

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

Xiaolu Wei<sup>1\*</sup>, Danna G Eickbush<sup>2</sup>, Iain Speece<sup>2</sup>, Amanda M Larracuenté<sup>2\*</sup>

<sup>1</sup>Department of Biomedical Genetics, University of Rochester Medical Center, Rochester, United States; <sup>2</sup>Department of Biology, University of Rochester, Rochester, United States

**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.

**\*For correspondence:**

xiaolu\_wei@urmc.rochester.edu (XW);  
alarracu@bio.rochester.edu (AML)

**Competing interests:** The authors declare that no competing interests exist.

**Funding:** See page 17

**Received:** 22 August 2020

**Accepted:** 08 July 2021

**Published:** 13 July 2021

**Reviewing editor:** Detlef Weigel, Max Planck Institute for Developmental Biology, Germany

© Copyright Wei et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## 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*). The de-repression of satDNA is associated with cellular senescence and various cancers (e.g., *Ting et al., 2011; Zhu et al., 2011*). However, satDNAs play roles in chromatin structure, chromosome segregation, and genome stability across a wide range of taxa (*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; Rošić 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*). Across organisms, satDNA-derived transcripts may generally be important for maintaining genome stability and integrity, yet the regulation and function of these transcripts remains poorly understood (reviewed in *Janssen et al., 2018*).

Insights might come from the small RNA pathways that protect genome integrity by silencing repeats. These RNA interference pathways play roles in heterochromatin formation and maintenance at repeats across species (*Hall et al., 2002; Volpe et al., 2002; Fukagawa et al., 2004;*

*Noma et al., 2004; Verdel et al., 2004; Novo et al., 2020*). 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) that target transposable elements (TEs)—genomic parasites that mobilize and can cause genome instability (*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 containing primarily truncated TE sequences, called piRNA clusters. The piRNAs derived from these loci repress TE activity through both post-transcriptional (*Brennecke et al., 2007; Gunawardane et al., 2007*) and transcriptional silencing. In ovaries, piRNAs guide Piwi to genomic locations with complementary nascent RNAs and recruit heterochromatin factors to silence TEs (*Wang and Elgin, 2011; Sienski et al., 2012; Le Thomas et al., 2013; Rozhkov et al., 2013*).

There are two main 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, small RNAs derived from satDNAs exist in germlines (e.g., *Aravin et al., 2003; Saito et al., 2006*). 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 (1–10 bp) or complex (>100 bp). We focus on two abundant families of complex satDNA in *Drosophila melanogaster*: *Responder (Rsp)* and satellites in the *1.688 g/cm<sup>3</sup>* family (*1.688*). We show that complex satDNAs are expressed and processed primarily into piRNAs in both testes and ovaries. In ovaries, this expression depends on the RDC complex and Moon. 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 findings add insight into the role of piRNA pathways in heterochromatin formation and genome stability.

## Results and discussion

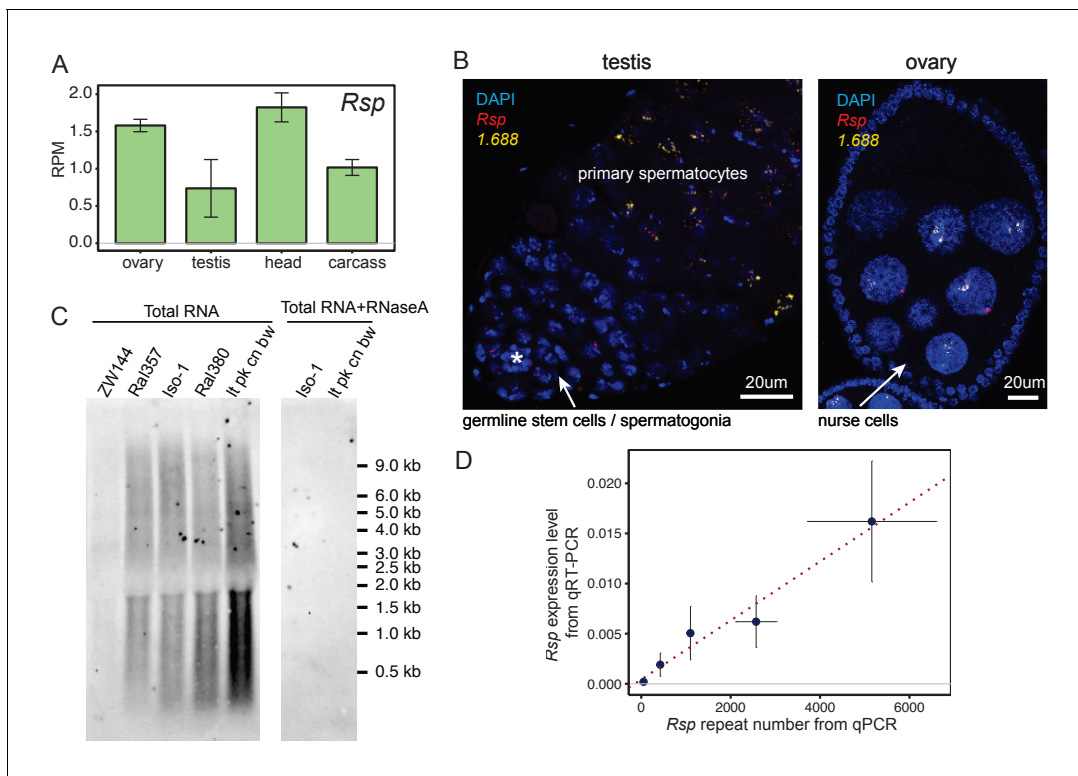
### SatDNA transcripts originate primarily from large heterochromatic satDNA blocks

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 is the most abundant complex satDNA in *D. melanogaster* (*Lohe and Roberts, 1988*). It comprises different subfamilies that exist as discrete tandem arrays in the pericentric

heterochromatin named after their repeat unit sizes on chromosome 2L (260-bp), chromosome 3L (353-bp and 356-bp), and the X chromosome (359-bp) (Losada and Villasante, 1996; Abad et al., 2000). Because there is high sequence similarity among these repeats, we analyzed all 1.688 subfamilies together unless stated otherwise.

We mined modENCODE datasets (Supplementary file 1 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, Figure 1—figure supplement 1). Both satDNA families are expressed in gonads, head, and other tissues (Figure 1A, Figure 1—figure supplement 1C). Their transcript abundance is low ( $\text{RPM}_{Rsp} < 10$  and  $\text{RPM}_{1.688} < 300$ ; Supplementary file 2) and generally increases throughout development and with adult age (Figure 1—figure supplement 1A, B). SatDNA-derived reads have very low abundance in the poly-A selected RNA-seq data ( $\text{RPM}_{Rsp} < 0.2$  and  $\text{RPM}_{1.688} < 10$ ; Supplementary file 2), indicating that the majority of satDNA transcripts are not polyadenylated.

To validate the presence of satDNA-derived transcripts in gonads, we used RNA fluorescence in situ hybridization (FISH). Both *Rsp* and 1.688 satellite transcripts are visible in testes and ovaries (Figure 1B, Figure 1—figure supplement 2A). These signals are undetectable after treating with



**Figure 1.** Satellite DNAs (satDNAs) are expressed in ovaries and testes. (A) *Rsp* satDNA transcription level in various tissues (corresponding result for 1.688 is shown in Figure 1—figure supplement 1C). Carcass: whole body without the head, reproductive organs, and digestive tract. Source data in Figure 1—source data 1. (B) RNA fluorescence in situ hybridization shows evidence for *Rsp* and 1.688-derived transcripts in testes and ovaries; asterisk indicates the hub. The probe for 1.688 recognizes all 1.688 subfamilies except for 260-bp on chromosome 2L. (C) Northern blot probed with *Rsp*. Total RNA was extracted from ovaries of fly lines with varying copy numbers of *Rsp*: ZW144 (200 copies), Ral357 (600 copies), Iso-1 (1100 copies), Ral380 (2300), and *It pk cn bw* (4100). There is no signal after RNaseA treatment. Signal quantification (shown in Figure 1—figure supplement 1D) 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 Supplementary file 4.

The online version of this article includes the following source data and figure supplement(s) for figure 1:

**Source data 1.** Satellite DNAs (SatDNA) transcription level in various developmental stages and tissues.

**Figure supplement 1.** Satellite DNAs (satDNAs) are expressed across tissues and developmental stages.

**Figure supplement 2.** RNA fluorescence in situ hybridization (FISH) signals of satellite DNAs (satDNAs) are from RNAs, not DNAs.

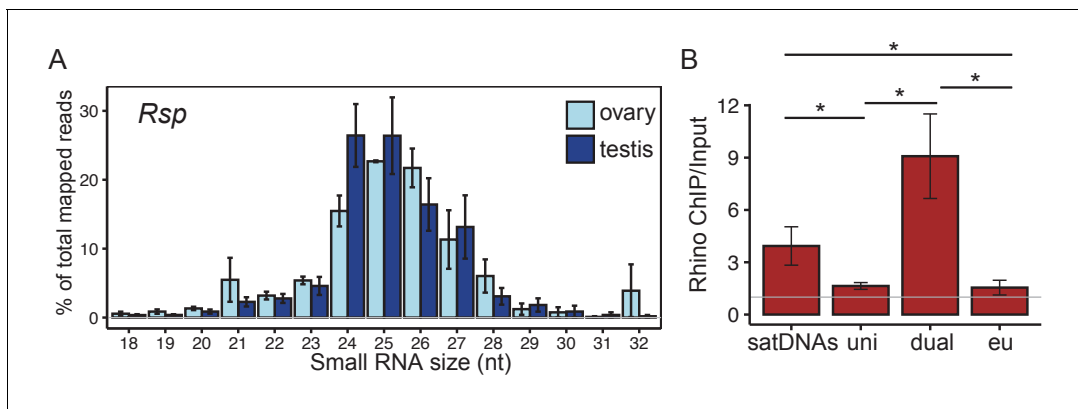
RNaseA prior to probe hybridization (**Figure 1—figure supplement 2B**), which degrades single-stranded RNAs, or RNaseH post-probe hybridization (**Figure 1—figure supplement 2C**), which degrades the RNA in DNA-RNA hybrids. This suggests that these signals are from RNA rather than DNA. We detected satDNA transcript foci in ovarian nurse cells and in pre-meiotic testicular germ cells. Interestingly, in testes we detected *Rsp* signal at earlier stages of spermatogenesis (i.e., germline stem cells/spermatogonia) than the 1.688 signals (i.e., primary spermatocytes; **Figure 1B**). The difference in timing is notable as *Rsp* is the specific target of *Segregation Distorter* (*SD*; **Sandler et al., 1959**): a well-known male meiotic drive system 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.

The bulk of satDNAs are found in large blocks of tandem repeats in the heterochromatin with small blocks occurring in the euchromatin (**Waring and Pollack, 1987; DiBartolomeis et al., 1992; Kuhn et al., 2012; Sproul et al., 2020**). Some of the euchromatic (**Menon et al., 2014; Joshi and Meller, 2017; Deshpande and Meller, 2018**) and heterochromatic loci in the 1.688 family (**Usakin et al., 2007**) are transcribed. To determine if satDNA-derived transcripts originating from large heterochromatic loci is a general feature of other complex satDNAs, we examined transcript size and abundance in total RNA from ovaries of flies that vary in *Rsp* repeat copy number (**Supplementary file 3; Khost et al., 2017**). We determined that, while transcript lengths were similar among these lines—ranging between <300 nt and >9000 nt (**Figure 1C**)—the abundance of *Rsp* transcripts correlated with genomic copy number (**Figure 1—figure supplement 1D** and **Supplementary file 4**, Pearson's correlation coefficient  $r^2 = 0.93$ , p-value=0.02). We validated these hybridization results using qPCR and qRT-PCR to quantify *Rsp* genomic DNA and RNA transcript abundance, respectively (**Figure 1D** and **Supplementary file 4**, Pearson's correlation coefficient  $r^2 = 0.98$ , p-value=0.003). The correlation between genomic copy number and transcript abundance is consistent with most transcripts originating from the large blocks of heterochromatic satDNA.

### SatDNA transcripts are processed into piRNAs in *Drosophila* germline

Many different repeat-derived transcripts are processed into piRNAs (**Aravin et al., 2003; Saito et al., 2006; Brennecke et al., 2007**) and endo-siRNAs (**Czech et al., 2008; Ghildiyal et al., 2008; Okamura et al., 2008; Menon et al., 2014**). To ask if complex satDNA-derived RNAs are processed into small RNAs, we reanalyzed published small RNA-seq data (**Supplementary file 1; Ghildiyal et al., 2010; Rozhkov et al., 2010; Fagegaltier et al., 2014; Mohn et al., 2014; Quénerch' 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—figure supplement 1A, B**). Our results suggest that the majority of these satDNA-derived small RNAs are piRNAs. First, these small RNAs are abundant in testes and ovaries, and their size distribution is typical for piRNA populations: an average of 90% of the RNAs range from 23 nt to 28 nt, with a peak at 24–26 nt in *D. melanogaster* (**Brennecke et al., 2007; Figure 2A** for *Rsp* and **Figure 2—figure supplement 1C** for 1.688). Second, the satDNA-derived small RNAs bear a signature of the piRNA-guided RNA cleavage process called the ping-pong cycle. piRNAs amplified through ping-pong have a 10 nt 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—figure supplement 1E, G**) and ~60–80% have 1U/10A (**Figure 2—figure supplement 1D, F**). Third, satDNA-derived small RNAs are bound by the PIWI proteins, as expected for piRNAs. Our reanalysis of published Piwi, Aubergine (Aub), and Argonaute3 (Ago3) RIP-seq data from ovaries (**Brennecke et al., 2007; Mohn et al., 2015; Sato et al., 2015**) shows that *Rsp* and 1.688 RNAs interact with each of these proteins (**Supplementary file 5**). For example, ~0.9% and 0.1% of Piwi-bound RNAs map to 1.688 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 (**Supplementary file 5**; e.g., 3.1% and 0.1% of Aub-bound RNAs map to 1.688 and *Rsp*, respectively; and 1.8% and 0.07% of Ago3-bound RNAs map to 1.688 and *Rsp*, respectively).





**Figure 2.** Satellite DNAs (satDNAs) 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—figure supplement 1C**). Source data in **Figure 2—source data 1**. (B) Rhino ChIP-seq result from ovaries showing the enrichment scores for satDNAs, uni-strand (uni) piRNA clusters, dual-strand (dual), and euchromatin (eu). The enrichment scores for each satDNA and piRNA cluster are shown in **Figure 2—figure supplement 5A**. p-values are estimated by pairwise t-tests with FDR correction (**Benjamini and Hochberg, 1995**). \* adjusted p-value < 0.05. Source data in **Figure 2—source data 2**.

The online version of this article includes the following source data and figure supplement(s) for figure 2:

**Source data 1.** Size distribution of small RNAs for *Rsp* and 1.688 in testes and ovaries.

**Source data 2.** Rhino/H3K9me3 ChIP-seq enrichment scores for *Rsp*, 1.688 heterochromatic loci, piRNA clusters, and euchromatin.

**Source data 3.** Rhino ChIP-seq enrichment scores for all repeats in the genome.

**Figure supplement 1.** Satellite-derived RNAs are mainly processed into piRNAs in the germline.

**Figure supplement 1—source data 1.** Small RNA levels (RPM) in ovary, testis, head, and body (minus head) for 1.688 and *Rsp*.

**Figure supplement 2.** Non-uniform distribution of piRNA reads along satellite DNA (satDNA) consensus sequences.

**Figure supplement 3.** Non-uniform distribution of piRNA reads along the germline-dominant transposable element (TE) consensus sequences.

**Figure supplement 4.** Satellite DNAs (satDNAs) are transcribed from both strands.

**Figure supplement 4—source data 1.** Percentages of reads mapped to plus and minus strands at all genomic copies of *Rsp* or 1.688 satellites.

**Figure supplement 5.** ChIP-seq result shows the chromatin state of satellite DNAs (satDNAs), uni-strand piRNA clusters, dual-strand piRNA clusters, and euchromatin (eu).

Together, these results indicate that satDNA-derived transcripts are processed into piRNAs in the female germline.

We examined the piRNA distribution along individual repeat units for *Rsp* and two subfamilies of 1.688 (359-bp and 260-bp) by blasting the corresponding sequencing reads to each consensus sequence. We find that the distribution of piRNA read depth is not uniform along the length of single monomers (359-bp and 260-bp) or dimers (*Rsp*) of these satDNA repeats (**Figure 2—figure supplement 2A, B**). This pattern could arise if these piRNAs derive from repeat fragments overrepresented in the genome. However, when we look at the alignment depth of all genomic repeat variants, it is more uniform across the monomers/dimers than the piRNA pileup (**Figure 2—figure supplement 2C**). We observe similar non-uniform patterns of piRNA pileup over germline-dominant TEs (e.g., *invader6*, *mdg3*, and *Het-A*; **Figure 2—figure supplement 3**), suggesting that these uneven distributions may arise from piRNA processing. The piRNA read pileup pattern also differs between ovaries and testes (**Figure 2—figure supplement 2A, B**), suggesting that there may be differences in transcription machinery, precursor production, or precursor processing between these tissues.

### SatDNA transcription resembles dual-strand piRNA clusters

*D. melanogaster* ovarian piRNAs originate primarily from uni- or dual-strand piRNA clusters. To determine which pathway controls the expression of satDNA-derived piRNA precursors, we asked whether transcripts come from one or both strands. We mapped total RNAseq reads from ovary and testis to the genome assembly. Collectively, for all genomic copies of *Rsp* or 1.688 satDNA (all subfamilies), we find a nearly 1:1 ratio of reads mapping to the plus and minus strands (**Figure 2—figure supplement 4A**; all mapped and uniquely mapped reads). However, the highly repetitive nature of satDNAs makes confidently assigning satellite-derived reads to a genomic location difficult. We

therefore take advantage of our assemblies for two representative satDNA loci: the major *Rsp* locus on chromosome 2R and the 260-bp locus, a subfamily of 1.688, on chromosome 2L (Klattenhoff et al., 2017). For these two loci, we confirm that reads map uniquely to both strands of the contigs (Figure 2—figure supplement 4B). Together, these results suggest that satDNAs are transcribed from both strands, similar to dual-strand piRNA clusters.

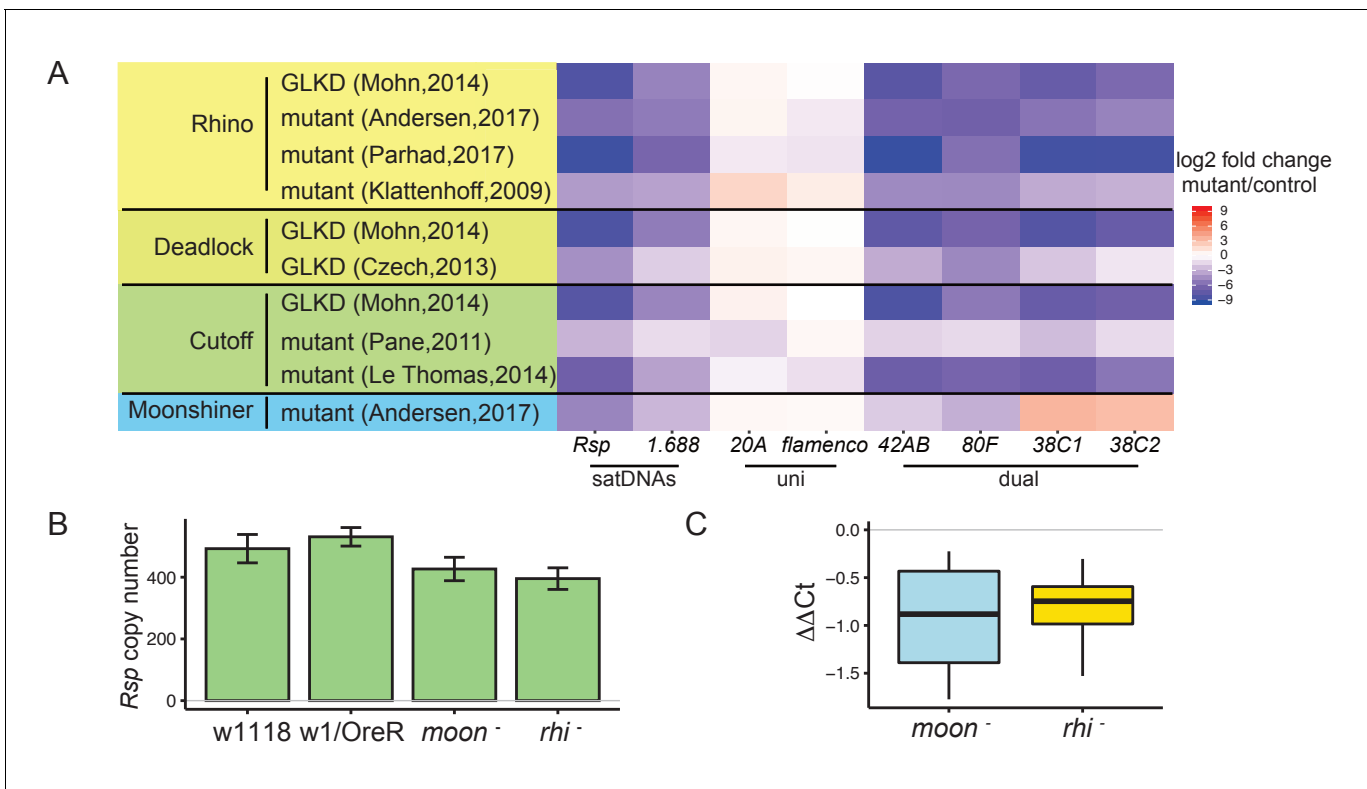
Dual-strand piRNA clusters are associated with the heterochromatin binding protein Rhi (Klattenhoff et al., 2009; Zhang et al., 2014). We therefore reanalyzed publicly available ChIP-seq datasets from ovaries (Mohn et al., 2014; Zhang et al., 2014; Parhad et al., 2017) to determine if satDNA regions are also Rhi-associated. Our results for piRNA clusters are consistent with previous studies (Klattenhoff et al., 2009; Mohn et al., 2014; Andersen et al., 2017): the dual-strand piRNA clusters have higher Rhi enrichment (mean enrichment ChIP/Input  $E_{\text{dual}} = 9.08$ ) compared to uni-strand piRNA clusters ( $E_{\text{uni}} = 1.69$ ; pairwise t-test with Benjamini–Hochberg; Benjamini and Hochberg, 1995 adjusted p-value  $P_{\text{adj}}=0.01$ ) and euchromatic genes ( $E_{\text{euch}} = 1.55$ ;  $P_{\text{adj}}=0.01$ ). We found that complex satDNAs are in the top 30% of all repeats enriched in Rhi (full data in Figure 2—source data 3). The level of Rhi enrichment for satDNAs ( $E_{\text{sat}} = 4.70$ ) is intermediate between the highly enriched dual-strand piRNA clusters ( $P_{\text{adj}}=0.1$ ) and the minimally Rhi enriched uni-strand piRNA clusters ( $P_{\text{adj}}=0.01$ ) or euchromatin ( $P_{\text{adj}}=0.01$ ) (Figure 2B and Figure 2—figure supplement 5A). Unlike the uneven distribution of piRNAs along satellite monomers/dimers (Figure 2—figure supplement 2A, B), the distribution of Rhi ChIP-seq reads (Figure 2—figure supplement 2D) is similar to the alignment depth of genomic repeats (Figure 2—figure supplement 2C). 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—figure supplement 2C) or in potentially unannotated piRNA clusters.

### SatDNA transcription is regulated by RDC complex and Moon

Because we find that satDNAs generate piRNAs in the female germline and their chromatin is associated with Rhi, we asked if the same transcription and RNA processing machinery are used by both satDNAs and 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—figure supplement 1), 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 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). We find that the pools of complex satDNA-derived piRNAs are also reduced in RDC and Moon mutants (Figure 3A, Figure 3—figure supplement 1). In *rhi* mutants, *Rsp* piRNA abundance is 0.2–6.3% their levels in wild-type datasets. Similarly, piRNA abundance for 1.688 is 1.4–7.8% their levels in wild-type datasets (complete list of log<sub>2</sub> fold change for satDNAs in Supplementary files 6 and 7). The reduction in satDNA piRNA level is robust to normalization method (miRNA in Figure 3A; flamenco cluster in Figure 3—figure supplement 1). While the expression of simple satellite repeats like AAGAG was not decreased in these mutants (Supplementary file 6 and Supplementary file 7), the low abundance of AAGAG reads (the number of reads mapping to AAGAG are only ~0.5% of *Rsp* and ~0.03% of 1.688) and known sources of bias for simple repeats (e.g., PCR bias in RNA-seq library preparation; 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 reanalyzed 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 the dual-strand piRNA clusters 42AB and 80F (Mohn et al., 2014; Andersen et al., 2017). Consistent with published reports, we detected decreases in steady-state long RNA transcript levels at dual-strand piRNA clusters (Figure 3—figure supplement 2). However, we did not observe a significant decrease in steady-state



**Figure 3.** Satellite DNA (satDNA) loci are regulated by the heterochromatin-dependent transcription machinery in *Drosophila* ovaries. (A) Heatmap showing the quantification of changes in piRNA abundance in small RNA-seq data from mutants of *rhino*, *cutoff*, *deadlock*, and *moonshiner* compared to controls for satDNAs and piRNA clusters, normalized by miRNA level. GLKD: germline knockdown. Complete list of log2 fold changes in **Supplementary file 6**. (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.  $\Delta\Delta Ct = \Delta Ct(\text{wild type}) - \Delta Ct(\text{mutant})$ , a negative value indicates lower expression in mutant. Student's t-test, p-value=0.077, 0.048. Source data in **Figure 3—source data 1**.

The online version of this article includes the following source data and figure supplement(s) for figure 3:

**Source data 1.** *Rsp* copy number and expression level estimated from qPCR and q-RT-PCR.

**Figure supplement 1.** Satellite DNA (satDNA) loci are regulated by the heterochromatin-dependent transcription machinery in *Drosophila* ovaries.

**Figure supplement 2.** *Rhino*, *Deadlock*, and *Cutoff* (RDC) and *moon* mutants affect piRNA precursor transcription at piRNA clusters.

**Figure supplement 2—source data 1.** Log2 fold changes of total RNA abundance for piRNA clusters from *Rhino*, *Deadlock*, and *Cutoff* (RDC) and *Moon* mutants.

**Figure supplement 3.** Satellite DNA (satDNA) piRNA production is affected in mutants of pathways involving piRNA precursor export, primary piRNA biogenesis, and the ping-pong cycle.

**Figure supplement 3—source data 1.** Log2 fold changes of small RNA abundance for satellite DNAs (satDNAs) and piRNA clusters from mutants of proteins in the primary piRNA pathway, pathway for piRNA precursor export from the nucleus and ping-pong pathway.

long RNA transcript levels for satDNAs (**Supplementary file 8**). To confirm this finding, we performed qRT-PCR using total RNA from ovaries of *rhino* (*rhi*<sup>-</sup>) and *moonshiner* (*moon*<sup>-</sup>) 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 in *rhi* and *moon* mutants compared to wild type (p-value=0.048 and 0.077; **Figure 3C**). Because satDNAs have generally low expression levels (*Rsp* and 1.688 total RNA levels are ~3% and ~25%, respectively, of both 42AB and 80F), we may have insufficient power to detect decreased expression in the mutants. It is also possible that the signal is masked by non-precursor transcripts. That is, there may be two kinds of transcription at satDNA loci: (1) RDC-regulated transcription that generates non-polyadenylated piRNA precursors and, (2) non-precursor transcription, which is not well characterized and may also largely lack polyadenylation. In this context, it would be difficult to distinguish precursor from non-precursor transcripts derived from satDNA. However, when we reanalyzed the total and poly-A selected RNA-seq data from the *rhi* mutant (**EIMaghraby et al., 2019**), we find that the abundance of poly-A

transcripts (which are likely a subset of non-precursors) is increased for *Rsp* and unchanged for 1.688 (**Supplementary file 9**) relative to wild type. This result suggests that changes in piRNA precursor levels may be masked by the non-precursor levels, similar to reports on piRNA cluster transcription in embryonic *piwi* knockdown ovaries (**Akkouche et al., 2017**). This situation might arise if only a subset of satDNA repeats are RDC-regulated. Alternatively, the proportion of piRNA precursor-to-non-precursor transcripts in these mutants might shift such that the abundance of piRNA precursors decreases but the total RNA level does not.

We also asked if the satDNA-derived piRNA pool is affected in mutants of 12 genes involved in piRNA precursor export from the nucleus, primary piRNA biogenesis, and the ping-pong cycle (**Figure 3—figure supplement 3**; **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**; **Supplementary file 1**). For each of the datasets analyzed, we recapitulate previously reported results for all known piRNA clusters (**Figure 3—figure supplement 3**; **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). Some of our reanalysis 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 increased levels in an independent Zucchini dataset (**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 dual-strand piRNA pathway.

## Heterochromatin establishment at satDNAs requires Piwi

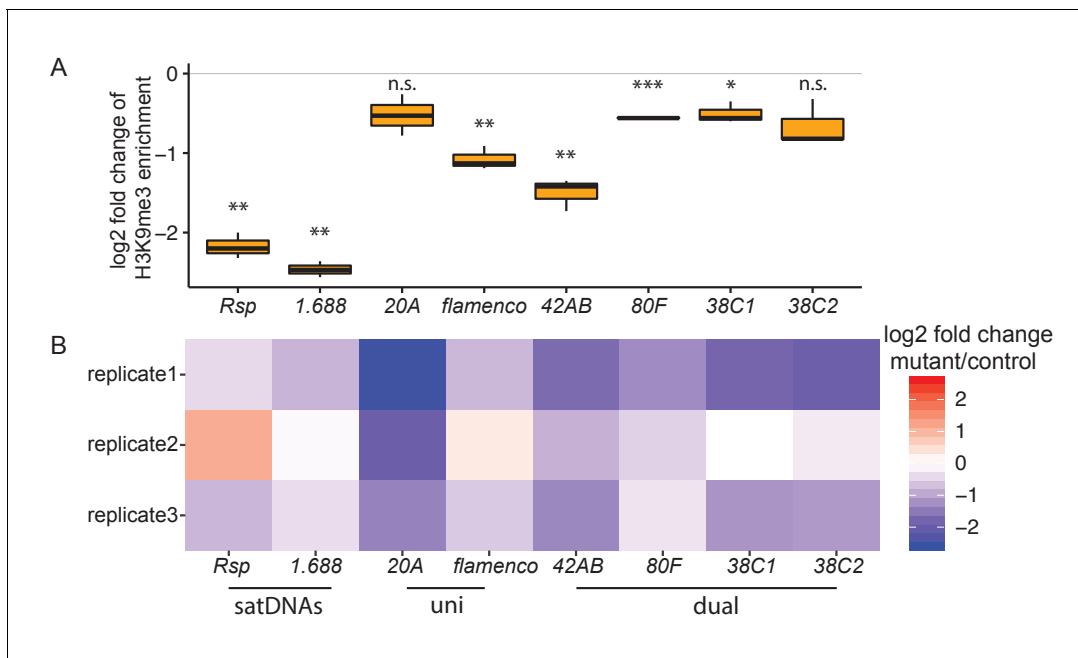
Consistent with their Rhi enrichment, we find that satDNAs are enriched for H3K9me3 in ovaries (**Figure 2—figure supplement 5B**; datasets from **Klenov et al., 2014**; **Le Thomas et al., 2014**; **Mohn et al., 2014**). Piwi plays an important role in establishing H3K9 methylation on euchromatic TEs in ovaries (**Mohn et al., 2014**) and heterochromatin more generally in embryos (**Akkouche et al., 2017**). Transiently knocking down *piwi* expression early in the embryonic germline leads to a general decrease in H3K9me3 in the adult ovary, and a specific decrease in piRNA production and increase in spliced non-precursor transcripts at dual-strand piRNA clusters (**Akkouche et al., 2017**). We therefore reanalyzed H3K9me3 ChIP-seq data from embryonic *piwi* knock down ovaries (**Akkouche et al., 2017**). We detected a decrease of H3K9me3 at satDNAs (**Figure 4A**), suggesting that Piwi is also required for the establishment of heterochromatin at these loci. Consistent with the decrease in H3K9me3, piRNA production from satDNAs is also reduced (with some variation among replicates observed for *Rsp*; **Figure 4B**); and satDNA total RNA levels are increased (**Figure 4—figure supplement 1**), similar to dual-strand piRNA clusters (**Akkouche et al., 2017**). However, it is again difficult to distinguish between satDNA precursor and non-precursor RNAs.

While Piwi is important for heterochromatin establishment at piRNA clusters, it appears to be dispensable for heterochromatin maintenance (**Czech et al., 2018**). 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**). Similar to piRNA clusters, our reanalysis of these data shows that the level of H3K9me3 on satDNAs is largely unchanged in the knockdown or mutant ovaries (with some variation observed among datasets; **Supplementary file 10**). These analyses suggest a role for Piwi in establishing, but not maintaining, heterochromatin at satDNAs in early embryos, which is important for producing piRNAs later in adult ovaries.

## Conclusions

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, our findings support a more general role for these pathways. Here we show that transcription from satDNAs is regulated by the heterochromatin-dependent RDC





**Figure 4.** Heterochromatin establishment disrupted at satellite DNAs (satDNAs) in *piwi* embryonic knockdown ovaries. (A) Log<sub>2</sub> fold change of H3K9me3 ChIP/input enrichment shows satDNA H3K9me3 levels decrease in *piwi* embryonic knockdown ovaries compared to control. Source data in **Figure 4—source data 1**. p-values are estimated by one-sample t-test ( $\mu = 0$ ) with FDR corrections (**Benjamini and Hochberg, 1995**). \* adjusted p-value < 0.05, \*\* adjusted p-value < 0.01, \*\*\* adjusted p-value < 0.001. (B) Log<sub>2</sub> fold change of small RNA abundance shows satDNA small RNA levels decrease compared to controls, with variation observed for replicate2. Small RNA abundance is normalized to the number of reads mapped to miRNAs. Source data in **Figure 4—source data 2**.

The online version of this article includes the following source data and figure supplement(s) for figure 4:

**Source data 1.** Log<sub>2</sub> fold change of H3K9me3 ChIP/input enrichment for satellite DNAs (satDNAs) and piRNA clusters in *piwi* embryonic knockdown ovaries.

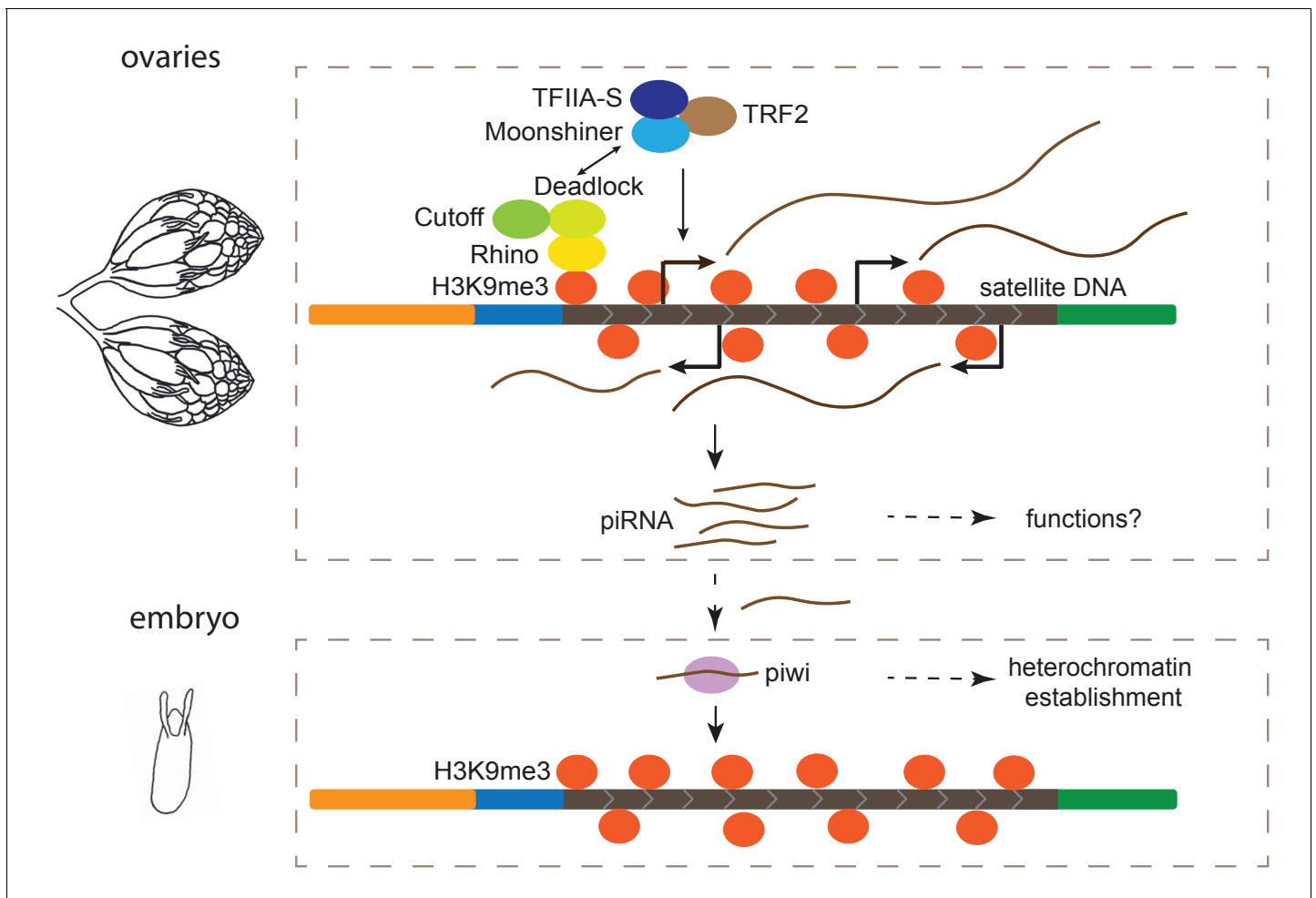
**Source data 2.** Log<sub>2</sub> fold change of small RNA abundance for satellite DNAs (satDNAs) and piRNA clusters in *piwi* embryonic knockdown ovaries.

**Figure supplement 1.** Log<sub>2</sub> fold change of total RNA abundance shows satellite DNA (satDNA) long RNA levels increase in *piwi* embryonic knockdown ovaries compared to control.

**Figure supplement 1—source data 1.** Log<sub>2</sub> fold change of total RNA abundance for satellite DNA (satDNA) and piRNA clusters in *piwi* embryonic knockdown ovaries.

machinery and Moon in ovaries and these transcripts are processed into piRNAs. Thus, complex satDNA transcription is regulated in a manner similar to dual-strand piRNA clusters in the female germline (**Figure 5**).

Our findings are consistent with a study that detected bidirectional transcription of the 1.688 satDNA family in ovaries (**Usakin et al., 2007**) and a recent analysis of satDNA-derived piRNAs in RDC mutants (**Chen et al., 2020**). Usakin et al. found that 1.688 transcript abundance is elevated in mutants of two piRNA processing genes, *spn-E* and *aub* (**Usakin et al., 2007**), suggesting that 1.688 is targeted by piRNAs, similar to TEs. However, the origins of the 1.688 piRNAs and how the transcription of precursors is regulated were unclear (**Usakin et al., 2007**). Here we provide evidence that most satellite-derived transcripts and small RNAs reported in previous studies (**Aravin et al., 2003; Saito et al., 2006; Usakin et al., 2007; Rošić et al., 2014; Chen et al., 2020; Chen et al., 2021**) come from the heterochromatin-dependent transcription of the large satDNA blocks. The role of these piRNAs in ovaries remains unknown, and we understand even less about piRNA biogenesis and function in *D. melanogaster* testes, where we also detect satDNA-derived piRNAs. Proportionally, far fewer piRNAs in the male germline are derived from TEs than in the female germline (**Nishida et al., 2007; Nagao et al., 2010; Quénerch’du et al., 2016**), suggesting roles outside of TE repression. For example, recent studies implicate piRNA pathways in intragenomic conflicts (e.g., male meiotic drive; **Gell and Reenan, 2013; Courret et al., 2019**), with satDNAs often at the center of these conflicts.



**Figure 5.** Model for maintenance of satellite DNA (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. While their functions in ovaries are unclear, 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.

While it will take more work to understand the role of satDNA-derived transcripts in the germline, we hypothesize that the maternal deposition of these piRNAs contributes to heterochromatin establishment at satDNAs in the early embryo (**Figure 5**). Maternal deposition of Piwi contributes to heterochromatin establishment in the embryo (**Gu and Elgin, 2013**), and Piwi-dependent H3K9me3 deposition at canonical piRNA clusters is important for subsequent piRNA production at piRNA clusters (**Akkouche et al., 2017**). Similar to piRNA clusters, we found evidence that both H3K9me3 chromatin and piRNA production from complex satDNA is reduced when transiently depleting Piwi in the embryos, suggesting a role for the piRNA pathway in heterochromatin establishment at satDNA loci (**Figure 5**). We propose a simple model of self-regulation, where Piwi, guided by satDNA-derived piRNAs, establishes H3K9me3 at satDNA, marking the satDNAs as piRNA production sites later in development (**Figure 5**). While a contributor, Piwi might not be the only factor necessary for heterochromatin establishment in embryos (**Wei et al., 2021a**). And once established, the maintenance of heterochromatin at piRNA clusters and satDNAs is not Piwi-dependent (**Supplementary file 10; Klenov et al., 2014; Mohn et al., 2014**). Therefore, the piRNA pathway is likely to be one of several factors important for proper packaging and regulation of repeat-rich regions of the genome (**Pal-Bhadra et al., 2004; Gu and Elgin, 2013**).

The consequences of disrupting satDNA packaging/regulation are likely to be complicated. The ramifications could be especially serious if a reduction in heterochromatin at satDNA in early

embryos affects heterochromatin in all tissues (reviewed in [Janssen et al., 2018](#)) and/or if establishing heterochromatin at satDNA loci serves as nucleation points for the larger-scale heterochromatinization of pericentric regions. 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](#)). Unregulated satDNA may disrupt this organization ([Novo et al., 2020](#)) and lead to cell death ([Jagannathan et al., 2019](#)). Second, de-repressed satDNA may lead to genome instability ([Peng and Karpen, 2007](#)) including chromosomal structural rearrangements (reviewed in [Janssen et al., 2018](#)). In the short term, rearrangements involving satDNA may lead to mitotic defects in the developing embryo as they can affect chromosome segregation ([Ferree and Barbash, 2009](#); [Ferree, 2014](#)). Over longer evolutionary time scales, these rearrangements contribute to variation in satDNA organization between individuals and species, and may cause genetic incompatibilities between closely related species ([Ferree and Barbash, 2009](#)). SatDNAs are indeed among the most rapidly evolving sequences in genomes (reviewed in [Ferree and Prasad, 2012](#); [Plohl et al., 2012](#)).

Many mysteries remain surrounding the functions of the piRNA pathway outside of its role in controlling TE activity. Our finding that the piRNA pathway regulates satDNA suggests a general role for the piRNA pathway and for maternal satDNA-derived RNAs in remodeling chromatin in the developing embryo. This initial establishment of heterochromatin may be an important step in ensuring genome integrity throughout development and in adult tissues, but this remains an open question. 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

### Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Gene ( <i>Drosophila melanogaster</i> )	<i>rhi</i>	Flybase	Flybase: FBgn0004400	
Gene ( <i>D. melanogaster</i> )	<i>moon</i>	Flybase	Flybase: FBgn0030373	
Gene ( <i>D. melanogaster</i> )	<i>del</i>	Flybase	Flybase: FBgn0086251	
Gene ( <i>D. melanogaster</i> )	<i>cuff</i>	Flybase	Flybase: FBgn0260932	
Gene ( <i>D. melanogaster</i> )	<i>piwi</i>	Flybase	Flybase: FBgn0004872	
Strain, strain background ( <i>D. melanogaster</i> , female and male)	Iso-1	Bloomington Drosophila Stock Center (BDSC)	BDSC: 2057; RRID:BDSC_2057	
Strain, strain background ( <i>D. melanogaster</i> , female)	ZW144	doi:10.1534/g3.114.015883 <a href="#">Grenier et al., 2015</a>		
Strain, strain background ( <i>D. melanogaster</i> , female)	Ral357	BDSC	BDSC:25184; RRID:BDSC_25184	
Strain, strain background ( <i>D. melanogaster</i> , female)	Ral380	BDSC	BDSC:25190; RRID:BDSC_25190	
Strain, strain background ( <i>D. melanogaster</i> , female)	It pk cn bw	<a href="#">Ganetzky, 1977</a>		
Strain, strain background ( <i>D. melanogaster</i> , female)	w <sup>1118</sup>	BDSC	BDSC:5905; RRID:BDSC_5905	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Strain, strain background ( <i>D. melanogaster</i> , female)	w <sup>1</sup>	BDSC	BDSC:2390; RRID:BDSC_2390	
Strain, strain background ( <i>D. melanogaster</i> , female)	OregonR	BDSC	BDSC:2376; RRID:BDSC_2376	
Genetic reagent ( <i>D. melanogaster</i> )	<i>rhi</i> mutant	Vienna Drosophila Resource Center (VRDC)	VDRC:313487	
Genetic reagent ( <i>D. melanogaster</i> )	<i>rhi</i> mutant	VRDC	VDRC:313488	
Genetic reagent ( <i>D. melanogaster</i> )	<i>moon</i> mutant	VRDC	VDRC:313735	
Genetic reagent ( <i>D. melanogaster</i> )	<i>moon</i> mutant	VRDC	VDRC:313738	
Sequence-based reagent	RPS3 forward	IDT	qPCR primer	AGTTGTACGCCGAGAAGGTG
Sequence-based reagent	RPS3 Reverse	IDT	qPCR primer	TGTAGCGGAGCACACCATAG
Sequence-based reagent	tRNA forward	IDT	qPCR primer	CTAGCTCAGTCGGTAGAGCATGA
Sequence-based reagent	tRNA Reverse	IDT	qPCR primer	CCAACGTGGGGCTCGAAC
Sequence-based reagent	<i>Rsp</i> forward	IDT	qPCR primer	GGAAAATCACCCATTTTGATCGC
Sequence-based reagent	<i>Rsp</i> Reverse	IDT	qPCR primer	CCGAATTCAAGTACCAGAC
Sequence-based reagent	Probe for 1.688	IDT	RNA FISH probe	Cy5TTTTCCAAATTTTCGGT CATCAAATAATCAT
Sequence-based reagent	Probe for <i>Rsp</i>	Stellaris	RNA FISH probe	Custom Stellaris FISH probes with 45 sequences listed in <b>Supplementary file 11</b>
Sequence-based reagent	T7_rsp2	IDT	Northern blot probe synthesis primer	TAATACGACTCACTATAGGG CCGAATTCAAGTACCAGAC
Sequence-based reagent	rsp1	IDT	Northern blot probe synthesis primer	GGAAAATCACCCATTTTGATCGC
Sequence-based reagent	<i>Rsp</i> primer_F	IDT	Slot blot probe synthesis primer	TAATACGACTCACTATAGGG AAAATCACCCATTTTGATCGC
Sequence-based reagent	<i>Rsp</i> primer_R	IDT	Slot blot probe synthesis primer	CCGAATTCAAGTACCAGAC
Sequence-based reagent	rp49 primer_F	IDT	Slot blot probe synthesis primer	TAATACGACTCACTATAGGG CAGTAAACGCGTTCTGCATG
Sequence-based reagent	rp49 primer_R	IDT	Slot blot probe synthesis primer	CAGCATACAGGCCCAAGATC
Software, algorithm	Bowtie2	doi:10.1038/nmeth.1923.	RRID:SCR_016368	
Software, algorithm	Bowtie	doi:10.1002/0471250953.bi1107s32.	RRID:SCR_005476	
Software, algorithm	DESeq2	doi:10.1186/s13059-014-0550-8.	RRID:SCR_015687	
Software, algorithm	piPipes	doi:10.1093/bioinformatics/btu647.		
Software, algorithm	BLAST	NCBI	RRID:SCR_004870	
Software, algorithm	R	R core team	RRID:SCR_001905	
Software, algorithm	Customized Python scripts	This paper		<b>Wei et al., 2021b</b> GitHub ( <a href="https://github.com/LarracuentaLab/Dmelanogaster_satDNA_regulation">https://github.com/LarracuentaLab/Dmelanogaster_satDNA_regulation</a> )

## Drosophila stocks

Iso-1 (RRID:BDSC\_2057) was used as the wild-type strain, unless stated otherwise. In the qPCR validation experiment, *rhi* mutants (*rhi*<sup>-</sup>) are transheterozygotes from the Vienna Drosophila Resource Center (VDRC 313487 and 313488) as are the *moonshiner* mutants (*moon*<sup>-</sup>) (VDRC 313735 and 313738) as described in Andersen et al., 2017. Based on the origin and genetic background of these mutants, *w*<sup>1118</sup> (RRID:BDSC\_5905) or the progeny from OregonR (Ore) (RRID:BDSC\_2376) crossed to *w*<sup>1</sup> (RRID:BDSC\_2390) were used as the wild-type controls for *rhi*<sup>-</sup> and *moon*<sup>-</sup>. All flies were maintained at 23°C on cornmeal medium.

## Small RNA-seq

6–8-day-old testes were dissected in RNase-free PBS buffer. Total RNA was extracted using *mirVana* miRNA Isolation Kit (Ambion) with procedures for isolating RNA fractions enriched for small RNAs (<200 nt), then treated with RNase free DNase I (Promega) at 37°C for 1 hr. 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

6–8-day-old testes were dissected in RNase-free PBS buffer. Total RNA was extracted using *mirVana* miRNA Isolation Kit (Ambion) with procedures for isolating RNA fractions enriched for long RNAs (>200 nt), then treated with RNase free DNase I (Promega) at 37°C for 1 hr. 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 Larracunte, 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, unless stated otherwise. Reads were then counted based on the annotations of their assigned mapping locations. Because there is high-sequence similarity among the *1.688* subfamily repeats (*260-bp*, *359-bp*, *353-bp*, *356-bp*), all *1.688* subfamilies were combined, unless stated otherwise. 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 **Supplementary file 1**. Total RNA-seq reads were trimmed for adaptors and then mapped to the genome using Bowtie2 (RRID:SCR\_016368) (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 (Wei et al., 2021b [https://github.com/LarracunteLab/Dmelanogaster\\_satDNA\\_regulation](https://github.com/LarracunteLab/Dmelanogaster_satDNA_regulation); Wei, 2020, `htseq_bam_count_proportional.py`; Wei et al., 2021b). For the *1.688* subfamilies, all subfamilies were combined into one *1.688* category, although analyzing each by subfamily (e.g., *353-bp*, *356-bp*, *359-bp*, *260-bp*) does not change our conclusions ([https://github.com/LarracunteLab/Dmelanogaster\\_satDNA\\_regulation](https://github.com/LarracunteLab/Dmelanogaster_satDNA_regulation)) (Wei, 2020; Wei et al., 2021b). For results shown in **Supplementary file 8**, DESeq2 (RRID:SCR\_015687) (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 log<sub>2</sub> 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](https://github.com/Larracuentelab/Dmelanogaster_satDNA_regulation); Wei, 2020; htseq\_bam\_count\_proportional.py; Wei et al., 2021b), and subsequent differential expression analysis between mutants and wildtype was done using DESeq2 (RRID:SCR\_015687) (Love et al., 2014; results shown in Figure 3—figure supplement 2).

### Small RNA-seq analysis

All small RNA-seq datasets reanalyzed in our study are listed in **Supplementary file 1**. Small RNA-seq reads were trimmed for adaptors, then mapped to the genome using Bowtie (RRID:SCR\_005476) (Langmead, 2010). A customized Python script ([https://github.com/Larracuentelab/Dmelanogaster\\_satDNA\\_regulation](https://github.com/Larracuentelab/Dmelanogaster_satDNA_regulation); Wei, 2020, htseq\_bam\_count\_proportional.py; Wei et al., 2021b) 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 log<sub>2</sub> fold changes of these normalized counts in mutants compared to wild type (i.e.,  $\log_2(\text{count}_{\text{mutant}}/\text{count}_{\text{WT}})$ ) 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](https://github.com/Larracuentelab/Dmelanogaster_satDNA_regulation); Wei, 2020, extract\_sequence\_by\_feature\_gff.py; Wei et al., 2021b). 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—figure supplement 2). Reads (ChIP or piRNA) mapping to a particular satDNA or genomic satDNA variant (as a control) were BLAST-ed to the consensus dimer (for 1.688 satellite) or trimer (for *Rsp* because it has left and right consensus sequences), and then coordinates were converted along a dimer/trimer to coordinates along a monomer/dimer consensus sequence. All plots were made in R (R Development Core Team, 2017).

### ChIP/RIP-seq analysis

All total ChIP-seq and RIP-seq datasets reanalyzed in our study are listed in **Supplementary file 1**. ChIP-seq and RIP-seq reads were trimmed for adaptors and mapped to the genome using Bowtie2 (RRID:SCR\_016368) (Langmead and Salzberg, 2012). A customized Python script ([https://github.com/Larracuentelab/Dmelanogaster\\_satDNA\\_regulation](https://github.com/Larracuentelab/Dmelanogaster_satDNA_regulation); Wei, 2020, htseq\_bam\_count\_proportional.py; Wei et al., 2021b) 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 ChIP-seq analyses, consider satDNA as discrete loci rather than repeat unit types is appropriate because some loci are composed of several repeat types. To examine the large blocks of heterochromatic satDNA chromatin for the Rhi and H3K9me3 ChIP-seq analyses, euchromatic 1.688 satDNAs were excluded and only reads that map uniquely to satDNA loci were analyzed. Heterochromatic satDNA loci were defined as discrete loci on chromosomes: 2L (2L\_2: 402701–460225; the 260-bp locus), 3L (3L\_3: 46695–106272; primarily 353-bp and 356-bp repeats), and the unmapped contigs (Contig101 and Contig9; 353-bp, 356-bp, and 359-bp repeats). Our conclusions do not change when we look at all reads (not just uniquely mapped; [https://github.com/Larracuentelab/Dmelanogaster\\_satDNA\\_regulation](https://github.com/Larracuentelab/Dmelanogaster_satDNA_regulation); Wei, 2020; Wei et al., 2021b). These analyses were repeated by combining all 1.688 subfamilies into a single category, and each subfamily was analyzed separately (e.g., all 353-bp repeats combined) but the conclusions do not change ([https://github.com/Larracuentelab/Dmelanogaster\\_satDNA\\_regulation](https://github.com/Larracuentelab/Dmelanogaster_satDNA_regulation); Wei, 2020; Wei et al., 2021b). Euchromatic controls are included for the Rhi and H3K9me3 ChIP-seq analyses. Here, the euchromatic control corresponds to the median enrichment score for protein coding genes that are 5 Mb distal from heterochromatin boundaries (Riddle et al., 2011) and piRNA clusters.

For the RIP-seq analyses, reported was the percentage of reads mapped to each repeat and piRNA cluster with miRNAs as the negative control. For the 1.688 subfamilies, all subfamilies were

combined into one 1.688 category, although analyzing each by subfamily (e.g., 353-bp, 356-bp, 359-bp, 260-bp) does not change the conclusions ([https://github.com/Larracuentelab/Dmelanogaster\\_satDNA\\_regulation](https://github.com/Larracuentelab/Dmelanogaster_satDNA_regulation); Wei, 2020; Wei et al., 2021b).

## RNA FISH

A Cy5-labeled oligo probe (5'-Cy5TTTTCCAAATTTTCGGTCATCAAATAATCAT-3') previously described in Ferree and Barbash, 2009 was used to detect 1.688 transcripts from all subfamilies except 260-bp on chromosome 2L. Custom Stellaris FISH probes were designed for *Rsp* (Supplementary file 11), and RNA FISH was performed following the manufacturer's instructions (Biosearch Technologies, Inc). 3–6-day-old ovaries and testes were dissected in RNase-free PBS buffer, fixed with 4% paraformaldehyde in PBS buffer at room temperature for 30 min, and then washed twice with PBS for 5 min. To permeabilize, tissues were kept in RNase free 70% ethanol at 4°C overnight. The ethanol was aspirated, and samples washed with Stellaris wash buffer on a nutating mixer for 3 min and kept still for 2 min at room temperature. Hybridization was then performed with each probe in Stellaris hybridization buffer in a humidity chamber at 37°C overnight. The working concentration was 100 nM for the oligo probe and 125 nM for the Stellaris probes. From this point, samples were kept in the dark. The samples were washed with Stellaris wash buffer twice at 37°C for 30 min 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 (100 µg/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 37°C for 1 hr and washed three times with Stellaris wash buffer at room temperature for 10 min before hybridization.

For RNase H controls, after probe hybridization and washing, tissues were treated with 1.5 µl RNase H (5000 units/ml; New England Biolabs) in 50 µl final volume in 1X RNase H buffer at 37°C for 2 hr and washed three times with Stellaris wash buffer at room temperature for 10 min before mounting and imaging.

## qPCR

For genomic DNA qPCR, 3–8-day-old flies were mashed with pipette tips for 5–10 s and incubated in buffer (10 mM Tris-Cl pH 8.2, 1 mM EDTA, 25 mM NaCl, 200 µg/ml Proteinase K) at 37°C for 30 min (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 45–60 s and then centrifuged for 3–5 min. The aqueous top layers were saved, an equal volume of Sevag added, and the samples vortexed for 30 s then centrifuged for 1 min. The aqueous top layers were saved and a second Sevag extraction performed. Diluted nucleic acid samples (concentration of 0.04 ng/µl) 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, 3–6-day-old ovaries were dissected in RNase-free PBS buffer, and total RNA was extracted using the *mirVana* miRNA Isolation Kit (Ambion). RNA samples were treated with RNase free DNase I (Promega) at 37°C for 1 hr. 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,  $\Delta\Delta\text{CT}$  was calculated (Livak and Schmittgen, 2001).

The replicate number for genomic DNA qPCR is 2–4 and for RNA qRT-PCR is 4–6. The sequences of primers used are: *Rsp* (forward: GGAAAATCACCCATTTTGATCGC, reverse: CCGAATCAAGTACCAGAC); 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 (0–20 hr old) and aged for 6 days. Ovaries were dissected from approximately 20 females (i.e., 6.0–6.8 days old) from each stock, and total nucleic acid isolated using a standard phenol/Sevag procedure (Khost et al., 2017). Total nucleic acid was then treated with DNase I as recommended (20 units; Promega), re-extracted with phenol/Sevag, 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 min. Samples were phenol/Sevag extracted, 10 µg of yRNA added as carrier, and ethanol precipitated.

### Northern blot analysis

Total RNA (10 µg)/RNase controls were suspended in 1× MOPS (0.04 M morpholinepropanesulfonic acid [MOPS] 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 hr. Standard RNA lanes were cut from the gel and stained with ethidium bromide to monitor electrophoresis. Gels were washed for 25 min in sterile water (with four 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 hr 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 hr 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 1 M sodium phosphate pH 6.8, 0.5 M EDTA, 5% SDS (2×, 10 min each) at 60°C and then at 1 M sodium phosphate pH 6.8, 0.5M EDTA, 1% SDS (3×, 10 min 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 min at room temperature, the sample transferred to a tube with an equal volume of ice-cold loading buffer (0.1× 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 min with 2× SSC before being UV crosslinked (Stratalinker). The membrane was first hybridized with a biotinylated rp49 RNA probe in North2-South 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 0.1× SSC/0.5% SDS (three times for ~20 min each) and re-hybridized with a *Rsp* probe (60°C overnight) and processed as above. Signals were quantitated using the ImageLab software (Bio-Rad). We determined the relative signal compared to Iso-1 for each line (5–7 replicates), and then estimate the *Rsp* copy number by scaling the relative slot blot signal to our estimate of *Rsp* copy number in Iso-1 (1100 repeats). Our Iso-1 estimate is based on *Rsp* count in a long-read assembly, which is supported by empirical slot blots (Khost et al., 2017).

To make the biotinylated RNA probes, gel extracted PCR amplicons (primers: *Rsp* 5'-TAATACGACTCACTATAGGGGAAAATCACCCATTTTGATCGC-3' and 5'-CCGAATTCAAGTACCAGAC-3'; rp49 5'- TAATACGACTCACTATAGGGCAGTAAACGCGTTCTGCATG-3' and 5'-CAGCA

TACAGGCCCAAGATC-3') were transcribed using the Biotin RNA Labeling Mix (Roche) and T7 polymerase (Promega).

### 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 are deposited in GitHub ([https://github.com/LarracuenteLab/Dmelanogaster\\_satDNA\\_regulationWei](https://github.com/LarracuenteLab/Dmelanogaster_satDNA_regulationWei), 2020) and at the Dryad Digital Repository (<https://doi.org/10.5061/dryad.hdr7sqvj3>; Wei et al., 2021b).

### Acknowledgements

This work was supported by the National Institutes of Health General Medical Sciences (R35 GM119515 to AML), a Stephen Biggar and Elisabeth Asaro fellowship in Data Science to AML, and a University of Rochester Agnes M. and George Messersmith Dissertation 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.

---

### Additional information

#### Funding

Funder	Grant reference number	Author
National Institutes of Health	R35 GM119515	Amanda M Larracuente
University of Rochester	Stephen Biggar and Elisabeth Asaro fellowship	Amanda M Larracuente
University of Rochester	Agnes M. and George Messersmith Dissertation Fellowship	Xiaolu Wei

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

#### Author contributions

Xiaolu Wei, Conceptualization, Resources, Software, Formal analysis, Validation, Visualization, Methodology, Writing - original draft, Writing - review and editing; Danna G Eickbush, Formal analysis, Validation, Visualization, Methodology, Writing - review and editing; Iain Speece, Formal analysis, Writing - review and editing; Amanda M Larracuente, Conceptualization, Supervision, Funding acquisition, Investigation, Project administration, Writing - review and editing

#### Author ORCIDs

Xiaolu Wei  <https://orcid.org/0000-0001-9952-3757>

Amanda M Larracuente  <https://orcid.org/0000-0001-5944-5686>

#### Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.62375.sa1>

Author response <https://doi.org/10.7554/eLife.62375.sa2>

---

### Additional files

#### Supplementary files

- Supplementary file 1. List of datasets used in this paper.
- Supplementary file 2. Count in RPM values of *Rsp/1.688* transcripts in total and poly-A RNA-seq datasets from various tissues.

- Supplementary file 3. Slot blot estimate of *Rsp* copy number.
- Supplementary file 4. *Rsp* expression level correlates with its copy number in the genome. Estimates from slot blot, northern blot, qPCR, and qRT-PCR indicate correlation between *Rsp* genomic copy number and expression level.
- Supplementary file 5. RIP-seq results for Piwi, Aub, and Ago3 from ovaries.
- Supplementary file 6. Log2 fold changes of small RNA abundance for satDNAs and piRNA clusters normalized to miRNA abundance in Rhino, Deadlock, and Cutoff (RDC) and Moon mutants.
- Supplementary file 7. Log2 fold changes of small RNA abundance for satDNAs and piRNA clusters normalized to the *flamenco* piRNA cluster in Rhino, Deadlock, and Cutoff (RDC) and Moon mutants.
- Supplementary file 8. Log2 fold change of total RNA abundance for satDNAs and piRNA clusters in Rhino, Deadlock, and Cutoff (RDC) and Moon mutants.
- Supplementary file 9. Log2 fold change of total and polyA selected RNA abundance in *rhi* mutant for satellite DNAs (satDNAs) and piRNA clusters.
- Supplementary file 10. Log2 fold change of H3K9me3 ChIP/input enrichment levels in *piwi* germline knockdown/mutant for satellite DNAs (satDNAs) and piRNA clusters.
- Supplementary file 11. *Rsp* probe sequences for RNA fluorescence in situ hybridization (FISH).
- Transparent reporting form

### Data availability

Sequencing data generated in this study have been deposited in NCBI Sequence Read Archive (SRA) under project accession PRJNA647441. Published sequencing data used in this study are from NCBI SRA database, and the full list of accession numbers can be found in Supplementary File 1.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Wei X, Eickbush DG, Speece I, Larracuent AM	2020	Heterochromatin-dependent transcription of satellite DNAs in the <i>Drosophila melanogaster</i> female germline	<a href="https://www.ncbi.nlm.nih.gov/bioproject/PRJNA647441/">https://www.ncbi.nlm.nih.gov/bioproject/PRJNA647441/</a>	NCBI BioProject, PRJNA647441

## 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*. *Molecular Genetics and Genomics* **264**:371–377. DOI: <https://doi.org/10.1007/s004380000331>
- Akkouche A**, Mugat B, Barckmann B, Varela-Chavez C, Li B, Raffel R, Pélisson A, Chambeyron S. 2017. Piwi is required during *Drosophila* embryogenesis to license Dual-Strand piRNA clusters for transposon repression in adult ovaries. *Molecular Cell* **66**:411–419. DOI: <https://doi.org/10.1016/j.molcel.2017.03.017>
- Allshire RC**, Madhani HD. 2018. Ten principles of heterochromatin formation and function. *Nature Reviews Molecular Cell Biology* **19**:229–244. DOI: <https://doi.org/10.1038/nrm.2017.119>, PMID: 29235574
- Andersen PR**, Tirian L, Vunjak M, Brennecke J. 2017. A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* **549**:54–59. DOI: <https://doi.org/10.1038/nature23482>, PMID: 28847004
- 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. DOI: [https://doi.org/10.1016/S1534-5807\(03\)00228-4](https://doi.org/10.1016/S1534-5807(03)00228-4), PMID: 12919683
- Aravin A**, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T, Chien M, Russo JJ, Ju J, Sheridan R, Sander C, Zavolan M, Tuschl T. 2006. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442**:203–207. DOI: <https://doi.org/10.1038/nature04916>, PMID: 16751777
- 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. DOI: <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Biscotti MA**, Canapa A, Forconi M, Olmo E, Barucca M. 2015. Transcription of tandemly repetitive DNA: functional roles. *Chromosome Research* **23**:463–477. DOI: <https://doi.org/10.1007/s10577-015-9494-4>, PMID: 26403245



- Bouzinba-Segard H**, Guais A, Francastel C. 2006. Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *PNAS* **103**:8709–8714. DOI: <https://doi.org/10.1073/pnas.0508006103>, PMID: 16731634
- 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. DOI: <https://doi.org/10.1016/j.cell.2007.01.043>
- 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. DOI: <https://doi.org/10.1126/science.161.3841.529>, PMID: 4874239
- Brown JB**, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, Booth BW, Wen J, Park S, Suzuki AM, Wan KH, Yu C, Zhang D, Carlson JW, Cherbas L, Eads BD, Miller D, Mockaitis K, Roberts J, Davis CA, et al. 2014. Diversity and dynamics of the *Drosophila* transcriptome. *Nature* **512**:393–399. DOI: <https://doi.org/10.1038/nature12962>, PMID: 24670639
- Chang CH**, Larracuente AM. 2019. Heterochromatin-Enriched assemblies reveal the sequence and organization of the *Drosophila melanogaster* Y Chromosome. *Genetics* **211**:333–348. DOI: <https://doi.org/10.1534/genetics.118.301765>, PMID: 30420487
- Charlesworth B**, Sniegowski P, Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**:215–220. DOI: <https://doi.org/10.1038/371215a0>, PMID: 8078581
- Chen P**, Luo Y, Aravin AA. 2020. RDC complex executes a dynamic piRNA program during *Drosophila* spermatogenesis to safeguard male fertility. *bioRxiv*. DOI: <https://doi.org/10.1101/2020.08.25.266643>
- Chen P**, Kotov AA, Godneeva BK, Bazylev SS, Olenina LV, Aravin AA. 2021. piRNA-mediated gene regulation and adaptation to sex-specific transposon expression in *D. melanogaster* male germline. *Genes & Development* **35**:914–935. DOI: <https://doi.org/10.1101/gad.345041.120>
- Courret C**, Chang CH, Wei KH, Montchamp-Moreau C, Larracuente AM. 2019. Meiotic drive mechanisms: lessons from *Drosophila*. *Proceedings of the Royal Society B: Biological Sciences* **286**:20191430. DOI: <https://doi.org/10.1098/rspb.2019.1430>, PMID: 31640520
- Czech B**, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, Perrimon N, Kellis M, Wohlschlegel JA, Sachidanandam R, Hannon GJ, Brennecke J. 2008. An endogenous small interfering RNA pathway in *Drosophila*. *Nature* **453**:798–802. DOI: <https://doi.org/10.1038/nature07007>, PMID: 18463631
- 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. *Molecular Cell* **50**:749–761. DOI: <https://doi.org/10.1016/j.molcel.2013.04.007>
- Czech B**, Munafò M, Ciabrelli F, Eastwood EL, Fabry MH, Kneuss E, Hannon GJ. 2018. piRNA-Guided genome defense: from biogenesis to silencing. *Annual Review of Genetics* **52**:131–157. DOI: <https://doi.org/10.1146/annurev-genet-120417-031441>, PMID: 30476449
- Dernburg AF**, Sedat JW, Hawley RS. 1996. Direct evidence of a role for heterochromatin in meiotic chromosome segregation. *Cell* **86**:135–146. DOI: [https://doi.org/10.1016/S0092-8674\(00\)80084-7](https://doi.org/10.1016/S0092-8674(00)80084-7), PMID: 8689681
- Deshpande N**, Meller VH. 2018. Chromatin that guides dosage compensation is modulated by the siRNA pathway in *Drosophila melanogaster*. *Genetics* **209**:1085–1097. DOI: <https://doi.org/10.1534/genetics.118.301173>, PMID: 29921620
- DiBartolomeis SM**, Tartof KD, Jackson FR. 1992. A superfamily of *Drosophila* satellite related (SR) DNA repeats restricted to the X chromosome euchromatin. *Nucleic Acids Research* **20**:1113–1116. DOI: <https://doi.org/10.1093/nar/20.5.1113>, PMID: 1549474
- Doolittle WF**, Sapienza C. 1980. Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**:601–603. DOI: <https://doi.org/10.1038/284601a0>, PMID: 6245369
- EIMaghraby MF**, Andersen PR, Pühringer F, Hohmann U, Meixner K, Lendl T, Tirian L, Brennecke J. 2019. A Heterochromatin-Specific RNA export pathway facilitates piRNA production. *Cell* **178**:964–979. DOI: <https://doi.org/10.1016/j.cell.2019.07.007>
- Fagegaltier D**, König 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. DOI: <https://doi.org/10.1534/genetics.114.169268>, PMID: 25081570
- Ferree PM**. 2014. Mitotic misbehavior of a *Drosophila melanogaster* satellite in ring chromosomes: insights into intragenomic conflict among heterochromatic sequences. *Fly* **8**:101–107. DOI: <https://doi.org/10.4161/fly.29488>, PMID: 25483254
- Ferree PM**, Barbash DA. 2009. Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in *Drosophila*. *PLOS Biology* **7**:e1000234. DOI: <https://doi.org/10.1371/journal.pbio.1000234>, PMID: 19859525
- Ferree PM**, Prasad S. 2012. How can satellite DNA divergence cause reproductive isolation? let Us count the chromosomal ways. *Genetics Research International* **2012**:1–11. DOI: <https://doi.org/10.1155/2012/430136>, PMID: 22567387
- Ferreira D**, Meles S, Escudeiro A, Mendes-da-Silva A, Adegas F, Chaves R. 2015. Satellite non-coding RNAs: the emerging players in cells, cellular pathways and Cancer. *Chromosome Research* **23**:479–493. DOI: <https://doi.org/10.1007/s10577-015-9482-8>, PMID: 26293605
- Fukagawa T**, Nogami M, Yoshikawa M, Ikeno M, Okazaki T, Takami Y, Nakayama T, Oshimura M. 2004. Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nature Cell Biology* **6**:784–791. DOI: <https://doi.org/10.1038/ncb1155>, PMID: 15247924

- Ganetzky B.** 1977. On the components of segregation distortion in *Drosophila melanogaster*. *Genetics* **86**:321–355. DOI: <https://doi.org/10.1093/genetics/86.2.321>, PMID: 407128
- 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. DOI: <https://doi.org/10.1534/genetics.112.147561>, PMID: 23267055
- Ghildiyal M**, Seitz H, Horwich MD, Li C, Du T, Lee S, Xu J, Kittler EL, Zapp ML, Weng Z, Zamore PD. 2008. Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* **320**:1077–1081. DOI: <https://doi.org/10.1126/science.1157396>, PMID: 18403677
- 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. DOI: <https://doi.org/10.1261/rna.1972910>, PMID: 19917635
- Girard A**, Sachidanandam R, Hannon GJ, Carmell MA. 2006. A germline-specific class of small RNAs binds mammalian piwi proteins. *Nature* **442**:199–202. DOI: <https://doi.org/10.1038/nature04917>, PMID: 16751776
- 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, Brown JB, Chervas L, Davis CA, Dobin A, Li R, Lin W, Malone JH, Mattiuzzo NR, Miller D, Sturgill D, et al. 2011. The developmental transcriptome of *Drosophila melanogaster*. *Nature* **471**:473–479. DOI: <https://doi.org/10.1038/nature09715>, PMID: 21179090
- Grenier JK**, Arguello JR, Moreira MC, Gottipati S, Mohammed J, Hackett SR, Boughton R, Greenberg AJ, Clark AG. 2015. Global diversity lines - a five-continent reference panel of sequenced *Drosophila melanogaster* strains. *G3: Genes, Genomes, Genetics* **5**:593–603. DOI: <https://doi.org/10.1534/g3.114.015883>, PMID: 25673134
- Grivna ST**, Beyret E, Wang Z, Lin H. 2006. A novel class of small RNAs in mouse spermatogenic cells. *Genes & Development* **20**:1709–1714. DOI: <https://doi.org/10.1101/gad.1434406>, PMID: 16766680
- Gu T**, Elgin SC. 2013. Maternal depletion of piwi, a component of the RNAi system, impacts heterochromatin formation in *Drosophila*. *PLOS Genetics* **9**:e1003780. DOI: <https://doi.org/10.1371/journal.pgen.1003780>, PMID: 24068954
- 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. DOI: <https://doi.org/10.1126/science.1140494>, PMID: 17322028
- Halbach R**, Miesen P, Joosten J, Taşköprü E, Rondeel I, Pennings B, Vogels CBF, Merklings SH, Koenraadt CJ, Lambrechts L, van Rij RP. 2020. A satellite repeat-derived piRNA controls embryonic development of aedes. *Nature* **580**:274–277. DOI: <https://doi.org/10.1038/s41586-020-2159-2>, PMID: 32269344
- Hall IM**, Shankaranarayana GD, Noma K, Ayoub N, Cohen A, Grewal SI. 2002. Establishment and maintenance of a heterochromatin domain. *Science* **297**:2232–2237. DOI: <https://doi.org/10.1126/science.1076466>, PMID: 12215653
- 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. DOI: <https://doi.org/10.1093/bioinformatics/btu647>, PMID: 25342065
- 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. *The EMBO Journal* **30**:3977–3993. DOI: <https://doi.org/10.1038/emboj.2011.308>, PMID: 21863019
- Houwing S**, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB, Plasterk RH, Hannon GJ, Draper BW, Ketting RF. 2007. A role for piwi and piRNAs in germ cell maintenance and transposon silencing in zebrafish. *Cell* **129**:69–82. DOI: <https://doi.org/10.1016/j.cell.2007.03.026>, PMID: 17418787
- Hutvagner G**, Simard MJ. 2008. Argonaute proteins: key players in RNA silencing. *Nature Reviews Molecular Cell Biology* **9**:22–32. DOI: <https://doi.org/10.1038/nrm2321>, PMID: 18073770
- Jagannathan M**, Cummings R, Yamashita YM. 2018. A conserved function for pericentromeric satellite DNA. *eLife* **7**:e34122. DOI: <https://doi.org/10.7554/eLife.34122>, PMID: 29578410
- Jagannathan M**, Cummings R, Yamashita YM. 2019. The modular mechanism of chromocenter formation in *Drosophila*. *eLife* **8**:e43938. DOI: <https://doi.org/10.7554/eLife.43938>, PMID: 30741633
- Janssen A**, Colmenares SU, Karpen GH. 2018. Heterochromatin: guardian of the genome. *Annual Review of Cell and Developmental Biology* **34**:265–288. DOI: <https://doi.org/10.1146/annurev-cellbio-100617-062653>, PMID: 30044650
- Joshi SS**, Meller VH. 2017. Satellite repeats identify X chromatin for dosage compensation in *Drosophila melanogaster* Males. *Current Biology* **27**:1393–1402. DOI: <https://doi.org/10.1016/j.cub.2017.03.078>
- Khost DE**, Eickbush DG, Larracuent AM. 2017. Single-molecule sequencing resolves the detailed structure of complex satellite DNA loci in *Drosophila melanogaster*. *Genome Research* **27**:709–721. DOI: <https://doi.org/10.1101/gr.213512.116>, PMID: 28373483
- Klattenhoff C**, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A, Seitz H, Zamore PD, Weng Z, Theurkauf WE. 2009. The *Drosophila* HP1 homolog rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* **138**:1137–1149. DOI: <https://doi.org/10.1016/j.cell.2009.07.014>, PMID: 19732946

- 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 Research* **42**: 6208–6218. DOI: <https://doi.org/10.1093/nar/gku268>, PMID: 24782529
- Kuhn GC**, Küttler H, Moreira-Filho O, Heslop-Harrison JS. 2012. The 1.688 repetitive DNA of *Drosophila*: concerted evolution at different genomic scales and association with genes. *Molecular Biology and Evolution* **29**:7–11. DOI: <https://doi.org/10.1093/molbev/msr173>, PMID: 21712468
- Langmead B**. 2010. Aligning short sequencing reads with bowtie. *Curr Protoc Bioinformatics* **11**:17. DOI: <https://doi.org/10.1002/0471250953.bi1107s32>
- Langmead B**, Salzberg SL. 2012. Fast gapped-read alignment with bowtie 2. *Nature Methods* **9**:357–359. DOI: <https://doi.org/10.1038/nmeth.1923>, PMID: 22388286
- Larracuent AM**, Presgraves DC. 2012. The selfish segregation distorter gene complex of *Drosophila melanogaster*. *Genetics* **192**:33–53. DOI: <https://doi.org/10.1534/genetics.112.141390>, PMID: 22964836
- Larson AG**, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, Agard DA, Redding S, Narlikar GJ. 2017. Liquid droplet formation by HP1 $\alpha$  suggests a role for phase separation in heterochromatin. *Nature* **547**: 236–240. DOI: <https://doi.org/10.1038/nature22822>, PMID: 28636604
- 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. DOI: <https://doi.org/10.1126/science.1130164>, PMID: 16778019
- Le Thomas A**, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Tóth KF. 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes & Development* **27**:390–399. DOI: <https://doi.org/10.1101/gad.209841.112>, PMID: 23392610
- Le Thomas A**, Stuwe E, Li S, Du J, Marinov G, Rozhkov N, Chen YC, Luo Y, Sachidanandam R, Toth KF, Patel D, Aravin AA. 2014. Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes & Development* **28**:1667–1680. DOI: <https://doi.org/10.1101/gad.245514.114>, PMID: 25085419
- Lippman Z**, Gendrel A-V, Black M, Vaughn MW, Dedhia N, Richard McCombie W, Lavine K, Mittal V, May B, Kasschau KD, Carrington JC, Doerge RW, Colot V, Martienssen R. 2004. Role of transposable elements in Heterochromatin and epigenetic control. *Nature* **430**:471–476. DOI: <https://doi.org/10.1038/nature02651>
- Livak KJ**, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>(Delta delta C(T)) Method. *Methods* **25**:402–408. DOI: <https://doi.org/10.1006/meth.2001.1262>, PMID: 11846609
- Lohe AR**, Roberts PA. 1988. Evolution of satellite DNA sequences in *Drosophila*. In: Verma R. S (Ed). *Heterochromatin: Molecular and Structural Aspects*. Cambridge University Press. p. 1–10.
- Losada A**, Villasante A. 1996. Autosomal location of a new subtype of 1.688 satellite DNA of *Drosophila melanogaster*. *Chromosome Research* **4**:372–383. DOI: <https://doi.org/10.1007/BF02257273>
- Love MI**, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**:550. DOI: <https://doi.org/10.1186/s13059-014-0550-8>, PMID: 25516281
- 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. DOI: <https://doi.org/10.1016/j.cell.2009.03.040>, PMID: 19395010
- Menon DU**, Coarfa C, Xiao W, Gunaratne PH, Meller VH. 2014. siRNAs from an X-linked satellite repeat promote X-chromosome recognition in *Drosophila melanogaster*. *PNAS* **111**:16460–16465. DOI: <https://doi.org/10.1073/pnas.1410534111>
- 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**:e48940. DOI: <https://doi.org/10.7554/eLife.48940>, PMID: 31687931
- 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. DOI: <https://doi.org/10.1016/j.cell.2014.04.031>, PMID: 24906153
- Mohn F**, Handler D, Brennecke J. 2015. Noncoding RNA piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science* **348**:812–817. DOI: <https://doi.org/10.1126/science.aaa1039>, PMID: 25977553
- 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. DOI: <https://doi.org/10.1261/rna.2270710>
- 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. DOI: <https://doi.org/10.1261/rna.744307>, PMID: 17872506
- Noma K**, Sugiyama T, Cam H, Verdel A, Zofall M, Jia S, Moazed D, Grewal SI. 2004. RITS acts in Cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. *Nature Genetics* **36**:1174–1180. DOI: <https://doi.org/10.1038/ng1452>, PMID: 15475954
- Novo CL**, Wong E, Hockings C, Poudel C, Sheekey E, Walker S, Schierle GSK, Narlikar GJ, Rugg-Gunn PJ. 2020. Satellite repeat transcripts modulate heterochromatin condensates and safeguard chromosome stability in mouse embryonic stem cells. *bioRxiv*. DOI: <https://doi.org/10.1101/2020.06.08.139642>
- Okamura K**, Chung W-J, Ruby JG, Guo H, Bartel DP, Lai EC. 2008. The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* **453**:803–806. DOI: <https://doi.org/10.1038/nature07015>

- 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*. *Molecular Cell* **47**:954–969. DOI: <https://doi.org/10.1016/j.molcel.2012.07.021>, PMID: 22902557
- Orgel LE**, Crick FH. 1980. Selfish DNA: the ultimate parasite. *Nature* **284**:604–607. DOI: <https://doi.org/10.1038/284604a0>, PMID: 7366731
- Pal-Bhadra M**, Leibovitch BA, Gandhi SG, Chikka MR, Rao M, Bhadra U, Birchler JA, Elgin SC. 2004. Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* **303**:669–672. DOI: <https://doi.org/10.1126/science.1092653>, PMID: 14752161
- Pane A**, Jiang P, Zhao DY, Singh M, Schüpbach T. 2011. The cutoff protein regulates piRNA cluster expression and piRNA production in the *Drosophila* germline. *The EMBO Journal* **30**:4601–4615. DOI: <https://doi.org/10.1038/emboj.2011.334>, PMID: 21952049
- Parhad SS**, Tu S, Weng Z, Theurkauf WE. 2017. Adaptive evolution leads to Cross-Species incompatibility in the piRNA transposon silencing machinery. *Developmental Cell* **43**:60–70. DOI: <https://doi.org/10.1016/j.devcel.2017.08.012>
- Parhad SS**, Theurkauf WE. 2019. Rapid evolution and conserved function of the piRNA pathway. *Open Biology* **9**:180181. DOI: <https://doi.org/10.1098/rsob.180181>, PMID: 30958115
- 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 Biology* **10**:564–571. DOI: <https://doi.org/10.4161/rna.24326>, PMID: 23588056
- Peng JC**, Karpen GH. 2007. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. *Nature Cell Biology* **9**:25–35. DOI: <https://doi.org/10.1038/ncb1514>, PMID: 17159999
- 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. DOI: <https://doi.org/10.1093/genetics/121.4.765>, PMID: 2470640
- Plohl M**, Meštrović N, Mravinac B. 2012. Satellite DNA evolution. *Genome Dynamics* **7**:126–152. DOI: <https://doi.org/10.1159/000337122>, PMID: 22759817
- Plohl M**, Meštrović N, Mravinac B. 2014. Centromere identity from the DNA point of view. *Chromosoma* **123**:313–325. DOI: <https://doi.org/10.1007/s00412-014-0462-0>, PMID: 24763964
- 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. DOI: <https://doi.org/10.1261/rna.034405.112>, PMID: 22753781
- Quénerch' du E**, Anand A, Kai T. 2016. The piRNA pathway is developmentally regulated during spermatogenesis in *Drosophila*. *RNA* **22**:1044–1054. DOI: <https://doi.org/10.1261/rna.055996.116>, PMID: 27208314
- R Development Core Team**. 2017. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <http://www.r-project.org>
- Riddle NC**, Minoda A, Kharchenko PV, Alekseyenko AA, Schwartz YB, Tolstorukov MY, Gorchakov AA, Jaffe JD, Kennedy C, Linder-Basso D, Peach SE, Shanower G, Zheng H, Kuroda MI, Pirrotta V, Park PJ, Elgin SC, Karpen GH. 2011. Plasticity in patterns of histone modifications and chromosomal proteins in *Drosophila* heterochromatin. *Genome Research* **21**:147–163. DOI: <https://doi.org/10.1101/gr.110098.110>, PMID: 21177972
- Rošić S**, Köhler F, Erhardt S. 2014. Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *Journal of Cell Biology* **207**:335–349. DOI: <https://doi.org/10.1083/jcb.201404097>, PMID: 25365994
- 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. DOI: <https://doi.org/10.1261/rna.2217810>, PMID: 20581131
- Rozhkov NV**, Hammell M, Hannon GJ. 2013. Multiple roles for piwi in silencing *Drosophila* transposons. *Genes & Development* **27**:400–412. DOI: <https://doi.org/10.1101/gad.209767.112>, PMID: 23392609
- Saito K**, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC. 2006. Specific association of piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes & Development* **20**:2214–2222. DOI: <https://doi.org/10.1101/gad.1454806>, PMID: 16882972
- 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. DOI: <https://doi.org/10.1093/genetics/44.2.233>, PMID: 17247821
- 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. *Molecular Cell* **59**:553–563. DOI: <https://doi.org/10.1016/j.molcel.2015.06.024>
- Schmidt T**. 1998. Genomes, genes and junk: the large-scale organization of plant chromosomes. *Trends in Plant Science* **3**:195–199. DOI: [https://doi.org/10.1016/S1360-1385\(98\)01223-0](https://doi.org/10.1016/S1360-1385(98)01223-0)
- Schueler MG**, Higgins AW, Rudd MK, Gustashaw K, Willard HF. 2001. Genomic and genetic definition of a functional human centromere. *Science* **294**:109–115. DOI: <https://doi.org/10.1126/science.1065042>, PMID: 11588252
- Sienski G**, Dönertas 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. DOI: <https://doi.org/10.1016/j.cell.2012.10.040>, PMID: 23159368



- Sproul JS**, Khost DE, Eickbush DG, Negm S, Wei X, Wong I, Larracuenta AM. 2020. Dynamic evolution of euchromatic satellites on the X chromosome in *Drosophila melanogaster* and the simulans Clade. *Molecular Biology and Evolution* **37**:2241–2256. DOI: <https://doi.org/10.1093/molbev/msaa078>, PMID: 32191304
- Strom AR**, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. 2017. Phase separation drives heterochromatin domain formation. *Nature* **547**:241–245. DOI: <https://doi.org/10.1038/nature22989>
- Swanson EC**, Manning B, Zhang H, Lawrence JB. 2013. Higher-order unfolding of satellite heterochromatin is a consistent and early event in cell senescence. *Journal of Cell Biology* **203**:929–942. DOI: <https://doi.org/10.1083/jcb.201306073>
- Ting DT**, Lipson D, Paul S, Brannigan BW, Akhavanfard S, Coffman EJ, Contino G, Deshpande V, Iafrate AJ, Letovsky S, Rivera MN, Bardeesy N, Maheswaran S, Haber DA. 2011. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* **331**:593–596. DOI: <https://doi.org/10.1126/science.1200801>, PMID: 21233348
- Ugarkovic D**. 2005. Functional elements residing within satellite DNAs. *EMBO Reports* **6**:1035–1039. DOI: <https://doi.org/10.1038/sj.embor.7400558>, PMID: 16264428
- 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. DOI: <https://doi.org/10.1534/genetics.107.071720>, PMID: 17409066
- Verdel A**, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI, Moazed D. 2004. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**:672–676. DOI: <https://doi.org/10.1126/science.1093686>, PMID: 14704433
- Volpe TA**, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**:1833–1837. DOI: <https://doi.org/10.1126/science.1074973>, PMID: 12193640
- 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. *Molecular Cell* **59**:819–830. DOI: <https://doi.org/10.1016/j.molcel.2015.08.007>, PMID: 26340424
- Wang SH**, Elgin SC. 2011. *Drosophila* piwi functions downstream of piRNA production mediating a chromatin-based transposon silencing mechanism in female germ line. *PNAS* **108**:21164–21169. DOI: <https://doi.org/10.1073/pnas.1107892109>, PMID: 22160707
- Waring GL**, Pollack JC. 1987. Cloning and characterization of a dispersed, multicopy, X chromosome sequence in *Drosophila melanogaster*. *PNAS* **84**:2843–2847. DOI: <https://doi.org/10.1073/pnas.84.9.2843>, PMID: 3106978
- Wehr K**, Swan A, Schüpbach 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. *Developmental Biology* **294**:406–417. DOI: <https://doi.org/10.1016/j.ydbio.2006.03.002>, PMID: 16616913
- Wei KH**, Lower SE, Caldas IV, Sless TJS, Barbash DA, Clark AG. 2018. Variable rates of simple satellite gains across the *Drosophila* phylogeny. *Molecular Biology and Evolution* **35**:925–941. DOI: <https://doi.org/10.1093/molbev/msy005>, PMID: 29361128
- Wei X**. 2020. *Drosophila melanogaster* satDNA regulation repository. [https://github.com/LarracuentaLab/Dmelanogaster\\_satDNA\\_regulation](https://github.com/LarracuentaLab/Dmelanogaster_satDNA_regulation).
- Wei KH**, Chan C, Bachtrog D. 2021a. Establishment of H3K9me3-dependent heterochromatin during embryogenesis in *Drosophila miranda*. *eLife* **10**:e55612. DOI: <https://doi.org/10.7554/eLife.55612>, PMID: 34128466
- Wei X**, Eickbush DG, Speece I, Larracuenta AM. 2021b. Data: Heterochromatin-Dependent Transcription of Satellite DNAs in the *Drosophila melanogaster* Female Germline Dryad. . <https://datadryad.org/stash/dataset/10.5061/dryad.hdr7sqvj3>
- Weiler KS**, Wakimoto BT. 1995. Heterochromatin and gene expression in *Drosophila*. *Annual Review of Genetics* **29**:577–605. DOI: <https://doi.org/10.1146/annurev.ge.29.120195.003045>, PMID: 8825487
- Wickersheim ML**, Blumenstiel JP. 2013. Terminator oligo blocking efficiently eliminates rRNA from *Drosophila* small RNA sequencing libraries. *BioTechniques* **55**:269–272. DOI: <https://doi.org/10.2144/000114102>, PMID: 24215643
- 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. DOI: [https://doi.org/10.1016/0092-8674\(88\)90550-8](https://doi.org/10.1016/0092-8674(88)90550-8), PMID: 2839299
- Zhang F**, Wang J, Xu J, Zhang Z, Koppetsch BS, Schultz N, Vreven T, Meignin C, Davis I, Zamore PD, Weng Z, Theurkauf WE. 2012. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* **151**:871–884. DOI: <https://doi.org/10.1016/j.cell.2012.09.040>, PMID: 23141543
- 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. DOI: <https://doi.org/10.1016/j.cell.2014.04.030>, PMID: 24906152
- 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. DOI: <https://doi.org/10.1038/nature10371>, PMID: 21901007