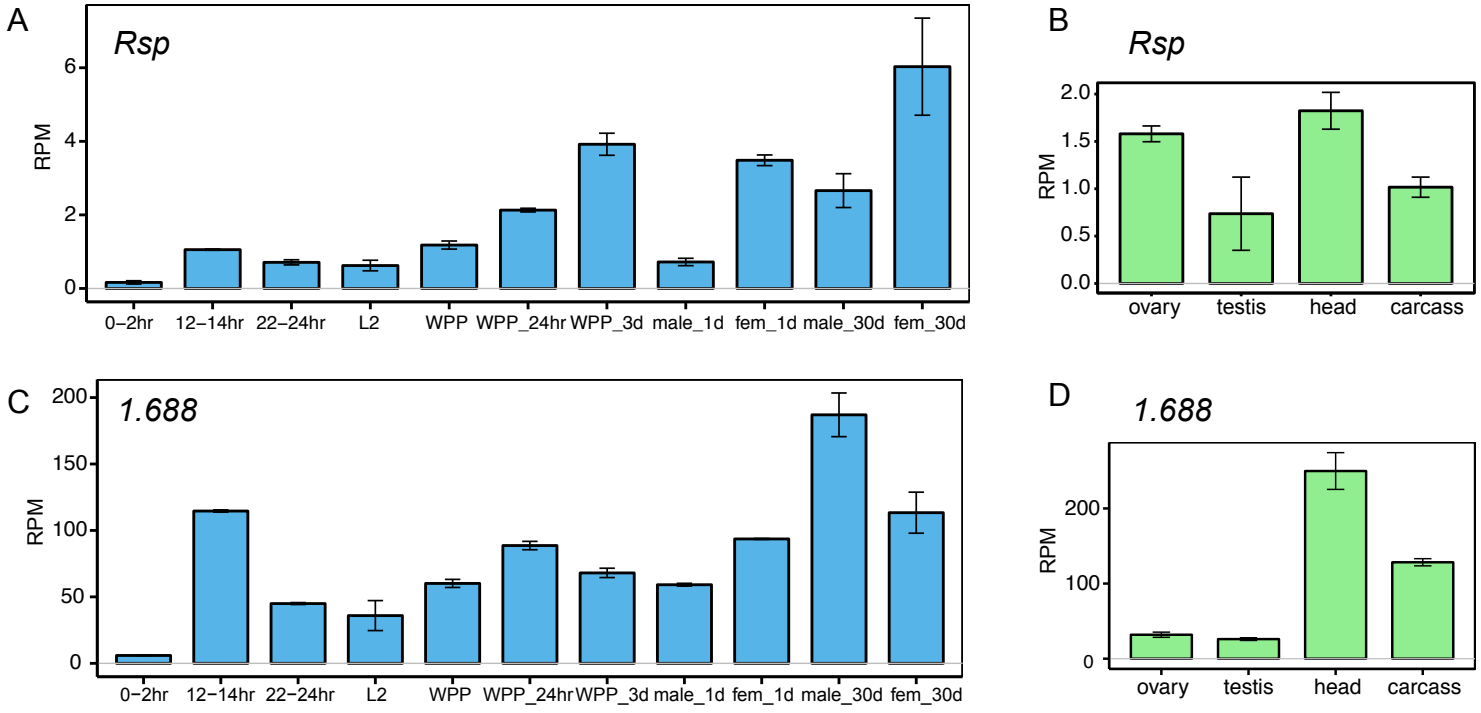


Figure 1-S1. SatDNAs are expressed across tissues and developmental stages. *Rsp* transcription level in various developmental stages (A) and tissues (B), and corresponding results for *1.688* satellite are shown in (C) and (D). L2: 2nd instar larvae. WPP: white prepupae. Fem: female. Carcass: whole body without the head, reproductive organs and digestive tract. Data from modENCODE (Graveley, et al. 2011; Brown, et al. 2014).

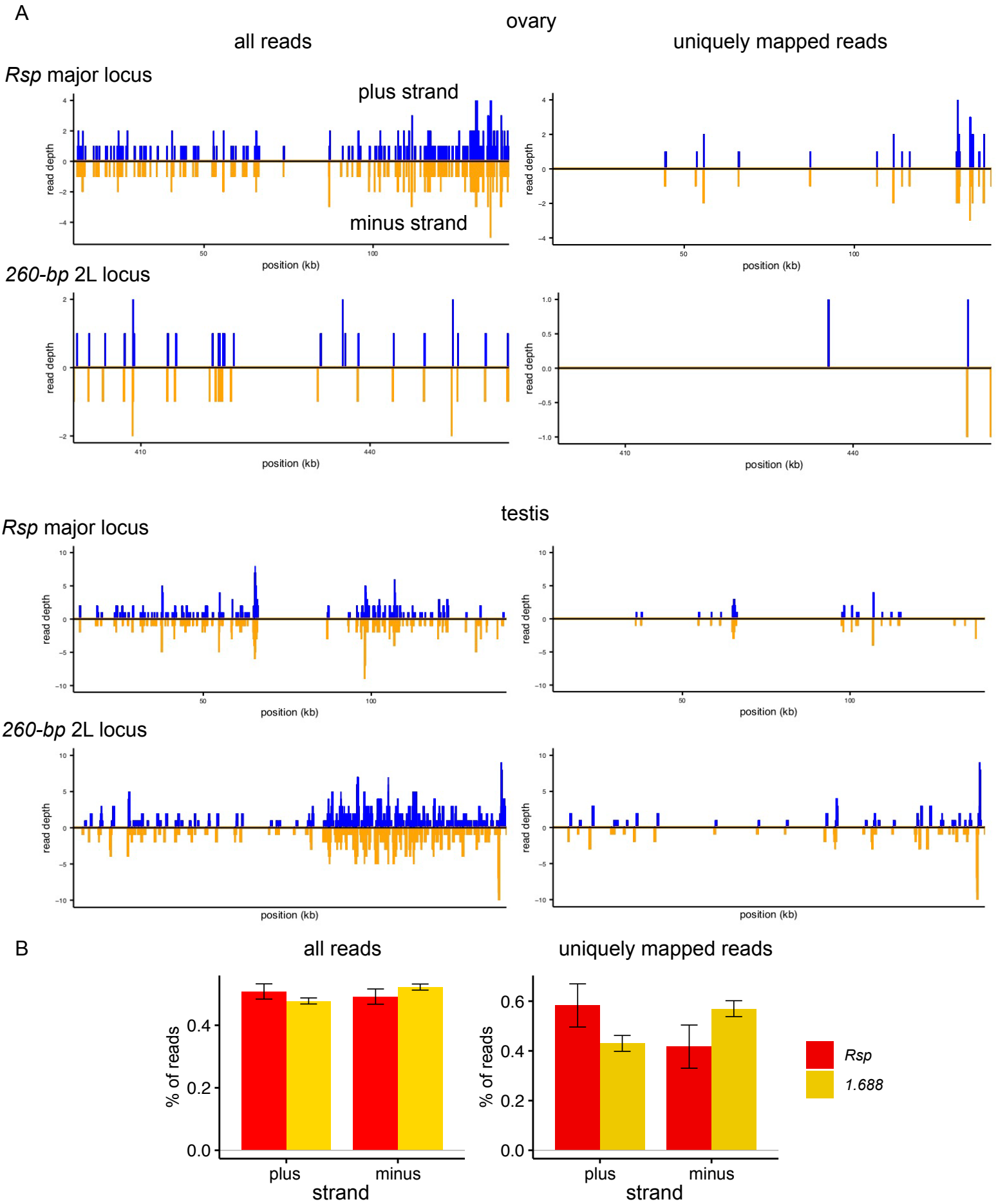
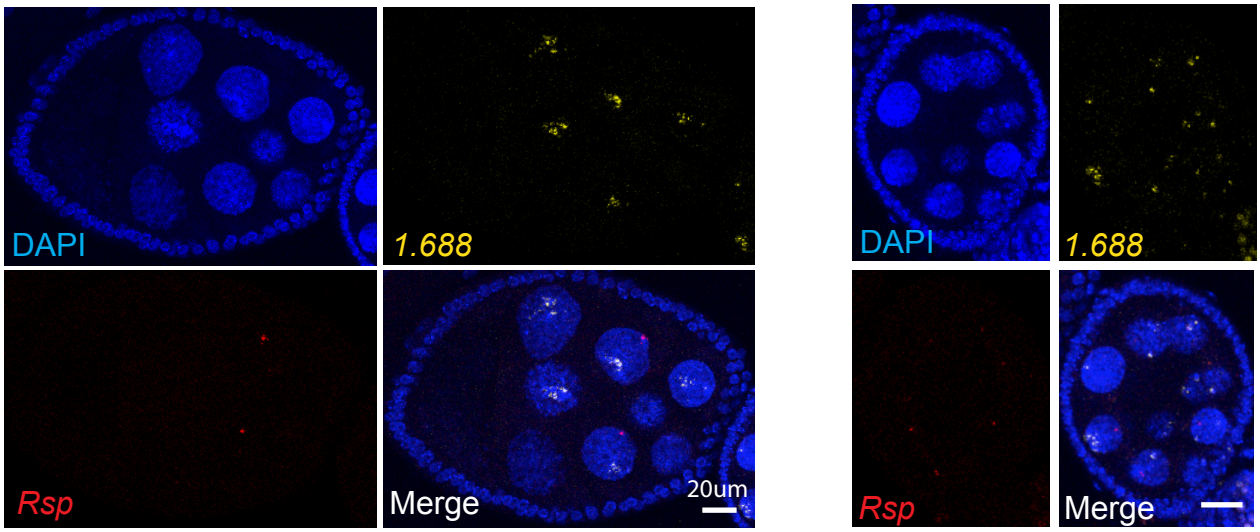
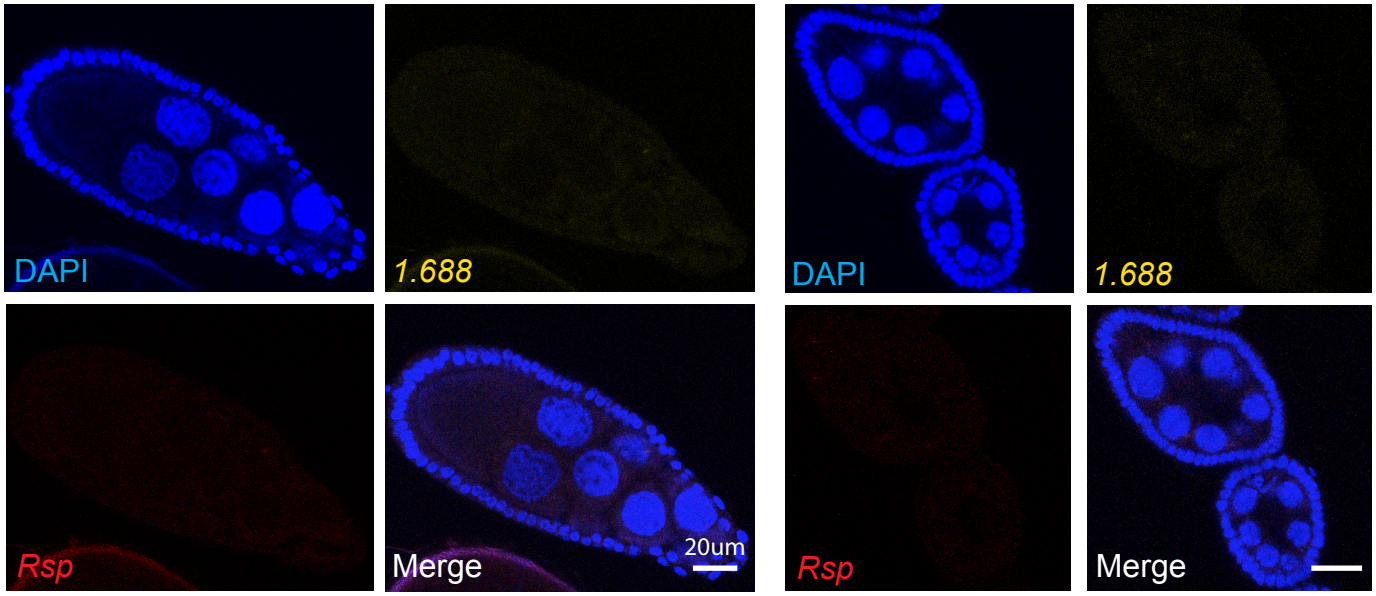


Figure 1-S2. SatDNAs are transcribed from both strands. (A) Read depth of both plus and minus strands for the *Rsp* 2R major locus and *260-bp* 2L locus. Read depth of all reads mapped to each locas is shown on the left, and depth of uniquely mapped reads is shown on the right. (B) Percentages of reads mapped to plus and minus strands for reads that mapped to all genomic copies of *Rsp* or *1.688* satellites, for all mapped reads (left) and uniquely mapped reads (right). Ovary data is from (Mohn, et al. 2014; Andersen, et al. 2017), and testis data is from this study.

A RNA FISH ovary



B RNaseA treated ovary



C RNaseH treated ovary (after probe hybrid)

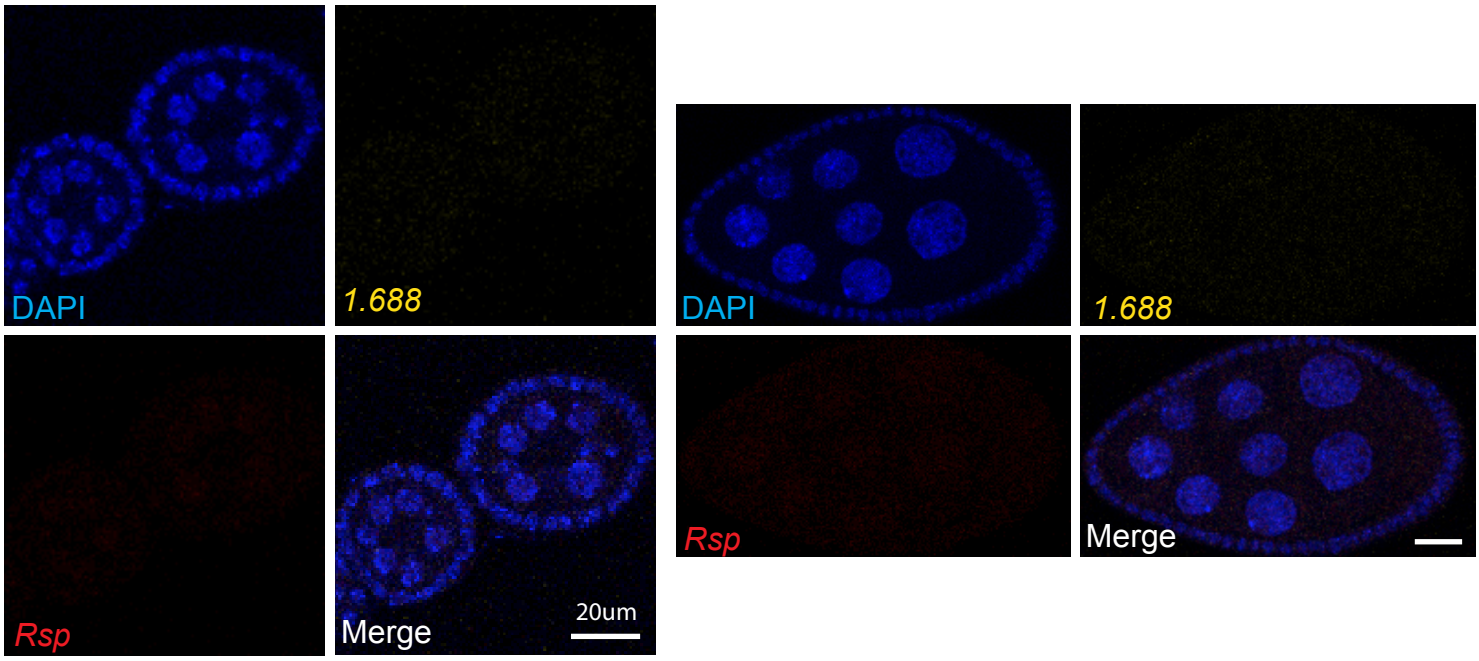


Figure 1-S3. RNA FISH signals of satDNAs are from RNAs, not DNAs. RNA FISH of *Rsp* and *1.688* in ovary (A), with RNaseA treated prior to probe hybridization (B) and with RNaseH treated after DNA probe hybridization (C).

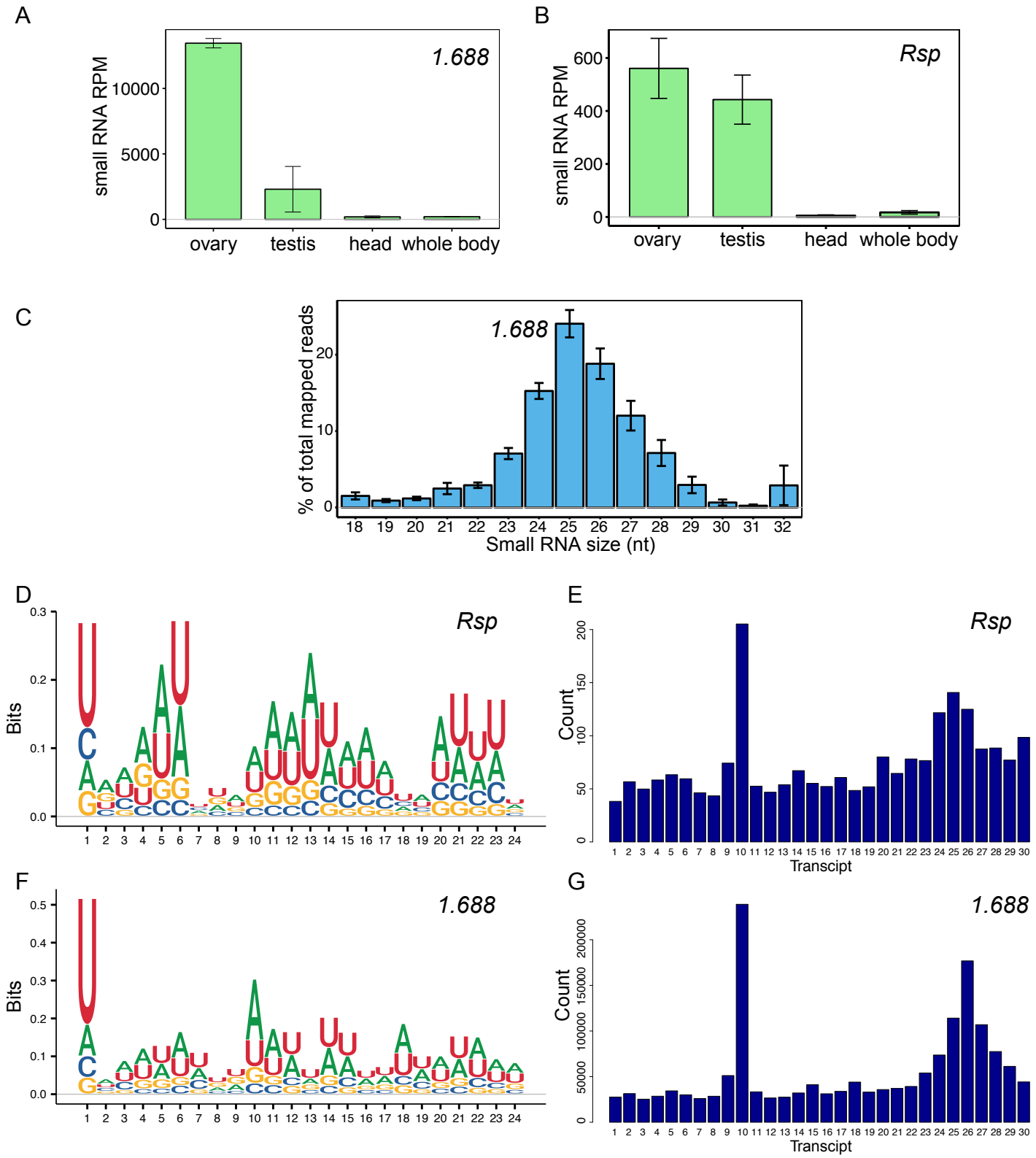


Figure 2-S1. Satellite derived RNAs are mainly processed into piRNAs in the germline. (A-B) Small RNA levels (RPM) in ovary, testis, head and whole body (head removed) for *Rsp* (A) and *1.688* (B). (C) Size distribution of *1.688* small RNAs in gonads. (D and F) Relative nucleotide bias of each position in small RNAs from *Rsp* (D) and *1.688* (F) satellite in ovary. (E and G) 5'-to-5' end distance analysis results of small RNAs from *Rsp* (E) and *1.688* (G) satellite in ovary show bias of 10-nt overlap. Z-score=4.55 for *Rsp* and 6.85 for *1.688* satellite. Data from (Ghildiyal, et al. 2010; Rozhkov, et al. 2010; Fagegaltier, et al. 2014; Mohn, et al. 2014; Quenerch'du, et al. 2016; Andersen, et al. 2017; Parhad, et al. 2017).

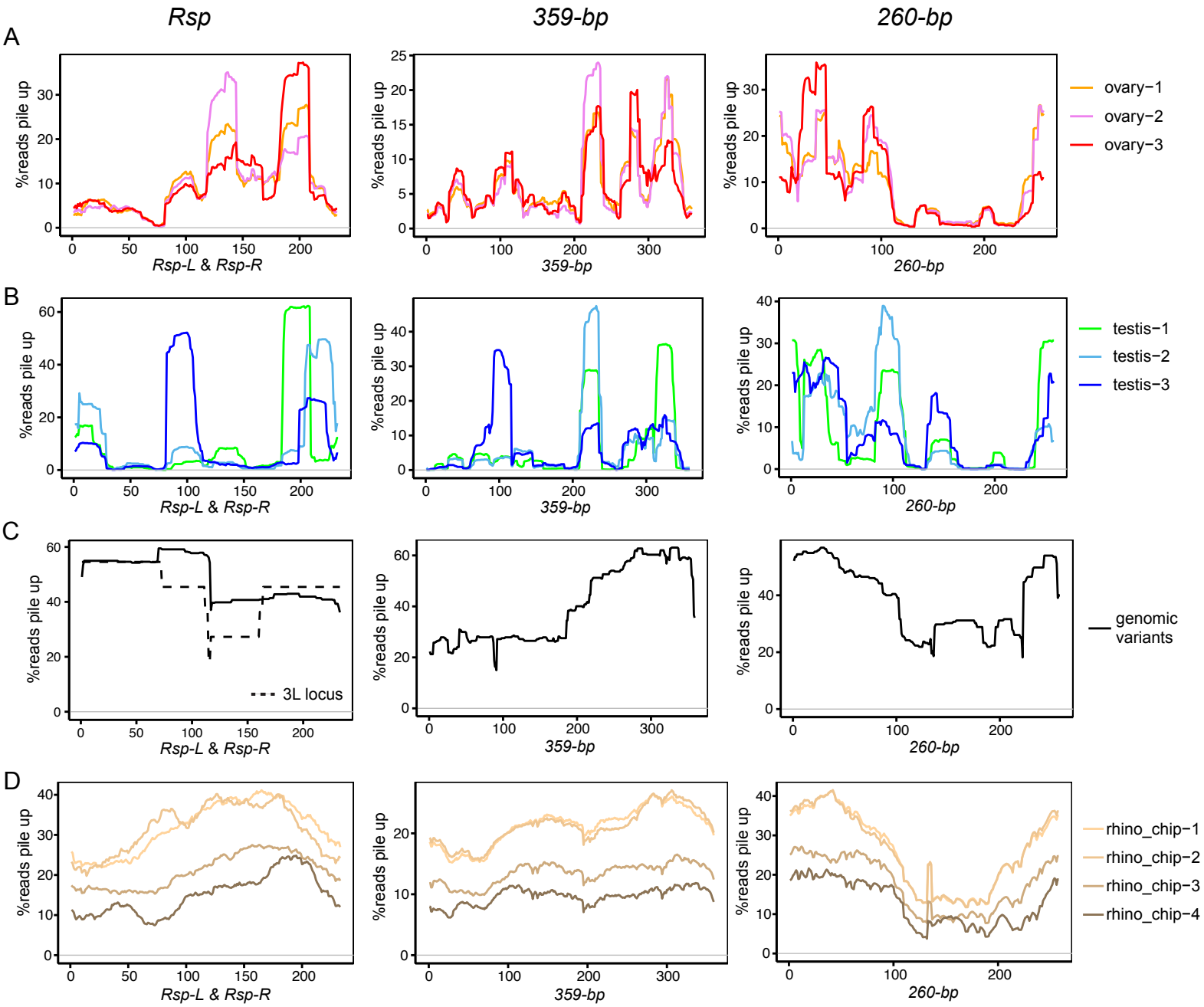


Figure 2-S2. Non-uniform distribution of piRNA reads along satDNA consensus sequences. Small RNA reads pileup along *Rsp*, *260-bp* and *359-bp* (*1.688* satellite) consensus sequences for ovary (A) and testis (B). (C) Alignment depth of genomic satDNA repeat variants extracted from the genome assembly along consensus sequences. *Rsp* variants resident inside the *Ago3* intron on chromosome 3L (3L locus) is shown separately in dashed line. (D) Rhino ChIP-seq reads pileup along the satellite consensus sequences. Data from (Rozhkov, et al. 2010; Mohn, et al. 2014; Quenerch'du, et al. 2016; Andersen, et al. 2017; Parhad, et al. 2017; Zhang, et al. 2014).

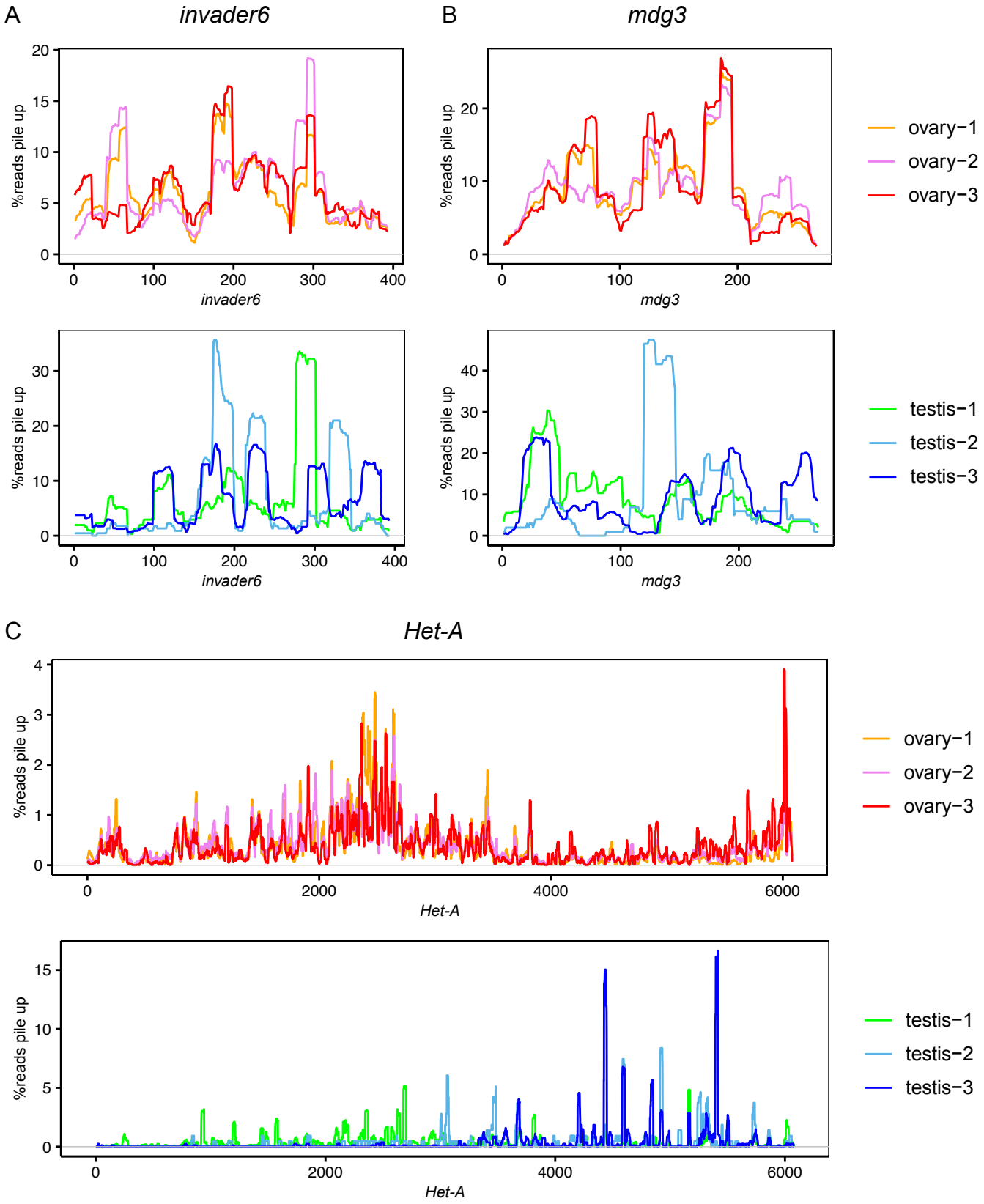


Figure 2-S3. Non-uniform distribution of piRNA reads along the germline-dominant TE consensus sequences. Small RNA reads pileup along *Invader6* (A), *mdg3* (B) and *Het-A* (C) consensus sequences for ovary and testis. Data from (Rozhkov, et al. 2010; Mohn, et al. 2014; Quenerch'du, et al. 2016; Parhad, et al. 2017; Andersen, et al. 2017)

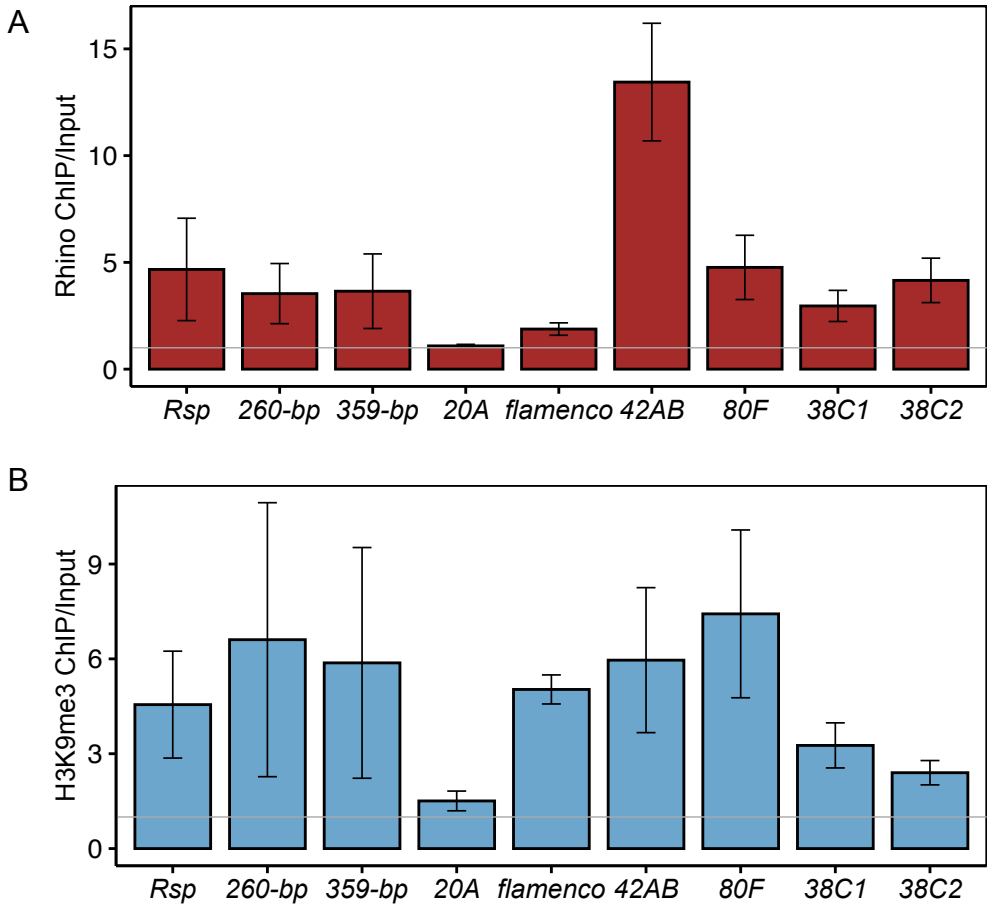


Figure 2-S4. ChIP-seq results shows the chromatin state of satDNAs, uni-strand piRNA clusters (*20A* and *flamenco*) and dual-strand piRNA clusters (*42AB*, *80F* and *38C1/2*). (A) Rhino ChIP-seq enrichment scores for each satDNA and piRNA cluster indicate that the chromatin state of satellite regions resembles dual-strand piRNA clusters. (B) H3K9me3 ChIP-seq enrichment scores indicate that satDNAs and most piRNA clusters are enriched for H3K9me3. Data from (Parhad, et al. 2017; Klenov, et al. 2014; Le Thomas, et al. 2014; Mohn, et al. 2014; Zhang, et al. 2014).

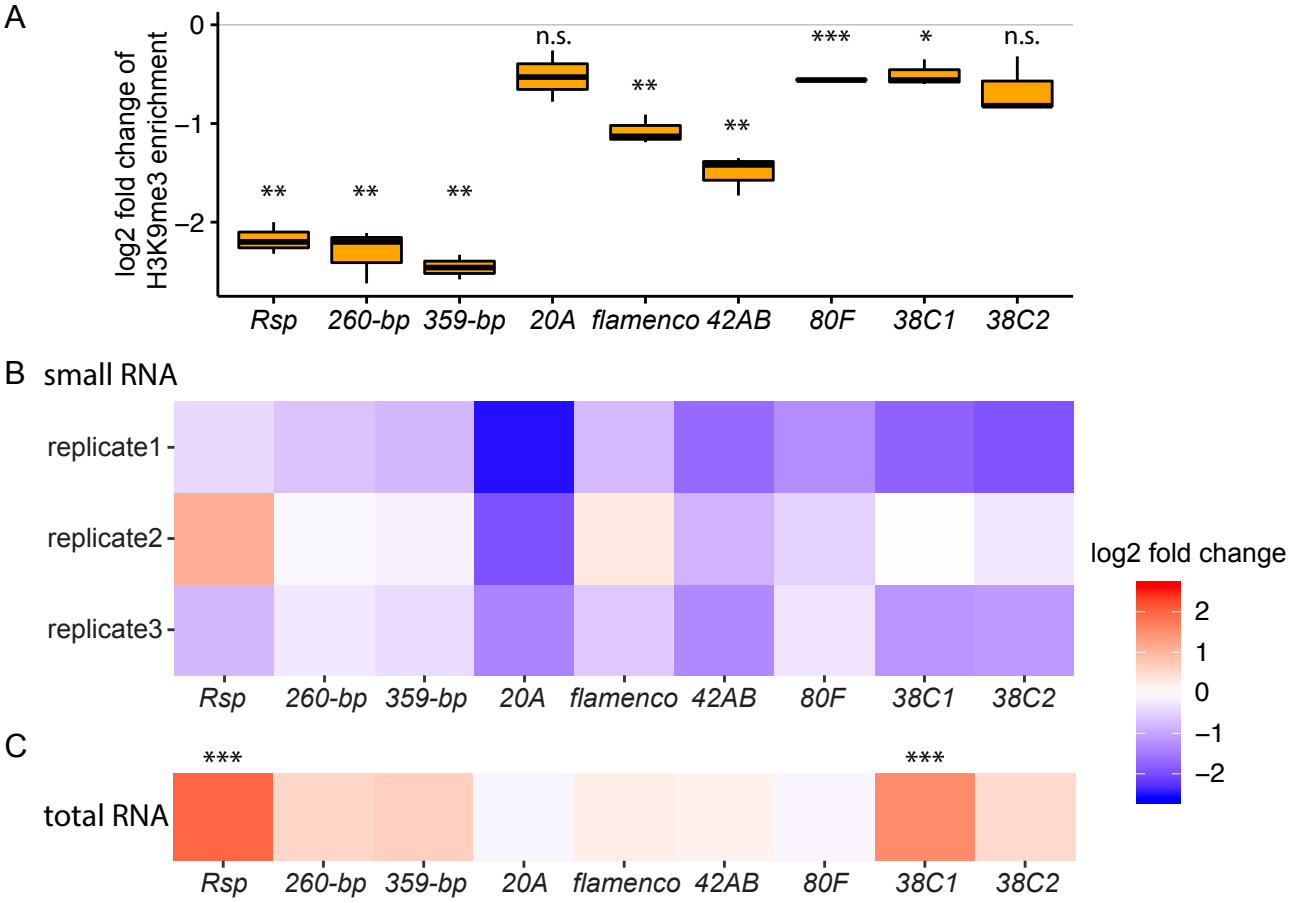


Figure 2-S5. H3K9me3 and RNA level changes in *piwi* embryonic knockdown ovaries compared to control. (A) Log2 fold change of H3K9me3 ChIP/input enrichment shows satDNA H3K9me3 levels decrease. P-values are estimated by one sample t-test (mu=0) with FDR correction (Benjamini 1995). (B) Log2 fold change of small RNA abundance shows satDNA small RNA levels decrease with variation observed for replicate2. Small RNA aundance is normalized to the number of reads mapped to miRNAs. (C) Log2 fold change of total RNA abundance shows satDNA long RNA levels increase. Adjusted p-values for total RNA changes are reported by DESeq2. *20A* and *flamenco* are uni-strand piRNA clusters, *42AB*, *80F* and *38C1/2* are dual-strand piRNA clusters. * adjusted p-value<0.05, ** adjusted p-value<0.01. *** adjusted p-value< 0.001. Data from (Akkouche, et al. 2017).

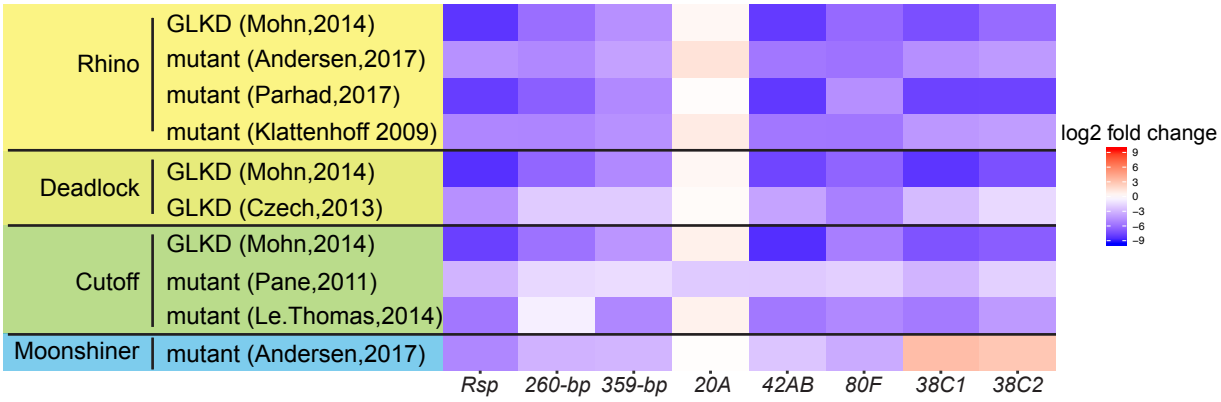


Figure 3-S1. SatDNA loci are regulated by the heterochromatin-dependent transcription machinery in Drosophila ovaries. Heatmap showing the quantification of small RNA abundance from mutants of *rhino*, *cutoff*, *deadlock*, and *moonshiner* for satDNAs and piRNA clusters, normalized to the *flamenco* piRNA cluster. GLKD: germline knockdown. Data from (Klattenhoff, et al. 2009; Pane, et al. 2011; Czech, et al. 2013; Le Thomas, et al. 2014; Mohn, et al. 2014; Andersen, et al. 2017; Parhad, et al. 2017).

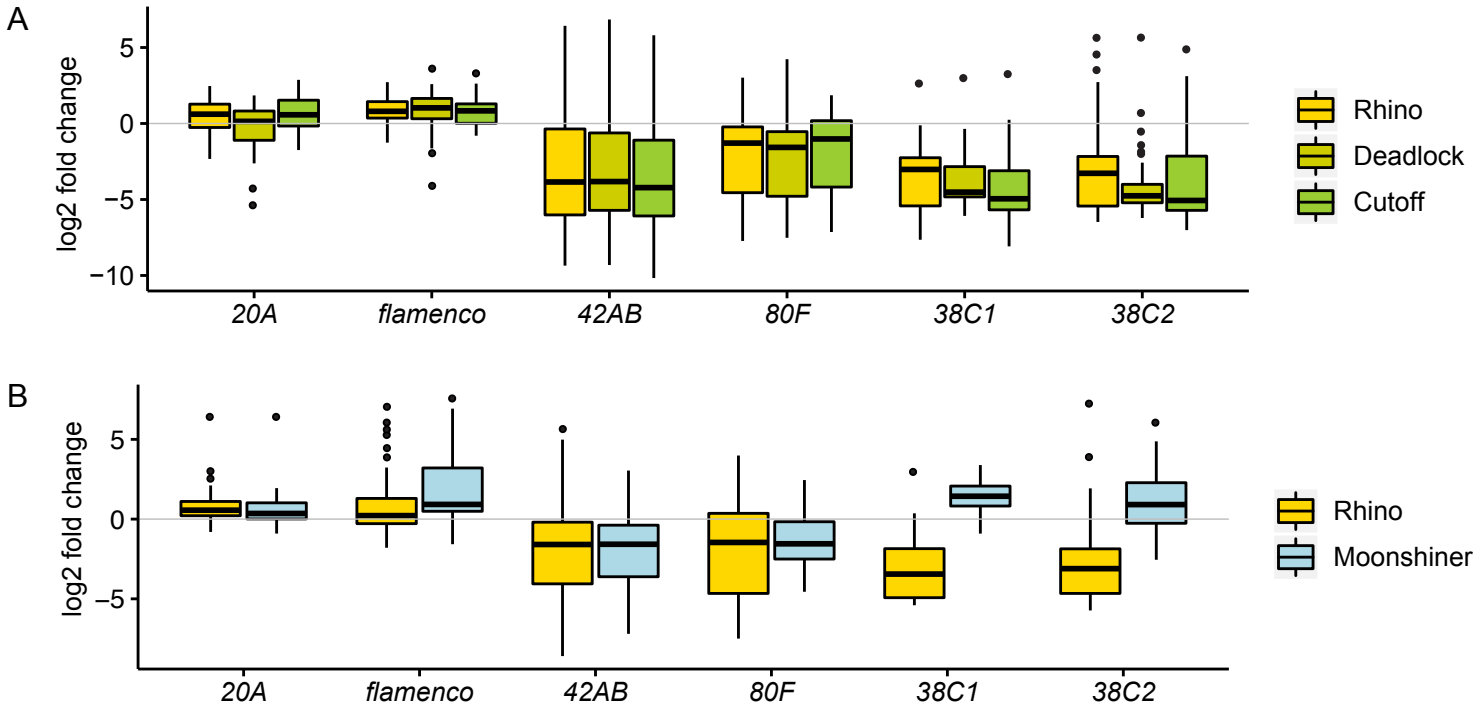


Figure 3-S2. RDC and Moon mutants affect piRNA precursor transcription at piRNA clusters. Boxplots showing the quantification of total RNA abundance for piRNA clusters in RDC (A) and Rhi/Moon (B) mutant ovaries relative to wild type (log2 fold change of 1-kb windows). *20A* and *flamenco* are uni-strand piRNA clusters, *42AB*, *80F* and *38C1/2* are dual-strand piRNA clusters. Data from (Mohn, et al. 2014) for (A) and (Andersen, et al. 2017) for (B). A sliding window method is not feasible for analysing satDNAs because it relies on uniquely mapped reads, of which there are relatively few at satDNA loci. Therefore, we instead count reads mapping to all genomic repeat variants for each satDNA, and piRNA cluster (the whole locus). The locus-wide results (Table S8) are comparable with the sliding window analysis, but more conservative.

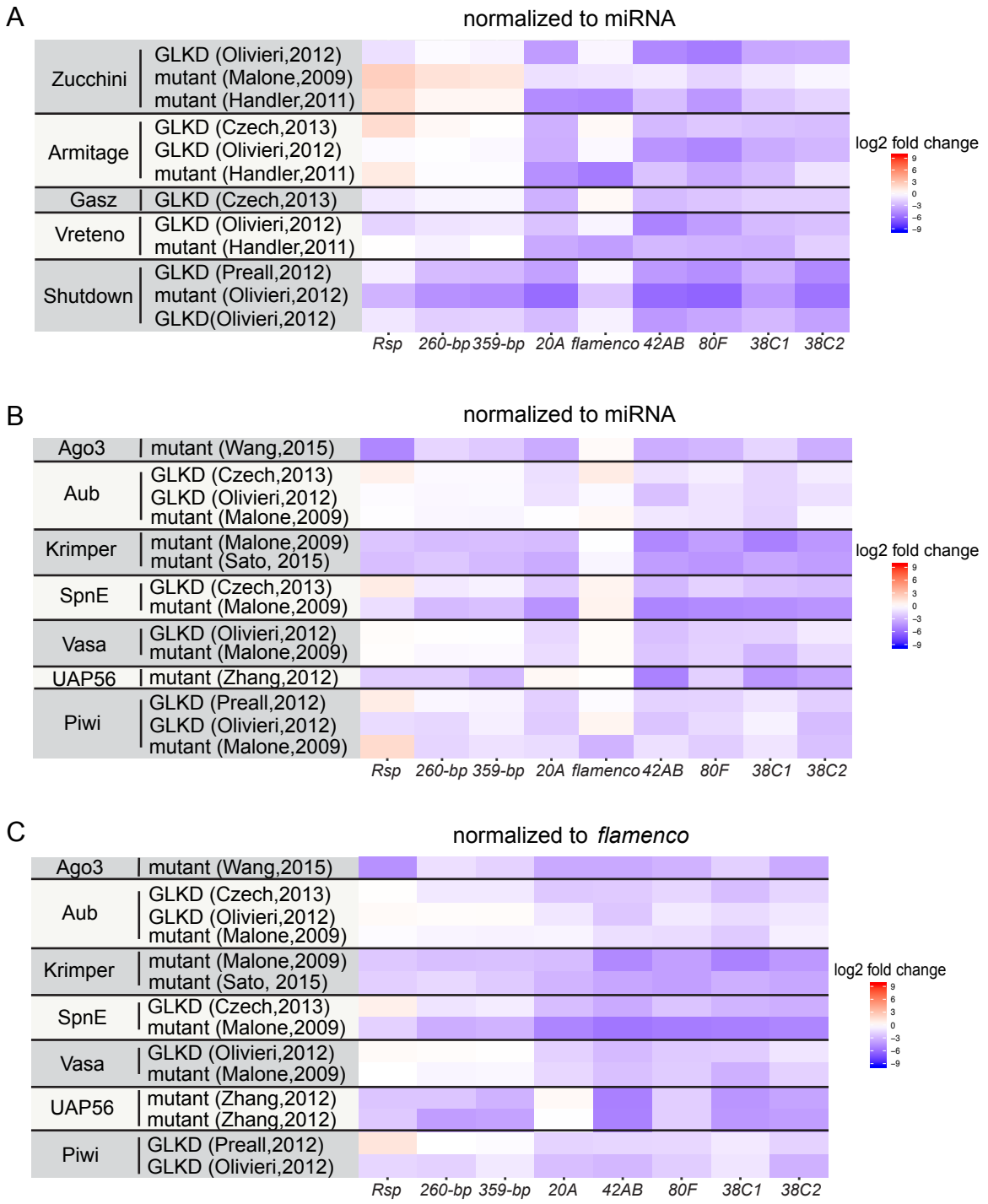


Figure 3-S3. SatDNA piRNA production is affected in mutants of pathways involving piRNA precursor export, primary piRNA biogenesis, and the ping-pong cycle. Heatmap showing the quantification of small RNA abundance from mutants of proteins in the primary piRNA pathway (A), pathway for piRNA precursor export from the nucleus and ping-pong pathway (B-C). The results are consistent between replicates/studies about Vreteno, Shutdown, Krimper and UAP56, suggesting that complex satDNA piRNA production is regulated by these proteins. For some mutants including Aub, SpnE, and Piwi, results differed between complex satDNAs — *e.g.* the *1.688* family of satDNAs (*260-bp* and *359-bp*) show consistently decreased piRNA levels, but not *Rsp* for Piwi. For other proteins such as Zucchini and Armitage, the patterns of change for satDNAs are variable between different datasets — *e.g.* satDNAs show decreased piRNA levels in one of the datasets, but increased levels in the other two datasets for Zucchini. For all the datasets analyzed, all piRNA clusters behave as reported previously. GLKD: germline knockdown. *Flamenco* is only expressed in somatic tissues, GLKD doesn't affect its abundance. Data from (Malone, et al. 2009; Handler, et al. 2011; Olivieri, et al. 2012; Preall, et al. 2012; Zhang, et al. 2012; Czech, et al. 2013; Sato, et al. 2015; Wang, et al. 2015).

Table S1. List of datasets used in this paper

accession	tissue	citation
small RNA-seq		
SRR060645	testis	Rozhkov et al., 2010
SRR2046463	testis	Quenerch'du et al., 2016
	testis	this study
SRR1259520	head	Fagegaltier et al., 2014
SRR1259521	head	Fagegaltier et al., 2014
SRR031692	head	Ghildiyal et al., 2010
SRR1259523	whole body (head removed)	Fagegaltier et al., 2014
SRR1259522	whole body (head removed)	Fagegaltier et al., 2014
SRR1187947	ovary	Mohn et al., 2014
SRR1187948	ovary	Mohn et al., 2014
SRR1187949	ovary	Mohn et al., 2014
SRR1187950	ovary	Mohn et al., 2014
SRR5445299	ovary	Andersen et al., 2017
SRR5445300	ovary	Andersen et al., 2017
SRR5445301	ovary	Andersen et al., 2017
SRR5445307	ovary	Andersen et al., 2017
SRR5445318	ovary	Andersen et al., 2017
SRR891254	ovary	Pane et al., 2011
SRR891255	ovary	Pane et al., 2011
SRR891256	ovary	Pane et al., 2011
SRR891257	ovary	Pane et al., 2011
SRR5803096	ovary	Parhad et al., 2017
SRR5803097	ovary	Parhad et al., 2017
SRR1524976	ovary	Le.Thomas et al., 2014
SRR1524977	ovary	Le.Thomas et al., 2014
SRR037091	ovary	Klattenhoff et al., 2009
SRR037092	ovary	Klattenhoff et al., 2009
SRR827714	ovary	Czech et al., 2013
SRR827716	ovary	Czech et al., 2013
SRR827717	ovary	Czech et al., 2013
SRR827718	ovary	Czech et al., 2013
SRR827719	ovary	Czech et al., 2013
SRR827720	ovary	Czech et al., 2013
SRR499784	ovary	Preall et al., 2012
SRR499785	ovary	Preall et al., 2012
SRR499786	ovary	Preall et al., 2012
SRR513384	ovary	Olivieri et al, 2012
SRR513385	ovary	Olivieri et al, 2012
SRR513386	ovary	Olivieri et al, 2012
SRR513387	ovary	Olivieri et al, 2012
SRR513388	ovary	Olivieri et al, 2012
SRR513389	ovary	Olivieri et al, 2012

SRR513390	ovary	Olivieri et al, 2012
SRR513391	ovary	Olivieri et al, 2012
SRR513392	ovary	Olivieri et al, 2012
SRR513395	ovary	Olivieri et al, 2012
SRR513396	ovary	Olivieri et al, 2012
SRR016828	ovary	Malone et al., 2009
SRR016829	ovary	Malone et al., 2009
SRR016822	ovary	Malone et al., 2009
SRR016823	ovary	Malone et al., 2009
SRR016826	ovary	Malone et al., 2009
SRR016827	ovary	Malone et al., 2009
SRR016832	ovary	Malone et al., 2009
SRR016833	ovary	Malone et al., 2009
SRR016836	ovary	Malone et al., 2009
SRR016837	ovary	Malone et al., 2009
SRR016838	ovary	Malone et al., 2009
SRR016821	ovary	Malone et al., 2009
SRR407284	ovary	Zhang et al., 2012
SRR407285	ovary	Zhang et al., 2012
SRR407286	ovary	Zhang et al., 2012
SRR407287	ovary	Zhang et al., 2012
SRR407288	ovary	Zhang et al., 2012
SRR317110	ovary	Handler et al., 2011
SRR317111	ovary	Handler et al., 2011
SRR317112	ovary	Handler et al., 2011
SRR317113	ovary	Handler et al., 2011
SRR317116	ovary	Handler et al., 2011
SRR317117	ovary	Handler et al., 2011
SRR2042568	ovary	Sato et al., 2015
SRR2042572	ovary	Sato et al., 2015
SRR1944886	ovary	Wang et al., 2015
SRR1944887	ovary	Wang et al., 2015
SRR1944875	ovary	Wang et al., 2015
SRR3657660	ovary	Akkouche et al., 2017
SRR3657661	ovary	Akkouche et al., 2017
SRR3657662	ovary	Akkouche et al., 2017
SRR3657663	ovary	Akkouche et al., 2017
SRR3657664	ovary	Akkouche et al., 2017
ChIP-seq for Rhino		
SRR5803091	ovary	Parhad et al., 2017
SRR5803102	ovary	Parhad et al., 2017
SRR5806898	ovary	Parhad et al., 2017
SRR5806896	ovary	Parhad et al., 2017
SRR1024309	ovary	Zhang et al., 2014
SRR1024325	ovary	Zhang et al., 2014

SRR1024304	ovary	Zhang et al., 2014
SRR1024307	ovary	Zhang et al., 2014
ChIP-seq for H3K9me3		
SRR653343	ovary	Le Thomas et al., 2013
SRR653344	ovary	Le Thomas et al., 2013
SRR653350	ovary	Le Thomas et al., 2013
SRR653351	ovary	Le Thomas et al., 2013
SRR1187981	ovary	Mohn et al., 2014
SRR1187982	ovary	Mohn et al., 2014
SRR1187971	ovary	Mohn et al., 2014
SRR1187972	ovary	Mohn et al., 2014
SRR1213172	ovary	Klenov et al., 2014
SRR1213173	ovary	Klenov et al., 2014
SRR1213168	ovary	Klenov et al., 2014
SRR1213169	ovary	Klenov et al., 2014
SRR3657644	ovary	Akkouche et al., 2017
SRR3657645	ovary	Akkouche et al., 2017
SRR3657646	ovary	Akkouche et al., 2017
SRR3657647	ovary	Akkouche et al., 2017
SRR3657648	ovary	Akkouche et al., 2017
SRR3657649	ovary	Akkouche et al., 2017
SRR3657652	ovary	Akkouche et al., 2017
SRR3657653	ovary	Akkouche et al., 2017
SRR3657654	ovary	Akkouche et al., 2017
SRR3657655	ovary	Akkouche et al., 2017
RIP-seq or Aub, Ago3 and Piwi		
GSM154620	ovary	Brennecke et al, 2007
GSM154621	ovary	Brennecke et al, 2007
GSM154622	ovary	Brennecke et al, 2007
SRR2042569	ovary	Sato et al, 2015
SRR2042570	ovary	Sato et al, 2015
SRR2042571	ovary	Sato et al, 2015
SRR2042575	ovary	Sato et al, 2015
SRR1746862	ovary	Mohn et al, 2015
SRR1746863	ovary	Mohn et al, 2015
SRR1746864	ovary	Mohn et al, 2015
SRR1746865	ovary	Mohn et al, 2015
total RNA-seq		
SRR1187952	ovary	Mohn et al., 2014
SRR1187953	ovary	Mohn et al., 2014
SRR1187954	ovary	Mohn et al., 2014
SRR1187955	ovary	Mohn et al., 2014
SRR1187956	ovary	Mohn et al., 2014
SRR5445322	ovary	Andersen et al., 2017
SRR5445323	ovary	Andersen et al., 2017

SRR5445324	ovary	Andersen et al., 2017
SRR5445325	ovary	Andersen et al., 2017
SRR5445326	ovary	Andersen et al., 2017
SRR5445327	ovary	Andersen et al., 2017
	testis	this study
SRR3657656	ovary	Akkouche et al., 2017
SRR3657657	ovary	Akkouche et al., 2017
SRR3657658	ovary	Akkouche et al., 2017
SRR3657659	ovary	Akkouche et al., 2017
SRR767606	0-2hr embryo	modENCODE
SRR1197337	0-2hr embryo	modENCODE
SRR767613	12-14hr embryo	modENCODE
SRR1197332	12-14hr embryo	modENCODE
SRR767625	22-24hr embryo	modENCODE
SRR1197328	22-24hr embryo	modENCODE
SRR767623	L2	modENCODE
SRR1197425	L2	modENCODE
SRR1197324	L2	modENCODE
SRR767617	WPP	modENCODE
SRR1197290	WPP	modENCODE
SRR767604	WPP_24hr	modENCODE
SRR1197288	WPP_24hr	modENCODE
SRR767611	WPP_3d	modENCODE
SRR1197285	WPP_3d	modENCODE
SRR1197317	adult_fem_1d	modENCODE
SRR767621	adult_fem_1d	modENCODE
SRR1197314	adult_fem_30d	modENCODE
SRR1197411	adult_fem_30d	modENCODE
SRR1197315	adult_male_1d	modENCODE
SRR767619	adult_male_1d	modENCODE
SRR1197413	adult_male_30d	modENCODE
SRR1197311	adult_male_30d	modENCODE
SRR1197486	adult_fem_4d_head	modENCODE
SRR1197374	adult_fem_4d_head	modENCODE
SRR1197294	adult_fem_4d_head	modENCODE
SRR1197480	adult_fem_4d_ovary	modENCODE
SRR1197376	adult_fem_4d_ovary	modENCODE
SRR1197296	adult_fem_4d_ovary	modENCODE
SRR1197301	adult_mal_4d_testes	modENCODE
SRR1197381	adult_mal_4d_testes	modENCODE
SRR1197377	adult_mal_4d_testes	modENCODE
SRR1197481	adult_mal_4d_head	modENCODE
SRR1197293	adult_mal_4d_head	modENCODE
SRR1197373	adult_mal_4d_head	modENCODE
SRR1197372	adult_mixed_sex_4d_carcass	modENCODE

SRR1197292	adult_mixed_sex_4d_carcass	modENCODE
SRR1197475	adult_mixed_sex_4d_carcass	modENCODE
mRNA-seq		
SRR100282	adult_fem_4d_head	modENCODE
SRR100278	adult_fem_4d_head	modENCODE
SRR070430	adult_fem_4d_head	modENCODE
SRR070417	adult_fem_4d_ovary	modENCODE
SRR070396	adult_fem_4d_ovary	modENCODE
SRR350960	adult_mal_4d_testes	modENCODE
SRR350961	adult_mal_4d_testes	modENCODE
SRR070422	adult_mal_4d_testes	modENCODE
SRR070400	adult_mal_4d_head	modENCODE
SRR070416	adult_mal_4d_head	modENCODE
SRR070387	adult_mixed_sex_4d_carcass	modENCODE
SRR070402	adult_mixed_sex_4d_carcass	modENCODE

note: modENCODE (Graveley et al., 2011; Brown et al., 2014)

Table S2. Count in RPM values of *Rsp*/*1.688* transcripts in total and poly-A RNA-seq datasets from various tissues (data from modENCODE (Graveley, et al. 2011; Brown, et al. 2014))

Repeat	RNA-seq type	Replicate	ovary	testis	carcass	head
Rsp	Total RNA	1	1.74	0.46	1.22	1.56
		2	1.46	1.5	0.97	1.57
		3	1.54	0.25	0.86	1.36
		4				2.54
		5				2.3
		6				1.61
	polyA RNA	1	0.17	0.01	0	0
		2	0.13	0	0.01	0
		3		0		0
		4				0.03
		5				0.08
1.688	Total RNA	1	37.17	25.09	137.87	191.56
		2	33.08	29.56	122.89	211.17
		3	25.42	23.55	124.27	228.41
		4				217.12
		5				316.69
		6				333.75
	polyA RNA	1	2.3	3.31	1.47	3.97
		2	1.3	2.72	2.33	0.95
		3		4.25		0.97
		4				9.41
		5				6.04

Table S3. Slot blot estimate of *Rsp* copy number

fly strains	replicates							average	copy number
	1	2	3	4	5	6	7		
ZW144	0.2	0.21	0.27	0.22	0.15	0.25		0.2167	238.33
Ral357	0.69	0.65	0.35	0.58	0.64			0.5820	640.2
Ral380	2.74	2.29	1.9	1.95	1.75			2.1260	2338.6
lt pk cn bw	5.46	5.19	2.26	1.78	4.12	3.6	3.59	3.7143	4085.71

replicates: ratio of *Rsp* signal over *rp49* signal relative to that of Iso-1
copy number: average ratio multiplied by the 1100 *Rsp* copy estimate for Iso-1 (Khost et al., 2017)

TableS4. *Rsp* expression level correlates with its copy number in the genome

Method I: northern blot and slot blot		
strains	*Rsp* DNA repeat number estimated from Slot blot	*Rsp* RNA expression level quantified from Northern blot
ZW144	238	0.966552
Ral357	640	6.798309
Iso1	1100	7.934042
Ral380	2339	8.456499
It pk cn bw	4086	17.28229
Method II: qPCR and qRT-PCR		
strains	*Rsp* DNA repeat number estimated from qPCR	*Rsp* RNA expression level quantified from qRT-PCR
ZW144	53	0.000195628
Ral357	424	0.00190355633333333
Iso1	1100	0.0050540994
Ral380	2569	0.0061967918
It pk cn bw	5158	0.01619652775

Method I (Pearson's correlation coefficient r^2=0.93, p-value=0.02).
Method II (Pearson's correlation coefficient r^2=0.98, p-value=0.003).

Table S5. RIP-seq results for Piwi, Aub and Ago3 from ovaries (data from (Brennecke, et al. 2007; Mohn, et al. 2015; Sato, et al. 2015))

Piwi

	Sato, 2015	Brennecke, 2007	Mohn, 2015	Mohn, 2015	**Average**
359-bp	0.391652	0.208933311	0.618672	0.431478	0.41268383
260-bp	0.042386	0.197639618	0.179258	0.104962	0.1310614
356-bp	0.05678	0.084702694	0.129075	0.094881	0.09135967
353-bp	0.219146	0.141171156	0.491425	0.345738	0.29937004
Rsp	0.097497	0.191992772	0.124947	0.063186	0.11940569
HETRP	0.485753	1.451239483	0.833975	0.556106	0.83176837
(AAGAG)n	0.000034	0	0.000033	0.000028	0.00002375
42AB	6.629327		27.433707	16.435909	16.832981
80F	1.680195		2.542423	1.88418	2.03559933
miRNA (avarage)	2.3866E-05	5.96659E-05	0.000441527	0.0014678	0.00049821

Aub

	Sato, 2015	Brennecke, 2007	Mohn, 2015	**Average**
359-bp	0.547824	2.17859064	2.20039	1.6422682
260-bp	0.03082	0.92119419	0.283161	0.4117251
356-bp	0.074607	0.262237762	0.344509	0.2271179
353-bp	0.363891	0.161377084	2.076218	0.867162
Rsp	0.035022	0.215169446	0.115852	0.1220145
HETRP	2.778344	3.308230231	2.199115	2.7618964
(AAGAG)n	0	0	0.000022	7.333E-06
42AB	10.858128		32.23298	21.545554
80F	2.672089		2.652625	2.662357
miRNA (avarage)	3.58E-05	3.57995E-05	0.000417661	0.0001631

Ago3

	Sato, 2015	Sato, 2015	Brennecke, 2007	Mohn, 2015	**Average**
359-bp	0.562595	0.519152	0.487679671	1.4915	0.76523167
260-bp	0.036081	0.050702	0.487679671	0.181716	0.18904467
356-bp	0.040204	0.068967	0.855578371	0.1689	0.28341234
353-bp	0.395218	0.405893	0.205338809	1.418512	0.60624045
Rsp	0.042929	0.019702	0.106947296	0.111576	0.07028857
HETRP	0.852257	0.579957	4.372005476	2.388736	2.04823887
(AAGAG)n	0.000152	0	0	0.000031	0.00004575
42AB	9.303048	12.211887		27.706589	16.4071747
80F	1.444347	1.082895		2.449803	1.659015
miRNA (avarage)	5.9666E-05	5.96659E-06	5.96659E-05	0.0011098	0.00030877

#numbers show % of reads in RIP-seq data mapped to each satDNA or other repeat

Table S6. Log2 fold changes of small RNA abundance for satDNAs and piRNA clusters normalized to miRNA abundance (data from (Klattenhoff, et al. 2009; Pane, et al. 2011; Czech, et al. 2013; Le Thomas, et al. 2014; Mohn, et al. 2014; Andersen, et al. 2017; Parhad, et al. 2017))

#Protein	#Replicate	#Repeat	#Log2_fold_change
Cutoff	Cutoff_1	*Rsp*	-8.00107576384032
Cutoff	Cutoff_1	*359-bp*	-4.512234605431456
Cutoff	Cutoff_1	*353-bp*	-5.168600612208664
Cutoff	Cutoff_1	*356-bp*	-6.209575559376659
Cutoff	Cutoff_1	*260-bp*	-5.980064663451489
Cutoff	Cutoff_2	*Rsp*	-2.962524394418258
Cutoff	Cutoff_2	*359-bp*	-1.3315320466374843
Cutoff	Cutoff_2	*353-bp*	-1.6811952211914714
Cutoff	Cutoff_2	*356-bp*	-0.6459110515708952
Cutoff	Cutoff_2	*260-bp*	-1.4691026545250363
Cutoff	Cutoff_3	*Rsp*	-6.687187856081559
Cutoff	Cutoff_3	*359-bp*	-6.020437212044945
Cutoff	Cutoff_3	*353-bp*	-3.4596083914242635
Cutoff	Cutoff_3	*356-bp*	-4.620493296899111
Cutoff	Cutoff_3	*260-bp*	-1.7242887267858575
Deadlock	Deadlock_1	*Rsp*	-8.42656692958964
Deadlock	Deadlock_1	*359-bp*	-4.9132240441001915
Deadlock	Deadlock_1	*353-bp*	-5.560237409812577
Deadlock	Deadlock_1	*356-bp*	-7.185944848696442
Deadlock	Deadlock_1	*260-bp*	-6.368006762052888
Deadlock	Deadlock_2	*Rsp*	-4.365337258598803
Deadlock	Deadlock_2	*359-bp*	-1.9486804427799578
Deadlock	Deadlock_2	*353-bp*	-2.020787786854162
Deadlock	Deadlock_2	*356-bp*	-2.7036690165683934
Deadlock	Deadlock_2	*260-bp*	-1.943699748181694
Moonshiner	Moonshiner	*Rsp*	-4.830343390783226
Moonshiner	Moonshiner	*359-bp*	-2.716320370682399
Moonshiner	Moonshiner	*353-bp*	-3.0850179373994875
Moonshiner	Moonshiner	*356-bp*	-3.167924370314697
Moonshiner	Moonshiner	*260-bp*	-2.8643567719598013
Rhino	Rhino_1	*Rsp*	-8.285690813998857
Rhino	Rhino_1	*359-bp*	-4.63362100020403
Rhino	Rhino_1	*353-bp*	-5.359828570309815
Rhino	Rhino_1	*356-bp*	-6.871739214921723
Rhino	Rhino_1	*260-bp*	-6.043474633120904
Rhino	Rhino_2	*Rsp*	-5.658869862864648
Rhino	Rhino_2	*359-bp*	-4.999364680866015
Rhino	Rhino_2	*353-bp*	-5.565596995496925
Rhino	Rhino_2	*356-bp*	-6.629498797282487
Rhino	Rhino_2	*260-bp*	-6.097365042611733

Rhino	Rhino_3	*Rsp*	-9.007801950219708
Rhino	Rhino_3	*359-bp*	-5.940018651787192
Rhino	Rhino_3	*353-bp*	-6.896549873877682
Rhino	Rhino_3	*356-bp*	-5.109790817375375
Rhino	Rhino_3	*260-bp*	-7.691171573864062
Rhino	Rhino_4	*Rsp*	-3.9787578925365423
Rhino	Rhino_4	*359-bp*	-3.5208638110901234
Rhino	Rhino_4	*353-bp*	-4.011203407526927
Rhino	Rhino_4	*356-bp*	-3.9503127746354196
Rhino	Rhino_4	*260-bp*	-4.021575618887951
Cutoff	Cutoff_1	*20A*	0.7269206105610226
Cutoff	Cutoff_1	*flamenco*	0.014907817248020634
Cutoff	Cutoff_1	*42AB*	-8.42072637369969
Cutoff	Cutoff_1	*38C1*	-7.0624660647290645
Cutoff	Cutoff_1	*38C2*	-6.627519518102915
Cutoff	Cutoff_1	*80F*	-5.294822801473395
Cutoff	Cutoff_2	*20A*	-1.7006528082159598
Cutoff	Cutoff_2	*flamenco*	0.4559084104198677
Cutoff	Cutoff_2	*42AB*	-1.7419023147529307
Cutoff	Cutoff_2	*38C1*	-2.591279145841692
Cutoff	Cutoff_2	*38C2*	-1.4519089999281198
Cutoff	Cutoff_2	*80F*	-1.4773809443989123
Cutoff	Cutoff_3	*20A*	-0.6575134759039825
Cutoff	Cutoff_3	*flamenco*	-1.3089923978364073
Cutoff	Cutoff_3	*42AB*	-6.865561534314475
Cutoff	Cutoff_3	*38C1*	-6.774052959492507
Cutoff	Cutoff_3	*38C2*	-5.42547808196712
Cutoff	Cutoff_3	*80F*	-6.193362935664909
Deadlock	Deadlock_1	*20A*	0.5174603230041568
Deadlock	Deadlock_1	*flam*enco	0.1007440977469233
Deadlock	Deadlock_1	*42AB*	-7.5232208062969645
Deadlock	Deadlock_1	*38C1*	-8.091874542812112
Deadlock	Deadlock_1	*38C2*	-7.0818629463320715
Deadlock	Deadlock_1	*80F*	-6.2765196359020345
Deadlock	Deadlock_2	*20A*	0.7154801840088788
Deadlock	Deadlock_2	*flamenco*	0.5080837672212679
Deadlock	Deadlock_2	*42AB*	-3.2600076482724223
Deadlock	Deadlock_2	*38C1*	-2.2718781569017668
Deadlock	Deadlock_2	*38C2*	-1.0134751595268388
Deadlock	Deadlock_2	*80F*	-4.718569684098228
Moonshiner	Moonshiner	*20A*	0.47914343700052653
Moonshiner	Moonshiner	*flamenco*	0.345485943311131
Moonshiner	Moonshiner	*42AB*	-2.0437217345961995
Moonshiner	Moonshiner	*38C1*	3.4640286100796445
Moonshiner	Moonshiner	*38C2*	3.134364231151971

Moonshiner	Moonshiner	*80F*	-3.1495765849726705
Rhino	Rhino_1	*20A*	0.5207133848310788
Rhino	Rhino_1	*flamenco*	0.1258844576141151
Rhino	Rhino_1	*42AB*	-7.882269810888194
Rhino	Rhino_1	*38C1*	-7.108885027997908
Rhino	Rhino_1	*38C2*	-5.963714788857198
Rhino	Rhino_1	*80F*	-6.027868926998758
Rhino	Rhino_2	*20A*	0.5902807695512498
Rhino	Rhino_2	*flame*nco	-0.9111926648306102
Rhino	Rhino_2	*42AB*	-6.434634917166173
Rhino	Rhino_2	*38C1*	-5.503142761204401
Rhino	Rhino_2	*38C2*	-4.926067391106908
Rhino	Rhino_2	*80F*	-6.6321512623913925
Rhino	Rhino_3	*20A*	-0.9317842883091012
Rhino	Rhino_3	*flamenco*	-1.0947758707916255
Rhino	Rhino_3	*42AB*	-9.187048051763462
Rhino	Rhino_3	*38C1*	-8.822307483766116
Rhino	Rhino_3	*38C2*	-8.78632904351748
Rhino	Rhino_3	*80F*	-5.6775790562853405
Rhino	Rhino_4	*20A*	1.9321064055282942
Rhino	Rhino_4	*flamenco*	0.9297050063701647
Rhino	Rhino_4	*42AB*	-4.6529099870901245
Rhino	Rhino_4	*38C1*	-3.3324662845975035
Rhino	Rhino_4	*38C2*	-3.0834978651577902
Rhino	Rhino_4	*80F*	-4.692314029563647
Cutoff	Cutoff_1	HETRP_DM	-7.791632582453689
Cutoff	Cutoff_1	(AAGAG)n	0.9669231355751853
Cutoff	Cutoff_2	HETRP_DM	-1.0193513400563294
Cutoff	Cutoff_2	(AAGAG)n	2.0668356277883837
Cutoff	Cutoff_3	HETRP_DM	-8.265583837160406
Cutoff	Cutoff_3	(AAGAG)n	-0.8785308855004396
Deadlock	Deadlock_1	HETRP_DM	-7.725546749806859
Deadlock	Deadlock_1	(AAGAG)n	2.009281398273609
Deadlock	Deadlock_2	HETRP_DM	-4.61747239299394
Deadlock	Deadlock_2	(AAGAG)n	-6.629503796611773
Moonshiner	Moonshiner	HETRP_DM	-2.474386221122434
Moonshiner	Moonshiner	(AAGAG)n	0.17881079406221256
Rhino	Rhino_1	HETRP_DM	-7.572537883598184
Rhino	Rhino_1	(AAGAG)n	1.76900072027511
Rhino	Rhino_2	HETRP_DM	-7.487243060626127
Rhino	Rhino_2	(AAGAG)n	-0.11998105548403903
Rhino	Rhino_3	HETRP_DM	-8.05104262206187
Rhino	Rhino_3	(AAGAG)n	-0.02631398306916523
Rhino	Rhino_4	HETRP_DM	-4.909195985063195
Rhino	Rhino_4	(AAGAG)n	-0.4726039207689516

Table S7. Log2 fold changes of small RNA abundance for satDNAs and piRNA clusters normalized to the *flamenco* piRNA cluster (data from (Klattenhoff, et al. 2009; Pane, et al. 2011; Czech, et al. 2013; Le Thomas, et al. 2014; Mohn, et al. 2014; Andersen, et al. 2017; Parhad, et al. 2017))

#Protein	#Replicate	#Repeat	#Log2_fold_change
Cutoff	Cutoff_1	*Rsp*	-7.8122164002952905
Cutoff	Cutoff_1	*359-bp*	-4.322738608771259
Cutoff	Cutoff_1	*353-bp*	-4.9791081791045055
Cutoff	Cutoff_1	*356-bp*	-6.020206886713346
Cutoff	Cutoff_1	*260-bp*	-5.7906752012101785
Cutoff	Cutoff_2	*Rsp*	-3.0500056555060966
Cutoff	Cutoff_2	*359-bp*	-1.4188464391041287
Cutoff	Cutoff_2	*353-bp*	-1.7685202751151623
Cutoff	Cutoff_2	*356-bp*	-0.7332125112296148
Cutoff	Cutoff_2	*260-bp*	-1.5565112028873678
Cutoff	Cutoff_3	*Rsp*	-5.591737882249918
Cutoff	Cutoff_3	*359-bp*	-4.924118538087152
Cutoff	Cutoff_3	*353-bp*	-2.3632510332070473
Cutoff	Cutoff_3	*356-bp*	-3.5242721174675733
Cutoff	Cutoff_3	*260-bp*	-0.6279371573423145
Deadlock	Deadlock_1	*Rsp*	-8.367251177022622
Deadlock	Deadlock_1	*359-bp*	-4.853078200407767
Deadlock	Deadlock_1	*353-bp*	-5.500085370995209
Deadlock	Deadlock_1	*356-bp*	-7.126039044272359
Deadlock	Deadlock_1	*260-bp*	-6.3080028098836545
Deadlock	Deadlock_2	*Rsp*	-4.536128647800352
Deadlock	Deadlock_2	*359-bp*	-2.1193742098216144
Deadlock	Deadlock_2	*353-bp*	-2.191487368371124
Deadlock	Deadlock_2	*356-bp*	-2.874384950175744
Deadlock	Deadlock_2	*260-bp*	-2.11440352173192
Moonshiner	Moonshiner	*Rsp*	-4.932694644557066
Moonshiner	Moonshiner	*359-bp*	-3.00197331510481
Moonshiner	Moonshiner	*353-bp*	-3.3492079471043668
Moonshiner	Moonshiner	*356-bp*	-3.4505485122448643
Moonshiner	Moonshiner	*260-bp*	-3.115871819044168
Rhino	Rhino_1	*Rsp*	-8.191225401938219
Rhino	Rhino_1	*359-bp*	-4.538399156127625
Rhino	Rhino_1	*353-bp*	-5.264603650837064
Rhino	Rhino_1	*356-bp*	-6.7767130134620315
Rhino	Rhino_1	*260-bp*	-5.948366533872915
Rhino	Rhino_2	*Rsp*	-4.501855934146332
Rhino	Rhino_2	*359-bp*	-3.8214664762195505
Rhino	Rhino_2	*353-bp*	-4.372034886847178
Rhino	Rhino_2	*356-bp*	-5.45782952941583
Rhino	Rhino_2	*260-bp*	-4.884886755029063

Rhino	Rhino_3	*Rsp*	-7.9050999235196455
Rhino	Rhino_3	*359-bp*	-4.835499275148136
Rhino	Rhino_3	*353-bp*	-5.7920622174665635
Rhino	Rhino_3	*356-bp*	-4.00534456930642
Rhino	Rhino_3	*260-bp*	-6.587109314465846
Rhino	Rhino_4	*Rsp*	-4.952100310433019
Rhino	Rhino_4	*359-bp*	-4.494233332247884
Rhino	Rhino_4	*353-bp*	-4.984572310820784
Rhino	Rhino_4	*356-bp*	-4.9236817416682355
Rhino	Rhino_4	*260-bp*	-4.994936722686433
Cutoff	Cutoff_1	*20A*	0.7120055105883745
Cutoff	Cutoff_1	*flamenco*	1.442695747829688e-06
Cutoff	Cutoff_1	*42AB*	-8.4357099657614
Cutoff	Cutoff_1	*38C1*	-7.077599302605546
Cutoff	Cutoff_1	*38C2*	-6.642806990753876
Cutoff	Cutoff_1	*80F*	-5.309805975086021
Cutoff	Cutoff_2	*20A*	-2.1564119779703335
Cutoff	Cutoff_2	*flamenco*	7.213476935462569e-07
Cutoff	Cutoff_2	*42AB*	-2.1974187441438557
Cutoff	Cutoff_2	*38C1*	-3.0474786069947433
Cutoff	Cutoff_2	*38C2*	-1.9085541009914735
Cutoff	Cutoff_2	*80F*	-1.933320019869867
Cutoff	Cutoff_3	*20A*	0.6514785251119458
Cutoff	Cutoff_3	*flamenco*	0.0
Cutoff	Cutoff_3	*42AB*	-5.556629325232875
Cutoff	Cutoff_3	*38C1*	-5.465690143372289
Cutoff	Cutoff_3	*38C2*	-4.1168739399888095
Cutoff	Cutoff_3	*80F*	-4.884707882957589
Deadlock	Deadlock_1	*20A*	0.4167184475984708
Deadlock	Deadlock_1	*flamenco*	1.442695747829688e-06
Deadlock	Deadlock_1	*42AB*	-7.624001038730409
Deadlock	Deadlock_1	*38C1*	-8.193071010814773
Deadlock	Deadlock_1	*38C2*	-7.183126738634821
Deadlock	Deadlock_1	*80F*	-6.377400183570641
Deadlock	Deadlock_2	*20A*	0.2073919118190605
Deadlock	Deadlock_2	*flamenco*	0.0
Deadlock	Deadlock_2	*42AB*	-3.7680895768744005
Deadlock	Deadlock_2	*38C1*	-2.779969527272313
Deadlock	Deadlock_2	*38C2*	-1.5215655241817176
Deadlock	Deadlock_2	*80F*	-5.226703540860852
Moonshiner	Moonshiner	*20A*	0.11481810609189369
Moonshiner	Moonshiner	*flamenco*	-1.442695747863017e-06
Moonshiner	Moonshiner	*42AB*	-2.347125150442763
Moonshiner	Moonshiner	*38C1*	3.255364970790362
Moonshiner	Moonshiner	*38C2*	2.767379020687878

Moonshiner	Moonshiner	*80F*	-3.430119836986398
Rhino	Rhino_1	*20A*	0.3948308834592519
Rhino	Rhino_1	*flamenco*	0.0
Rhino	Rhino_1	*42AB*	-8.008200707817199
Rhino	Rhino_1	*38C1*	-7.235004187408406
Rhino	Rhino_1	*38C2*	-6.089838403327336
Rhino	Rhino_1	*80F*	-6.1538705712694375
Rhino	Rhino_2	*20A*	1.3800380392157199
Rhino	Rhino_2	*flamenco*	-1.442695747863017e-06
Rhino	Rhino_2	*42AB*	-5.589668670771764
Rhino	Rhino_2	*38C1*	-4.584892832458001
Rhino	Rhino_2	*38C2*	-4.145037904800569
Rhino	Rhino_2	*80F*	-5.785202935247456
Rhino	Rhino_3	*20A*	0.1629899654058821
Rhino	Rhino_3	*flamenco*	0.0
Rhino	Rhino_3	*42AB*	-8.092425049719873
Rhino	Rhino_3	*38C1*	-7.728184384848781
Rhino	Rhino_3	*38C2*	-7.692347614205799
Rhino	Rhino_3	*80F*	-4.583022580649129
Rhino	Rhino_4	*20A*	1.0024098181600634
Rhino	Rhino_4	*flamenco*	7.213493165986108e-06
Rhino	Rhino_4	*42AB*	-5.58260823083062
Rhino	Rhino_4	*38C1*	-4.262155785745151
Rhino	Rhino_4	*38C2*	-4.013189454278323
Rhino	Rhino_4	*80F*	-5.622005309232196
Cutoff	Cutoff_1	HETRP_DM	-7.6021653082639435
Cutoff	Cutoff_1	(AAGAG)n	1.1571794796445292
Cutoff	Cutoff_2	HETRP_DM	-1.1067738248972403
Cutoff	Cutoff_2	(AAGAG)n	1.9831017297734967
Cutoff	Cutoff_3	HETRP_DM	-7.1693121423474855
Cutoff	Cutoff_3	(AAGAG)n	0.21636437275093123
Deadlock	Deadlock_1	HETRP_DM	-7.665414185945266
Deadlock	Deadlock_1	(AAGAG)n	2.070628187982324
Deadlock	Deadlock_2	HETRP_DM	-4.788169975710487
Deadlock	Deadlock_2	(AAGAG)n	-10.372593033140745
Moonshiner	Moonshiner	HETRP_DM	-2.767545404292368
Moonshiner	Moonshiner	(AAGAG)n	0.12918803431931078
Rhino	Rhino_1	HETRP_DM	-7.4773310166223075
Rhino	Rhino_1	(AAGAG)n	1.8653447934059315
Rhino	Rhino_2	HETRP_DM	-6.3545552703453545
Rhino	Rhino_2	(AAGAG)n	1.3010063093403985
Rhino	Rhino_3	HETRP_DM	-6.946591564789813
Rhino	Rhino_3	(AAGAG)n	1.0780407381855284
Rhino	Rhino_4	HETRP_DM	-5.882563113846568
Rhino	Rhino_4	(AAGAG)n	-1.4459693018706132

Table S8. Log2 fold change of total RNA abundance for satDNAs and piRNA clusters (data from (Mohn, et al. 2014; Andersen, et al. 2017))

protein	repeatID	baseMean	log2FoldChange	lfcSE	pvalue	padj
Cutoff	*Rsp*	150.987264	-0.987310703	0.87494971	0.25914261	0.99949378
Deadlock	*Rsp*	104.421271	-1.320844886	0.91965724	0.15093518	0.99859633
Moonshiner	*Rsp*	94.4712611	-0.45233908	0.25871017	0.08038796	0.500139
Rhino	*Rsp*	136.432816	0.026681688	0.46998175	0.95472702	0.99909438
Cutoff	*356-bp*	134.199408	-1.483961187	0.8254925	0.07222969	0.99949378
Cutoff	*260-bp*	85.1432407	-0.771749272	0.8249678	0.34953574	0.99949378
Cutoff	*353-bp*	791.471319	0.311966817	0.67910052	0.6459596	0.99949378
Cutoff	*359-bp*	983.674787	-0.242730481	0.66939167	0.7168936	0.99949378
Cutoff	*360-bp*	418.739012	-0.035256767	0.69796595	0.95971312	0.99949378
Deadlock	*356-bp*	96.3189339	-1.543325643	0.87379615	0.07735714	0.99859633
Deadlock	*353-bp*	666.458809	0.82832386	0.68506796	0.22661984	0.99859633
Deadlock	*360-bp*	330.332279	0.319204424	0.7090205	0.65256279	0.99859633
Deadlock	*260-bp*	66.280278	-0.349631522	0.86554107	0.68625271	0.99859633
Deadlock	*359-bp*	754.274332	0.025077264	0.67696282	0.9704501	0.99859633
Moonshiner	*260-bp*	113.277268	-1.212597212	0.24194929	5.39E-07	1.04E-05
Moonshiner	*356-bp*	114.589469	-1.100593902	0.24946713	1.03E-05	0.00016956
Moonshiner	*353-bp*	724.004047	-0.43021101	0.11645277	0.0002205	0.00310635
Moonshiner	*359-bp*	1154.63428	-0.09541183	0.10000988	0.34007145	0.99914884
Moonshiner	*360-bp*	321.958738	0.083570622	0.15755908	0.59582898	0.99914884
Rhino	*356-bp*	109.876767	-1.473761577	0.35692421	3.64E-05	0.00078093
Rhino	*260-bp*	104.465366	-0.748862531	0.29967581	0.01245764	0.15530519
Rhino	*359-bp*	1001.7942	-0.22295042	0.21534307	0.30051638	0.99909438
Rhino	*360-bp*	340.229941	0.106519054	0.2522788	0.67285895	0.99909438
Rhino	*353-bp*	721.869348	-0.063994885	0.2661101	0.80995604	0.99909438
Cutoff	*38C2*	520.67409	-1.786200828	0.64852894	0.00588296	0.32141689
Cutoff	*38C1*	372.177442	-1.659978493	0.64666839	0.01025914	0.51438865
Cutoff	*42AB*	6215.62259	-0.752318038	0.63928244	0.23926881	0.99993393
Cutoff	*20A*	21556.3689	0.710747643	0.63388497	0.2621787	0.99993393
Cutoff	*80F*	8897.4974	0.422125053	0.64231907	0.51105937	0.99993393
Cutoff	*flamenco*	1971.03137	-0.01202246	0.58959385	0.98373139	0.99993393
Deadlock	*38C1*	275.989027	-1.405447559	0.65653685	0.03229835	0.98170676
Deadlock	*38C2*	391.452368	-1.38170335	0.65142057	0.0339166	0.98265044
Deadlock	*flamenco*	1652.8816	0.555418667	0.58184917	0.33979267	0.99850437
Deadlock	*42AB*	5846.75928	0.464278102	0.62897777	0.46042508	0.99850437
Deadlock	*20A*	14092.8244	0.371094367	0.6238851	0.55196912	0.99850437
Deadlock	*80F*	6269.81813	0.334042145	0.63268737	0.59751775	0.99850437
Moonshiner	*42AB*	4070.5203	-0.827022404	0.0407482	1.40E-91	6.42E-89
Moonshiner	*38C2*	874.010244	1.549946377	0.08985218	1.12E-66	2.57E-64
Moonshiner	*38C1*	448.265537	1.669922474	0.12585974	3.54E-40	4.65E-38

Moonshiner	*20A*	9184.66043	0.196953578	0.02863194	6.04E-12	2.82E-10
Moonshiner	*80F*	3004.54407	-0.308161266	0.04589612	1.89E-11	8.40E-10
Moonshiner	*flamenco*	11330.2141	-0.094635978	0.02607315	0.00028382	0.00488448
Rhino	*38C1*	210.296383	-1.467834345	0.2625671	2.27E-08	1.55E-06
Rhino	*38C2*	371.385374	-1.244826985	0.26373771	2.36E-06	0.0001234
Rhino	*42AB*	5087.68089	-0.300003619	0.25813685	0.245159	0.99986008
Rhino	*20A*	11945.7616	0.26830092	0.25704457	0.2965819	0.99986008
Rhino	*flamenco*	7934.23718	0.100869685	0.19366718	0.60247799	0.99986008
Rhino	*80F*	4483.92589	-0.109049761	0.26502443	0.68072749	0.99986008

Table S9. *Rsp* probe sequences

ATCGGAAATTCACCCATTTT
TGAAATCGGAAAATCACCAA
AATCGGAAAATCACCCATTT
TGAAATCGGAAATTCACCCA
CAGAATCTCTAGAAGGTGTC
CAGTTGAACAGAATCTCTAG
GTTGAACAGAATCTCTAGAA
TGCGTACCAGTTGAACAGAA
ATCGGAAAATTACCCATTTT
AATCGGAAAATCACCAAATT
CCAGGTTTAAATAATCTGTA
TGTTTTTGTGTACCAGGTTT
CTTAGTCAATTTATTGTTTT
TATCTTCTGTTTGTCTGGTA
CAGGTTTAAATAATCTGTAG
TCTGTAGAAGGTATCTTCTG
TGAAATCGGAAATTCAACCA
TCGAAAAATCACCCATTTTG
TGAAATCGGAAAATCATCAA
TCTAGAAGGTGTCTTCTGTT
TCGATAAATCACCCATTTTG
AAACTTAGGCAATTTACTGT
GTCAATTTATTGTGTTTGTG
AATCGGAAAATCCCCCATTT
TGAAATCGGAACATCACCCA
ATCGGAAAATCACCAACTTT
TCGGCAAATCACCAATTTTG
TGTCTGGTACTTGAAATCGG
TGCTTACCAGTTGAACAGAA
TGAAATCGGAAAATGACCCA
CGCCCATTACTTGAAATCGG
AAGTTGTGTCATTTTAAGCG
TCGCCTGGTACTTGAAATCG
TCGTCTGGTACTTGAAATCG
TTAAATAATCTGTAGAAGGT
TAGAAGGTGTCTTCTGTTCG
AGGTATCTTCTGTTTGTCTG
TTTGTCTGGTACTTGAAATC
TTGCGTACCAGTTGAACAGA
CTTACCAGTTGAACAGAATC
TCGCCTATTACTTGAAATCG
GGTATCTTCTGTTCGCCTGG
GAAGGTGTCTTCTGTTCGTC

GTCTATTACTTGAAATCGGA
TGTGTACCAGGTTTAAATAA
TGTTTTTGCTTACCAGTTGA
CATTTTGACCGCTTAAAATG
GGTACTTGAAATCGAAAAAT

References

Akkouche A, Mugat B, Barckmann B, Varela-Chavez C, Li B, Raffel R, Pelisson A, Chambeyron S. 2017. Piwi Is Required during Drosophila Embryogenesis to License Dual-Strand piRNA Clusters for Transposon Repression in Adult Ovaries. Mol Cell 66:411-419 e414.

Andersen PR, Tirian L, Vunjak M, Brennecke J. 2017. A heterochromatin-dependent transcription machinery drives piRNA expression. Nature 549:54-59.

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128:1089-1103.

Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, Booth BW, Wen J, Park S, Suzuki AM, et al. 2014. Diversity and dynamics of the Drosophila transcriptome. Nature 512:393-399.

Czech B, Preall JB, McGinn J, Hannon GJ. 2013. A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway. Mol Cell 50:749-761.

Fagegaltier D, Konig A, Gordon A, Lai EC, Gingeras TR, Hannon GJ, Shcherbata HR. 2014. A genome-wide survey of sexually dimorphic expression of Drosophila miRNAs identifies the steroid hormone-induced miRNA let-7 as a regulator of sexual identity. Genetics 198:647-668.

Ghildiyal M, Xu J, Seitz H, Weng Z, Zamore PD. 2010. Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway. RNA 16:43-56.

Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, et al. 2011. The developmental transcriptome of Drosophila melanogaster. Nature 471:473-479.

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

Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A, et al. 2009. The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell 138:1137-1149.

Klenov MS, Lavrov SA, Korbut AP, Stolyarenko AD, Yakushev EY, Reuter M, Pillai RS, Gvozdev VA. 2014. Impact of nuclear Piwi elimination on chromatin state in Drosophila melanogaster ovaries. Nucleic Acids Res 42:6208-6218.

Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Toth KF. 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev 27:390-399.

Le Thomas A, Stuwe E, Li S, Du J, Marinov G, Rozhkov N, Chen YC, Luo Y, Sachidanandam R, Toth KF, et al. 2014. Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. Genes Dev 28:1667-1680.

Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ. 2009. Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell 137:522-535.

Mohn F, Handler D, Brennecke J. 2015. piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. Science 348:812-817.

Mohn F, Sienski G, Handler D, Brennecke J. 2014. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila. Cell 157:1364-1379.

Olivieri D, Senti KA, Subramanian S, Sachidanandam R, Brennecke J. 2012. The cochaperone shutdown defines a group of biogenesis factors essential for all piRNA populations in Drosophila. Mol Cell 47:954-969.

Pane A, Jiang P, Zhao DY, Singh M, Schupbach T. 2011. The Cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline. EMBO J 30:4601-4615.

Parhad SS, Tu S, Weng Z, Theurkauf WE. 2017. Adaptive Evolution Leads to Cross-Species Incompatibility in the piRNA Transposon Silencing Machinery. Dev Cell 43:60-70 e65.

Preall JB, Czech B, Guzzardo PM, Muerdter F, Hannon GJ. 2012. shutdown is a component of the Drosophila piRNA biogenesis machinery. RNA 18:1446-1457.

Quenerch'du E, Anand A, Kai T. 2016. The piRNA pathway is developmentally regulated during spermatogenesis in Drosophila. RNA 22:1044-1054.

Rozhkov NV, Aravin AA, Zelentsova ES, Schostak NG, Sachidanandam R, McCombie WR, Hannon GJ, Evgen'ev MB. 2010. Small RNA-based silencing strategies for transposons in the process of invading Drosophila species. RNA 16:1634-1645.

Sato K, Iwasaki YW, Shibuya A, Carninci P, Tsuchizawa Y, Ishizu H, Siomi MC, Siomi H. 2015. Krimper Enforces an Antisense Bias on piRNA Pools by Binding AGO3 in the Drosophila Germline. Mol Cell 59:553-563.

Wang W, Han BW, Tipping C, Ge DT, Zhang Z, Weng Z, Zamore PD. 2015. Slicing and Binding by Ago3 or Aub Trigger Piwi-Bound piRNA Production by Distinct Mechanisms. Mol Cell 59:819-830.

Zhang F, Wang J, Xu J, Zhang Z, Koppetsch BS, Schultz N, Vreven T, Meignin C, Davis I, Zamore PD, et al. 2012. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. Cell 151:871-884.

Zhang Z, Wang J, Schultz N, Zhang F, Parhad SS, Tu S, Vreven T, Zamore PD, Weng Z, Theurkauf WE. 2014. The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. Cell 157:1353-1363.