PLOS GENETICS

RESEARCH ARTICLE

# Off-target piRNA gene silencing in *Drosophila melanogaster* rescued by a transposable element insertion

**Danny E. Miller[1,2,3,4☯], Ana P. Dorador[5☯], Kelley Van Vaerenberghe[5,6☯], Angela Li[5], Emily K. Grantham[5], Stefan Cerbin[5], Celeste Cummings[5], Marilyn Barragan[5], Rhonda R. Egidy[1], Allison R. Scott[1], Kate E. Hall[1], Anoja Perera[1], William D. Gilliland[7], R. Scott Hawley[1,2], Justin P. Blumenstiel[5]***

**1** Stowers Institute for Medical Research, Kansas City, Missouri, United States of America, **2** Department of Molecular and Integrative Physiology, University of Kansas Medical Center, Kansas City, Kansas, United States of America, **3** Division of Genetic Medicine, Department of Pediatrics, University of Washington and Seattle Children's Hospital, Seattle, Washington, United States of America, **4** Department of Laboratory Medicine and Pathology, University of Washington, Seattle, Washington, United States of America, **5** Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas, United States of America, **6** Division of Biological Sciences, University of Montana, Missoula, Montana, United States of America, **7** Department of Biological Sciences, DePaul University, Chicago, Illinois, United States of America

☯ These authors contributed equally to this work.
* jblumens@ku.edu





**Citation:** Miller DE, Dorador AP, Van Vaerenberghe K, Li A, Grantham EK, Cerbin S, et al. (2023) Off-target piRNA gene silencing in *Drosophila melanogaster* rescued by a transposable element insertion. PLoS Genet 19(2): e1010598. https://doi.org/10.1371/journal.pgen.1010598

**Editor:** Giovanni Bosco, Geisel School of Medicine at Dartmouth, UNITED STATES

**Received:** July 7, 2022

**Accepted:** January 4, 2023

**Published:** February 21, 2023



**Data Availability Statement:** All sequence data are available at the NCBI under BioProject PRJNA640639. All additional supporting data are within the Supporting Information Files and FlyBase (www.flybase.org). Original data underlying this manuscript can also be accessed from the Stowers Original Data Repository at: ftp://odr.stowers.org/LIBPB-1723.

**Funding:** This work was supported by the KU Genome Sequencing Core. This core lab is supported by the National Institute of General

## Abstract

Transposable elements (TE) are selfish genetic elements that can cause harmful mutations. In *Drosophila*, it has been estimated that half of all spontaneous visible marker phenotypes are mutations caused by TE insertions. Several factors likely limit the accumulation of exponentially amplifying TEs within genomes. First, synergistic interactions between TEs that amplify their harm with increasing copy number are proposed to limit TE copy number. However, the nature of this synergy is poorly understood. Second, because of the harm posed by TEs, eukaryotes have evolved systems of small RNA-based genome defense to limit transposition. However, as in all immune systems, there is a cost of autoimmunity and small RNA-based systems that silence TEs can inadvertently silence genes flanking TE insertions. In a screen for essential meiotic genes in *Drosophila melanogaster*, a truncated *Doc* retrotransposon within a neighboring gene was found to trigger the germline silencing of *ald*, the *Drosophila Mps1* homolog, a gene essential for proper chromosome segregation in meiosis. A subsequent screen for suppressors of this silencing identified a new insertion of a *Hobo* DNA transposon in the same neighboring gene. Here we describe how the original *Doc* insertion triggers flanking piRNA biogenesis and local gene silencing. We show that this local gene silencing occurs *in cis* and is dependent on *deadlock*, a component of the Rhino-Deadlock-Cutoff (RDC) complex, to trigger dual-strand piRNA biogenesis at TE insertions. We further show how the additional *Hobo* insertion leads to de-silencing by reducing flanking piRNA biogenesis triggered by the original *Doc* insertion. These results support a model of TE-mediated gene silencing by piRNA biogenesis in *cis* that depends on local determinants of transcription. This may explain complex patterns of off-target gene silencing triggered by TEs within populations and in the laboratory. It also provides a mechanism of

Medical Sciences (NIGMS) of the National Institutes of Health (NIH) under award number P20GM103638. This project was also supported by NSF MCB Awards 1413532 and 2025197 (PI: JPB), NIHNIGMS/COBRE Award P20GM103638, NIH-NIGMS/AREA Award R15GM099054 (PI: WDG) and NIH-NIGMS/KINBRE Award P20GM103418. Support to MB was provided by the NIH-NIGMS Initiative for Maximizing Student Development program at the University of Kansas (NIH 5R25GM062232). Support to CC was provided by NSF REU Program Award DBI 1560139. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

sign epistasis among TE insertions, illuminates the complex nature of their interactions and supports a model in which off-target gene silencing shapes the evolution of the RDC complex.

## Author summary

Transposable elements (TEs) are selfish DNA elements that can move through genomes and cause mutation. In some species, the vast majority of DNA is composed of this form of selfish DNA. Because TEs can be harmful, systems of genome immunity based on small RNA have evolved to limit the movement of TEs. However, like all systems of immunity, it can be challenging for the host to distinguish self from non-self. Thus, TE insertions occasionally cause the small RNA silencing machinery to turn off the expression of critical genes. The rules by which this inadvertent form of autoimmunity causes gene silencing are not well understood. In this article, we describe a phenomenon whereby a TE insertion, rather than silencing a nearby gene, rescues the silencing of a gene caused by another TE insertion. This reveals a mode of TE interaction *via* small RNA silencing that may be important for understanding how TEs exert their effects on gene expression in populations and across species.

## Introduction

Transposable elements (TE) are selfish elements that can cause DNA damage, mutation, chromosome rearrangements, and sterility. In *Drosophila*, even early investigations estimated that about half of the spontaneous mutations that cause visible phenotypes are caused by TE insertions [1]. Nonetheless, despite their harm, TEs can greatly proliferate in the genomes of sexually reproducing species [2]. A consequence of TE proliferation is that diverse systems of genome defense have evolved that limit transposition through DNA methylation, repressive chromatin, direct transcriptional repression, and small-RNA silencing. There is substantial crosstalk between these modes of genome defense. For example, small RNAs generated from harmful TE transcripts can silence TEs through cytoplasmic post-transcriptional silencing but also enter the nucleus to trigger DNA methylation and transcriptional repression [3,4].

In animals, small RNAs designated piwi-interacting RNAs (piRNAs) play a critical role in genome defense within reproductive tissues. piRNAs are derived from TE sequences recognized by the piRNA machinery and diverted from canonical mRNA processing into a piRNA generating pathway. By shunting TE transcripts toward piRNA biogenesis, the host is able to destroy TE transcripts and also generate a pool of antisense piRNAs that further repress TEs throughout the genome. Interestingly, like other systems of immunity, genomic immunity can be costly when the distinction between self and non-self is disrupted. For example, in *Arabidopsis thaliana*, selection can act against DNA-methylated TE insertions that reduce the expression of flanking genes [5]. Off-target gene silencing by systems of genome defense, and subsequent selective effects, has been observed in a variety of organisms [6–17]. However, genic silencing by flanking TEs is hardly universal within a genome. For example, in maize, the capacity to trigger the formation of flanking heterochromatin can vary significantly among TE families [13]. The cause of this variation is poorly understood.

Studies in *Drosophila*, where DNA methylation is absent and piRNAs are the primary line of defense against TEs, show that TE insertions can trigger the spreading of heterochromatin

and transcriptional silencing of genes [16–23]. In the germline, TE insertions also have the capacity to trigger the production of piRNAs from flanking sequences [24–27]. The mechanism of flanking piRNA biogenesis that spreads from a TE insertion can be explained by a general model whereby Piwi-piRNA complexes target nascent TE transcripts [28–31] followed by recruitment of the histone methyltransferase SETDB1/Egg [32–35]. Upon H3K9 methylation by SETDB1, germline TE insertions may be co-transcriptionally repressed and converted to piRNA generating loci by subsequent recruitment of the HP1 paralog Rhino [27,28,36]. In the germline, clusters of TE insertions transcribed in both directions become a source of sense and anti-sense piRNA and are known as dual-strand piRNA clusters [37,38]. Recruitment of Rhino coincides with non-canonical transcription within the TE insertion and transcripts are directed into a pathway of RNA processing that lacks standard capping, splicing and polyadenylation [27,37,39,40]. Since non-canonical transcription can ignore TE encoded termination signals [40] and extend beyond the target TE insertion, transcripts designated for piRNA processing can yield piRNAs from genomic regions outside the TE insertion. This occurs presumably through phased piRNA biogenesis [41–43], since transcripts derived from unique genomic regions flanking the TE will not be the initial target of TE-derived piRNAs that trigger ping-pong biogenesis.

In *Drosophila*, there is evidence that TEs with the capacity to induce flanking H3K9 methylation through piRNA targeting are deleterious due to the silencing of neighboring genes [14,15]. However, there is striking variation in the capacity for TE insertions to trigger these effects. Across two independent strains, only about half of euchromatic insertions show a signature of locally induced H3K9 methylation [15]. Why some TEs trigger local piRNA biogenesis and/or repressive chromatin and others do not is poorly understood, though a variety of factors are known to contribute. One factor is clearly the class of TE. In maize, only some TE families appear to induce local heterochromatin formation [13] and in *Drosophila*, the LTR class appears to exert a stronger effect on local chromatin compared to other families [15]. Such differences may be explained by regulatory sequences embedded within the particular TE family or class. For example, elements primarily expressed in somatic cells of the ovary trigger a greater degree of flanking H3K9 methylation in cultured ovarian somatic cells [29]. Additionally, TE insertions that lack a promoter and are thus not expressed can fail to trigger flanking piRNA biogenesis in the germline [44].

In the absence of regulatory sequences encoded within TE insertions, the capacity for a TE fragment to nucleate local repression and piRNA biogenesis depends on the interaction between the individual insertion and the transcriptional environment [45]. A recent investigation of flanking piRNA biogenesis triggered by transgenes showed that transcription in opposing directions (convergent transcription) may enhance conversion of TEs into standalone piRNA clusters with flanking piRNA biogenesis [26]. A recent study has also shown that zinc finger recruitment to DNA motifs can also mediate the nucleation of *Rhino* at standalone clusters [46]. However, there is no general model that explains variation in the capacity for TEs to become dual-strand piRNA clusters and why some TEs insertions have strong effects on the expression of flanking genes while others do not.

In a genetic analysis of the *Drosophila Mps1* locus, we identified TE insertions that have a complex influence on gene expression whereby insertions can either trigger local gene silencing or de-silencing. Using polyA mRNA-seq and small RNA sequencing, we show how the fate of transcripts from this locus shifts between canonical mRNA processing and piRNA biogenesis in the presence of different TE insertions. We further show that germline gene silencing induced by a *Doc* insertion depends on *deadlock*, a component of the Rhino-Deadlock-Cutoff (RDC) complex. The RDC complex plays a critical role in converting dispersed TE insertions into stand-alone dual-strand clusters within the germline [27,40,46].

Moreover, we show this gene silencing occurs in *cis*, suggesting that genic piRNAs themselves do not silence strongly in *trans*. Rather, the genic piRNAs appear primarily as a readout of silencing in *cis*. This complex effect of TE insertions supports a model in which the capacity for one TE to silence flanking genes depends on local patterns of transcription that can be altered by other TE insertions. This represents a case of compensatory mutation or sign epistasis between TE insertions, whereby the harm or benefit of an allele depends on genetic background [47].

## Results

### A DNA transposon insertion rescues a retrotransposon insertion allele of Mps1

The *Drosophila* homolog of *Mps1*, *ald*, is a conserved protein kinase that is a key component of the meiotic and mitotic spindle assembly checkpoint present in most organisms [48–51]. While *Mps1* has both mitotic and meiotic function, the *Drosophila* fully recessive $Mps1^{A15}$ allele acts only in the germline through an effect on meiosis and is caused by a *Doc* non-LTR retrotransposon insertion into the 3' end of the neighboring gene *alt*, rather than *Mps1* itself [52,53] (Fig 1). *alt* and *Mps1* are convergently transcribed with transcripts overlapping at the 3' end, a configuration that has been proposed to enable TE insertions to trigger flanking piRNA biogenesis and local gene silencing [26]. To understand why a transposon insertion in one gene could affect the function of another gene, a genetic screen was performed to identify suppressors of the *Doc* $Mps1^{A15}$ allele [54]. In this screen, seven stocks were isolated that suppressed nondisjunction caused by the $Mps1^{A15}$ allele yet retained the *Doc* insertion.

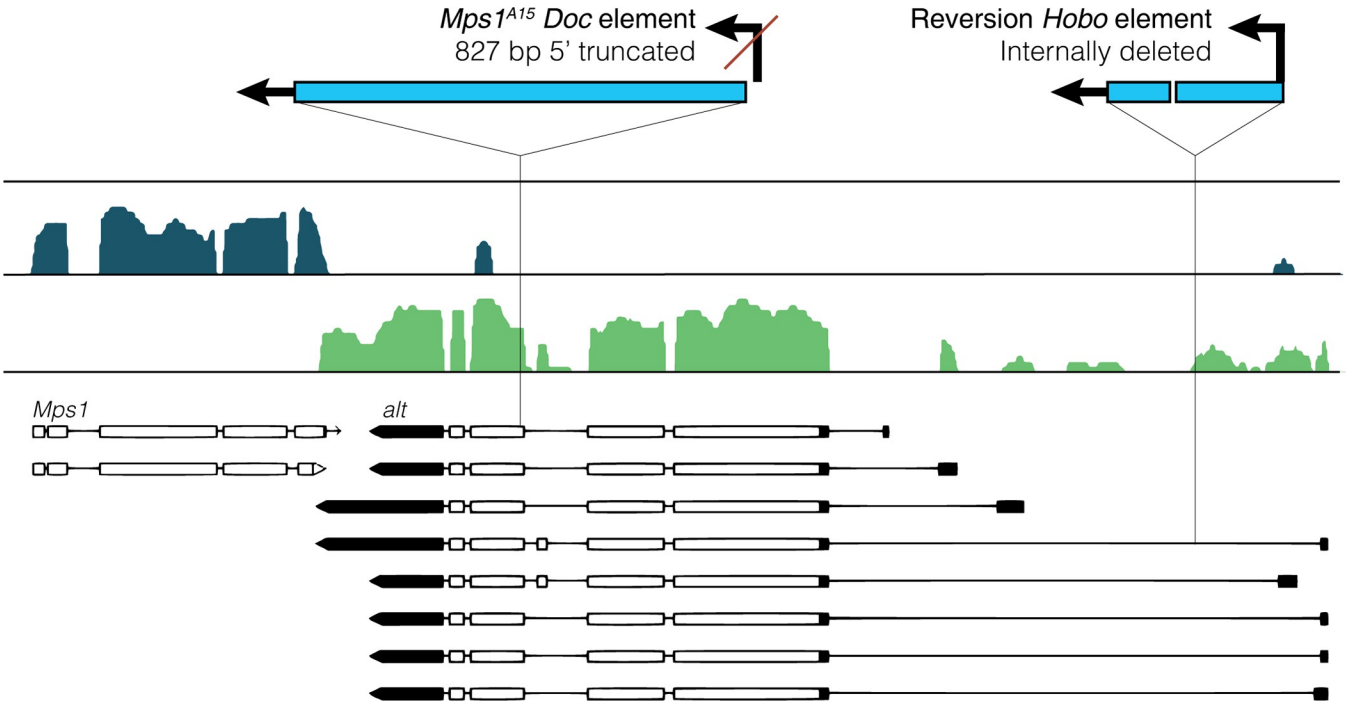



**Fig 1. Annotation of TE structure and insertion positions for $Mps1^{A15}$ and $Mps1^{A15.revertant}$ alleles within *alt*.** Transcript annotations and stranded mRNA-seq mappings from (blue[+] and green [–]density plots) from ovaries of mated females are from Flybase (www.flybase.org) and [56]. The $Mps1^{A15}$ *Doc* insertion is 5' truncated by 827 bp and located within the 4th/5th exon of *alt*, in the sense direction of the *alt* transcript. The $Mps1^{A15.rev}$ *Hobo* insertion is internally deleted, inserted in the sense direction of *alt* and located in the first intron, 1188 bp from the first *alt* TSS.

https://doi.org/10.1371/journal.pgen.1010598.g001

Subsequent genetic mapping performed on one of these stocks indicated that the revertant allele (reconfirmed rate of X chromosome non-disjunction: Control: 2/1813; $Mps1^{A15}$/$Mps1^{A15}$: 182/748; *Revertant.131/Revertant.131*: 1/557) was in close proximity to the original *Doc* insertion, so whole genome sequencing of five revertant lines was performed to identify the nature of the revertant lesion. This sequencing revealed no proximal nucleotide differences between the $Mps1^{A15}$ and revertant alleles and no differences in coverage that would be expected from gene duplication (S1 Fig), but did identify a new *Hobo* insertion within the flanking gene *alt* (Fig 1). In fact, this insertion was identified in all five revertant stocks, indicating that the identified stocks all carried the same lesion, designated $Mps1^{A15.rev}$.

PCR and Sanger sequencing confirmed the nature of the *Doc* and *Hobo* insertions (Fig 1 and S1 File). The original $Mps1^{A15}$ *Doc* insertion is 5' truncated and lacks the first 827 nucleotides containing the promoter [55]. The *Doc* is inserted within the fourth of six exons in the sense orientation with respect to *alt*, thus placing a target for germline antisense *Doc* piRNAs within the *alt* transcript. The *Hobo* insertion contains the 5' and 3' ends of the consensus *Hobo* element, but is internally deleted. Similar to the *Doc* insertion, it is in the sense orientation with respect to the *alt* transcript, but is inserted within the first intron, 1188 bp from the first TSS. Previous studies indicate that Piwi can repress gene promoter function *via* TE insertions near the TSS [29].

## Local gene silencing by a Doc insertion is ameliorated with insertion of the Hobo element

Since the $Mps1^{A15}$ allele has an effect on meiosis, but not mitosis [52], we determined how the two TE insertions influence the germline expression of flanking genes by performing polyA mRNA-seq on early 0–2 hour embryos. This approach enables an analysis of germline expression without contamination of somatic tissues of the ovary. In eggs laid by females homozygous for the $Mps1^{A15}$ allele, *Mps1* and the neighboring gene *alt* have no expression (Fig 2). Interestingly, the silencing effect of the *Doc* insertion spreads beyond *Mps1* to also cause silencing of *CG7524*, which is divergently transcribed with respect to *Mps1*, while additional genes are not affected. Thus, the *Doc* insertion into *alt* leads to the germline silencing of *alt* and two other genes. polyA mRNA-seq on embryos from $Mps1^{A15.rev}$ homozygous females indicates that the *Hobo* insertion near the 5' end of *alt* restored the expression of *Mps1* and the flanking gene *CG7524*. However, the silencing of *alt* persists. The *Hobo* insertion does not cause additional silencing of other flanking genes, such as *CG7655*, which is divergently transcribed with respect to *alt* (Fig 2).

## Altered gene expression is associated with altered genic piRNA profiles

Since TE insertions have the capacity to induce flanking piRNA biogenesis, we investigated small RNA profiles from whole ovaries in three different experiments. By comparing results with polyA mRNA-seq, we would be able to compare modes of transcript processing from this locus, between the standard RNA processing pathway that includes polyadenylation with the alternative pathway of germline dual-strand piRNA biogenesis mediated by the RDC complex that bypasses polyadenylation [27,40]. Small RNAs were classified as piRNAs based on size (23–30 nt) and showed both U-bias and ping-pong signatures (S2 Fig). In the absence of either the *Doc* or *Hobo* insertion, +/+ ovaries indicate a modest population of piRNAs derived from the site of convergent transcription between *Mps1* and *alt* (Fig 3A). Of note, these piRNAs are essentially derived from only one strand, in the sense orientation with respect to *alt* transcription. Across the entire region, there is no evidence that piRNAs are generated through bidirectional transcription since piRNAs derived from one strand do not have a corresponding

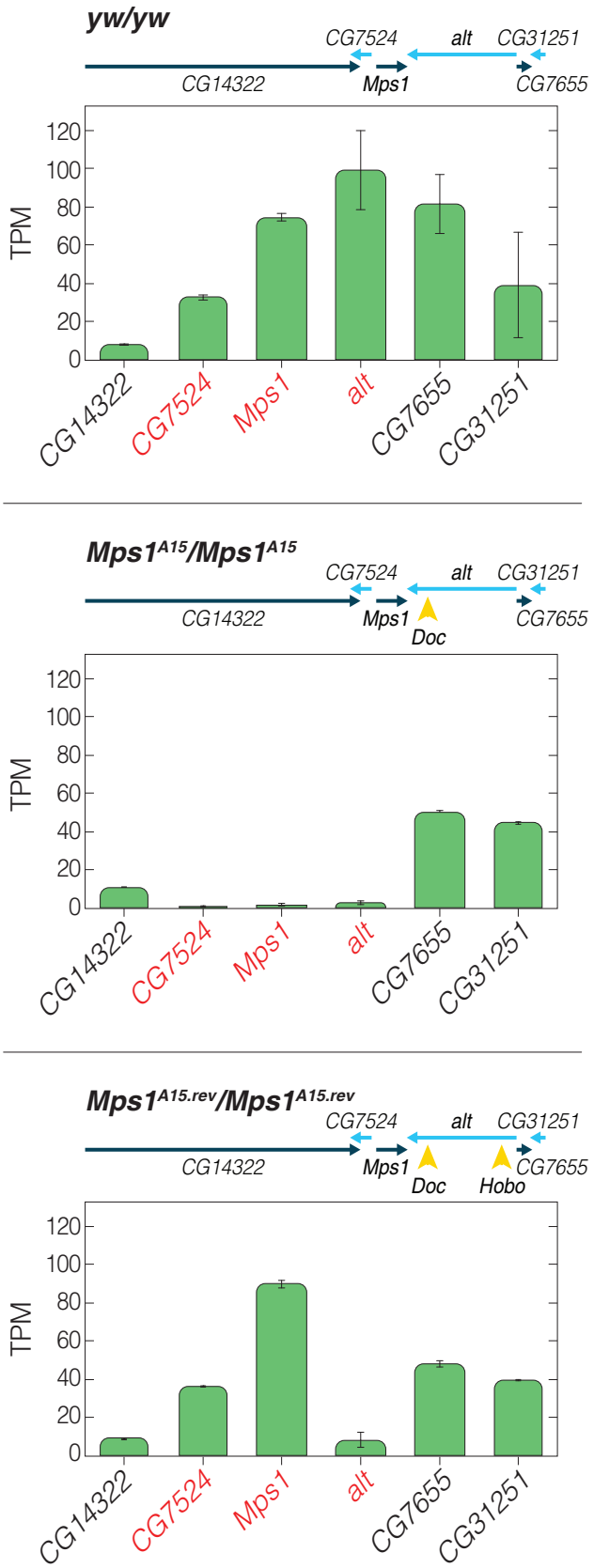

yw/yw
CG7524
alt
CG31251
CG14322
Mps1
CG7655
TPM
120
100
80
60
40
20
0
CG14322
CG7524
Mps1
alt
CG7655
CG31251
Mps1A15/Mps1A15
CG7524
alt
CG31251
CG14322
Mps1
Doc
CG7655
TPM
120
100
80
60
40
20
0
CG14322
CG7524
Mps1
alt
CG7655
CG31251
Mps1A15.rev/Mps1A15.rev
CG7524
alt
CG31251
CG14322
Mps1
Doc
Hobo
CG7655
TPM
120
100
80
60
40
20
0
CG14322
CG7524
Mps1
alt
CG7655
CG31251

**Fig 2. A *Hobo* insertion triggers germline de-silencing of two genes silenced by the $Mps1^{A15}$ *Doc* insertion.** Germline polyA mRNA TPM values from 0–2 hour embryos laid by homozygous mothers. *yw/yw* indicates the wildtype strain used for the original screen. The scheme above each graph describes which TE insertions (yellow arrowheads) were present in mothers of each experiment. Error bars are S.E. In the presence of only the $Mps1^{A15}$ *Doc* insertion, *CG7524*, *Mps1* and *alt* are silenced in the germline of $Mps1^{A15}$ homozygous mothers. In $Mps1^{A15.rev}$ homozygous mothers, the $Mps1^{A15.rev}$ *Hobo* insertion restores germline expression of *CG7524* and *Mps1* in the presence of th *Doc* insertion, but expression of *alt* is not restored. Red indicates the names of the genes affected by the *Doc* insertion.

https://doi.org/10.1371/journal.pgen.1010598.g002

population derived from the alternate strand. Thus, in the absence of TE insertions, piRNAs from this region appear to be generated through the pathway that generates sense 3'UTR genic piRNAs [57].

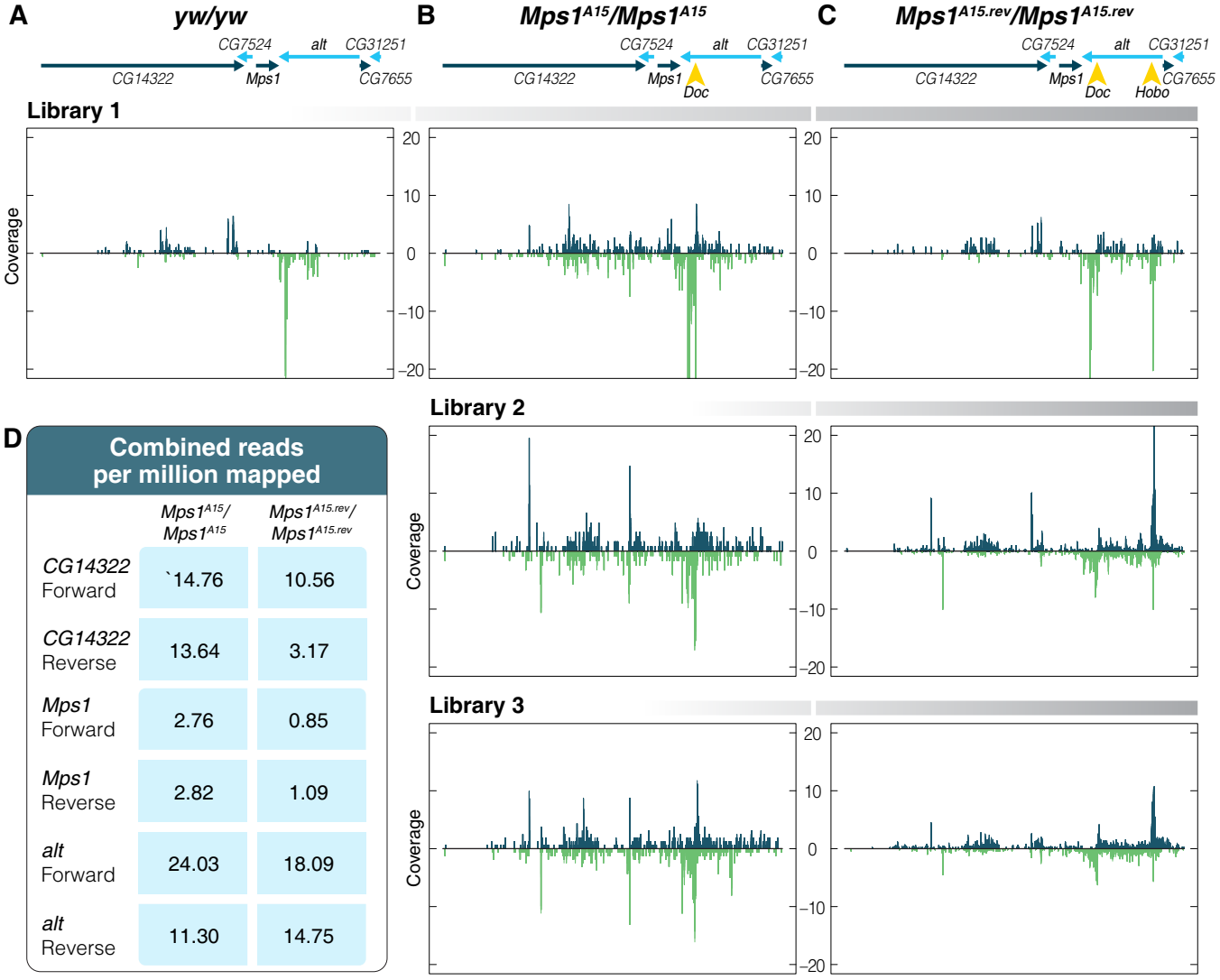



**Fig 3. *A Doc* insertion within *alt* triggers the formation of genic piRNA biogenesis from both strands which is disrupted by a *Hobo* insertion.** A) Small RNAs derived from convergently transcribed *Mps1* and *alt* are single stranded and in the same direction of the genic transcripts. B) Results across multiple libraries indicate that the *Doc* insertion of the $Mps1^{A15}$ allele triggers dual-strand piRNA biogenesis across multiple genes. C) The *Hobo* insertion of the $Mps1^{A15.rev}$ allele disrupts *Doc* triggered dual-strand genic piRNA biogenesis. D) Quantification of mapped reads from combined libraries indicates piRNAs derived from both strands of *Mps1* are reduced by approximately three-fold in $Mps1^{A15.rev}$ compared to $Mps1^{A15}$ but maintained at similar levels in both from *alt*.

https://doi.org/10.1371/journal.pgen.1010598.g003

The insertion of the *Doc* element in $Mps1^{A15}$ homozygotes changes this pattern dramatically (Fig 3B). In all three experiments (Fig 3B, Libraries 1,2,3), the *Doc* insertion is associated with conversion to piRNA biogenesis from both strands. This is evident in *alt*, where the *Doc* insertion is located, but extends across three genes on one side. Interestingly, while piRNAs derived from both strands are identified from *CG14322*, this gene is not silenced, perhaps because the 5' end of this gene is distant. 23–30 nucleotide RNAs produced from this region have a strong 5' U bias and a ping-pong signature (S2 Fig). This supports a model whereby piRNAs are generated through either phased-piRNA biogenesis on transcripts generated through bidirectional transcription from within the *Doc* element or spreading of a standalone dual-strand cluster from the *Doc* element. The ping-pong piRNA biogenesis signature supports the latter model, but there is also a weak phasing signature as noted by a modest peak of single nucleotide 3' end to 5' end distances on the plus strand. This signature is not present on the minus strand. Other dual strand clusters also have a signature of both modes of piRNA biogenesis [42].

The *Hobo* insertion changes this pattern of flanking piRNA biogenesis (Fig 3C, Libraries 1,2,3). At the site of the *Hobo* insertion near the *alt* TSS, a new population of sense and antisense flanking piRNA emerges and these dual strand piRNAs retain a ping-pong signature (S2 Fig). While there is an apparent shift in the location of piRNAs derived from *alt*, the total abundance of *alt* sense and antisense piRNA is similar in $Mps1^{A15}$ and $Mps1^{A15.rev}$ homozygotes (Fig 3A). However, the *Hobo* insertion is associated with a substantial, though incomplete, reduction of sense and antisense piRNAs derived from flanking genes. In particular, there is an approximate threefold reduction of sense and antisense piRNAs derived from the de-silenced *Mps1* (Fig 3D). Overall, while *alt* maintains a population of sense and antisense piRNAs with the *Hobo* insertion, sense and antisense piRNA biogenesis from flanking genes becomes greatly reduced. It is not apparent why the *Doc* insertion has the capacity to induce dual-strand cluster formation across multiple genes, but the *Hobo* element insertion lacks this capacity. One possibility is that endogenous piRNA abundance differs between these two elements and the abundance of piRNAs is important in triggering the formation of a dual-strand cluster. Indeed, in both strains, piRNAs that target the *Doc* element are about 10-fold more abundant than piRNAs that target the *Hobo* element (S3 Fig). This is consistent with a model in which abundant piRNAs trigger dual-strand cluster formation at the site of the *Doc* insertion, but the *Hobo* element, with fewer endogenous piRNAs, has reduced capacity to do the same.

## Doc mediated silencing in cis is dependent on the RDC complex component deadlock

As the RDC complex licenses the formation of dual-strand clusters [27], we sought to determine if the suppression of *Mps1* depended on the RDC component *deadlock* (Fig 4A). polyA mRNA-seq on ovaries revealed that the suppression of *Mps1* and *alt* is *deadlock* dependent. Low expression in *deadlock* heterozygotes that are homozygous for the $Mps1^{A15}$ allele is restored in $deadlock^{HN56}$ /$deadlock^{3}$ transheterozygotes [58–60]. We attribute the discrepancy in *CG7524* to the fact that this experiment was performed in ovaries rather than 0–2 Hour embryos (*deadlock* mutants are sterile). We also noticed through SNP analysis of polyA mRNA-seq reads that *deadlock* dependent silencing *via* the *Doc* insertion only happens in *cis* (Fig 4B). In $Mps1^{A15}$ heterozygotes, only one allele is expressed. However, in *deadlock* transheterozygotes, expression of the silenced allele is restored. Small RNA sequencing of ovaries also showed a dramatic change in local piRNA biogenesis in *deadlock* transheterozygotes. In particular, there was a striking increase in the production of small RNAs from the minus strand (S4 Fig) of the 3' end of the *alt* transcript near the site of the *Doc* insertion.

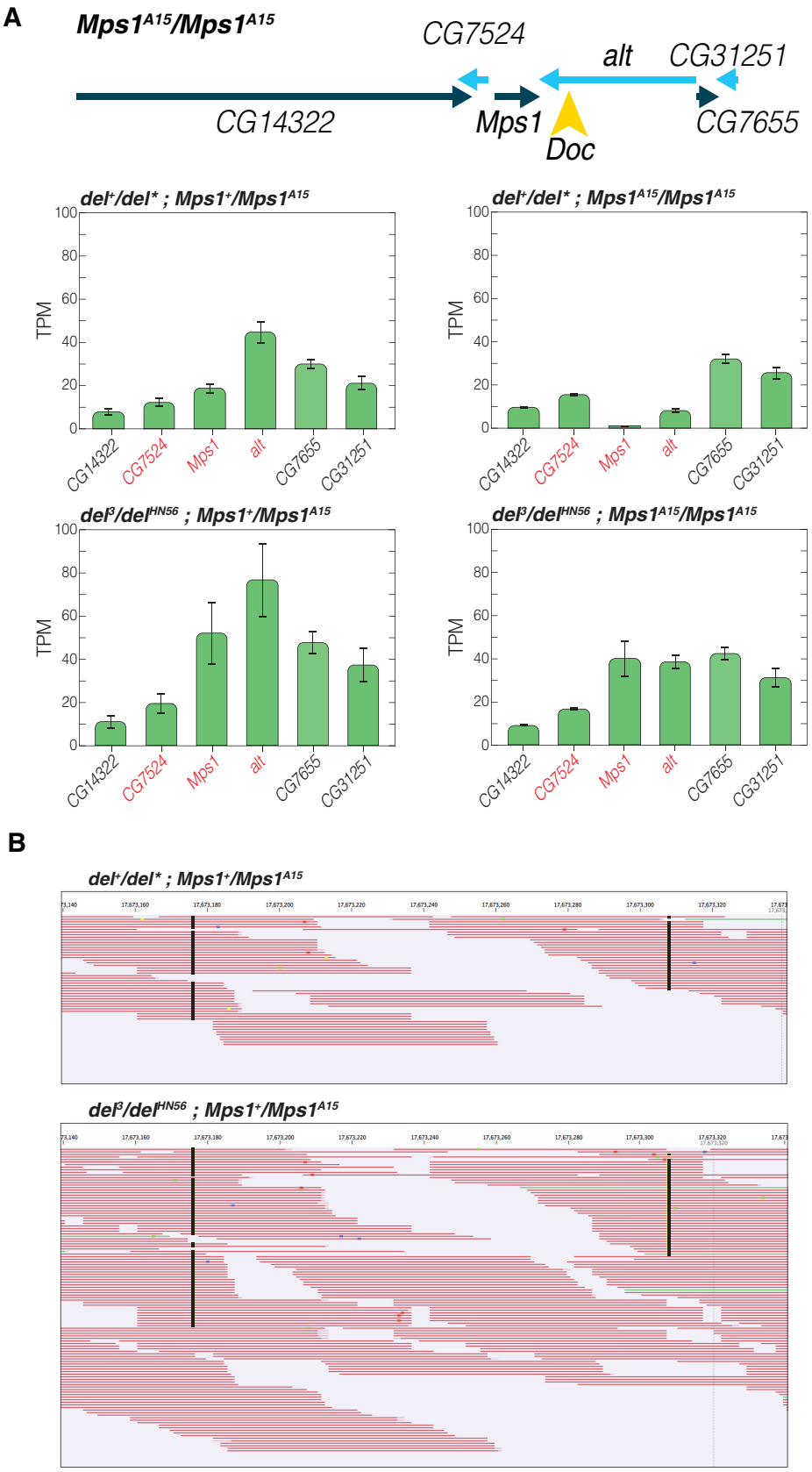

A
Mps1A15/Mps1A15
CG7524
alt
CG31251
CG14322
Mps1
Doc
CG7655
del+/del* ; Mps1+/Mps1A15
del+/del* ; Mps1A15/Mps1A15
del3/delHN56 ; Mps1+/Mps1A15
del3/delHN56 ; Mps1A15/Mps1A15
TPM
0
20
40
60
80
100
CG14322
CG7524
Mps1
alt
CG7655
CG31251
B
del+/del* ; Mps1+/Mps1A15
del3/delHN56 ; Mps1+/Mps1A15

**Fig 4. Silencing of *Mps1* by the *Doc* insertion is in *cis* and dependent on *deadlock*.** A) TPM of polyA mRNA-seq on whole ovaries. Silencing of *Mps1* and *alt* by the *Doc* insertion depends on functional *deadlock*. B) *Deadlock* dependent silencing of *Mps1* by the *Doc* insertion occurs in *cis*. In the absence of functional *Deadlock*, SNP analysis of polyA mRNA-seq reads reveals the de-silencing of the suppressed allele in $Mps1^{A15}$ / + heterozygotes.

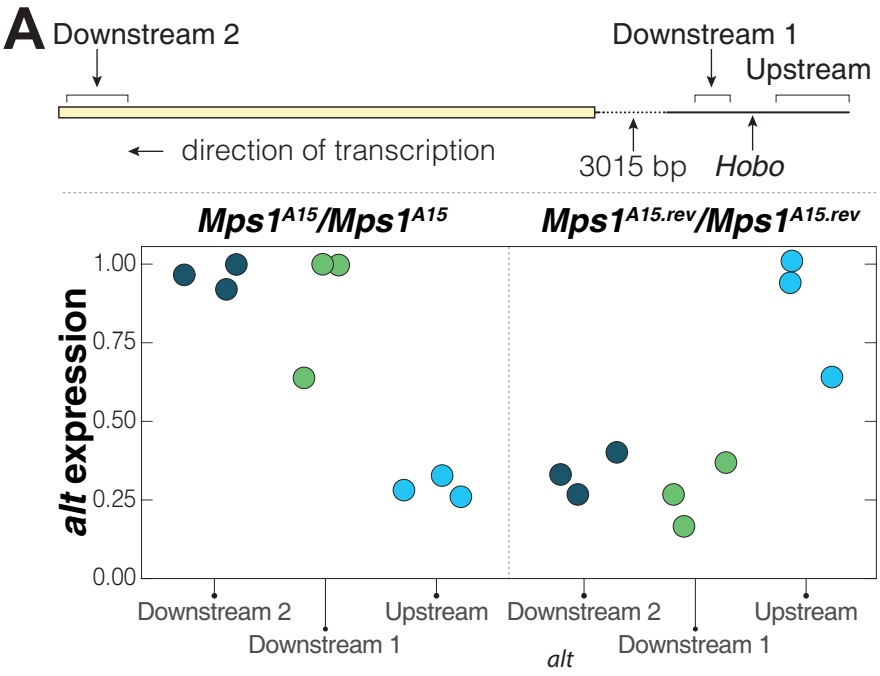
https://doi.org/10.1371/journal.pgen.1010598.g004

## The de-silencing Hobo insertion alters flanking intronic and exonic RNA expression

How does the *Hobo* element insertion alter the local landscape of dual-strand piRNA biogenesis and flanking gene silencing? One possibility is that the *Hobo* insertion triggers transcriptional silencing of *alt*. Alternatively, regulatory sequences embedded in the *Hobo* element may alter processivity of *alt* transcription, reducing the capacity for the truncated *Doc* insertion embedded in the *alt* transcript to serve as a piRNA target that nucleates the formation of a genic dual-strand cluster. In this case, the disruption of the *Doc* dual-strand cluster by *Hobo* would depend on canonical mRNA processing signals embedded in *Hobo* such as transcription termination (S1 File). This may be explained by the fact that piRNA abundance to the *Hobo* element is less than piRNA abundance to *Doc*. Fewer endogenous *Hobo* piRNAs may lack sufficient capacity to trigger the properties of a dual-strand cluster that ignore transcriptional termination signals at the *Hobo* insertion. To test the hypothesis that alteration of downstream expression of *alt* by *Hobo* mediates rescue of *Mps1* expression, we performed quantitative RT-PCR with random hexamers to quantify steady-state RNA transcript abundance of both intronic and exonic sequences flanking the *Hobo* insertion. One primer pair was located in the intron 66 nucleotides upstream of the *Hobo* insertion (S1 File), one primer pair was located in the intron 53 nucleotides downstream of the *Hobo* insertion and one primer pair was located in a downstream exon. Fig 5A shows that the insertion of the *Hobo* element has a significant effect on the expression of flanking intronic and exonic sequences ($p<0.001$ for the interaction between primer location and *Hobo* insertion). In particular, the insertion of the *Hobo* element in the same direction as *alt* leads to increased RNA abundance of intronic sequence immediately upstream of the *Hobo* transposon. This suggests that general transcriptional activity of *alt* upstream of *Hobo* is not likely to explain its effect on *Mps1* silencing. In contrast, the insertion of the *Hobo* element greatly decreases the abundance of downstream intronic and exonic RNA transcripts. These results support a model (Fig 5B) for the mechanism of how the *Hobo* insertion rescues *deadlock*-mediated *Doc* silencing of *Mps1*. In particular, lower expression of downstream RNA caused by the *Hobo* insertion renders the *Doc* element (also embedded within the *alt* transcript) as a weaker target and weaker trigger for piRNA-mediated gene silencing of flanking genes. We propose that transcription stop/polyadenylation sequences in the *Hobo* element limit RNA polymerase processivity through *alt*, though the *Hobo* element may also impact transcript stability downstream as well (See S1 File for description of putative transcription stop sequences in *Hobo*).

## Silencing and de-silencing act zygotically

piRNAs that repress TEs in *Drosophila* are transmitted maternally and maintain continuous silencing across generations [36,61–64]. piRNAs also have the capacity to maintain off-target gene silencing through maternal transmission [65,66]. This maternal transmission also can enable paramutation [67,68]. Therefore, we tested whether the silencing or de-silencing of *Mps1* depended on the maternal silencing state. This was achieved through quantitative RT-PCR of *Mps1* from ovarian mRNA collected from females generated through reciprocal crosses between wildtype, $Mps1^{A15}$ and $Mps1^{A15.rev}$ homozygotes. For each of the three pairs of crosses, reciprocal females did not show differences in the expression of *Mps1* (Fig 6). Thus,


A
Downstream 2
Downstream 1
Upstream
direction of transcription
3015 bp
Hobo
Mps1A15/Mps1A15
Mps1A15.rev/Mps1A15.rev
alt expression
1.00
0.75
0.50
0.25
0.00
Downstream 2
Upstream
Downstream 1
Downstream 2
Upstream
Downstream 1
alt

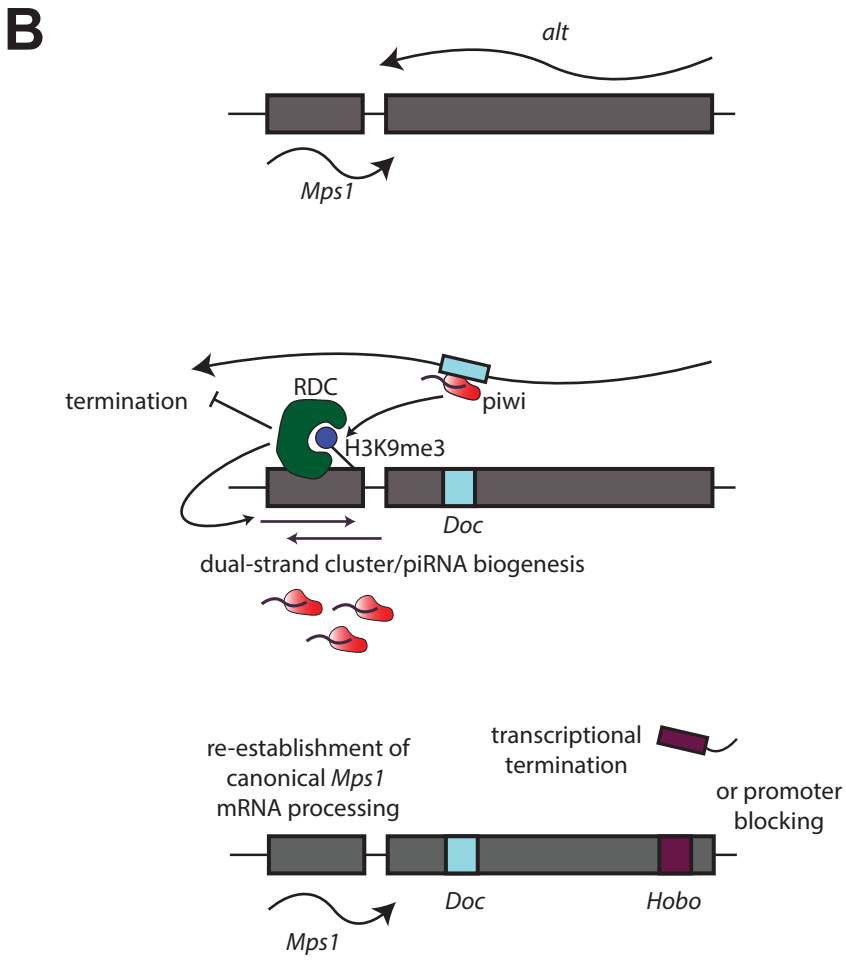

B
alt
Mps1
termination
RDC
piwi
H3K9me3
Doc
dual-strand cluster/piRNA biogenesis
re-establishment of
canonical Mps1
mRNA processing
transcriptional
termination
or promoter
blocking
Doc
Hobo
Mps1

**Fig 5. The insertion of the *Hobo* element alters local RNA expression of intronic and exonic sequence.** A) qPCR was performed on random hexamer cDNA generated from ovary total RNA. Values were normalized to *rp49* and maximum value for each amplicon made equal to 1 to enable analysis of relative expression. Upstream and Downstream 1 PCR amplicons are intronic and immediately flank the *Hobo* insertion. Downstream 2 amplicon is derived from an exon sequence. The *Hobo* insertion alters local expression by increasing transcript abundance of intronic sequence immediately upstream of the insertion, but decreases transcript abundance downstream. B) Model for disruption of *Doc* triggered silencing of *Mps1*. The presence of the *Doc* element within the *alt* transcript triggers the formation of a dual-strand cluster that spreads into *Mps1*. The *Hobo* insertion blocks processivity of the *alt* transcript, presumably through transcriptional termination. In this case, the *Doc* insertion is no longer a target for cluster formation.

https://doi.org/10.1371/journal.pgen.1010598.g005

there is no maternal effect on either silencing or de-silencing. This indicates that genic piRNAs generated from one allele neither silence nor de-silence through a maternal effect or zygotically in *trans*, though maternal *Doc* and *Hobo* piRNAs may certainly play an important role in establishing the silencing in the first place. Additionally, +/$Mps1^{A15.rev}$ genotypes have similar

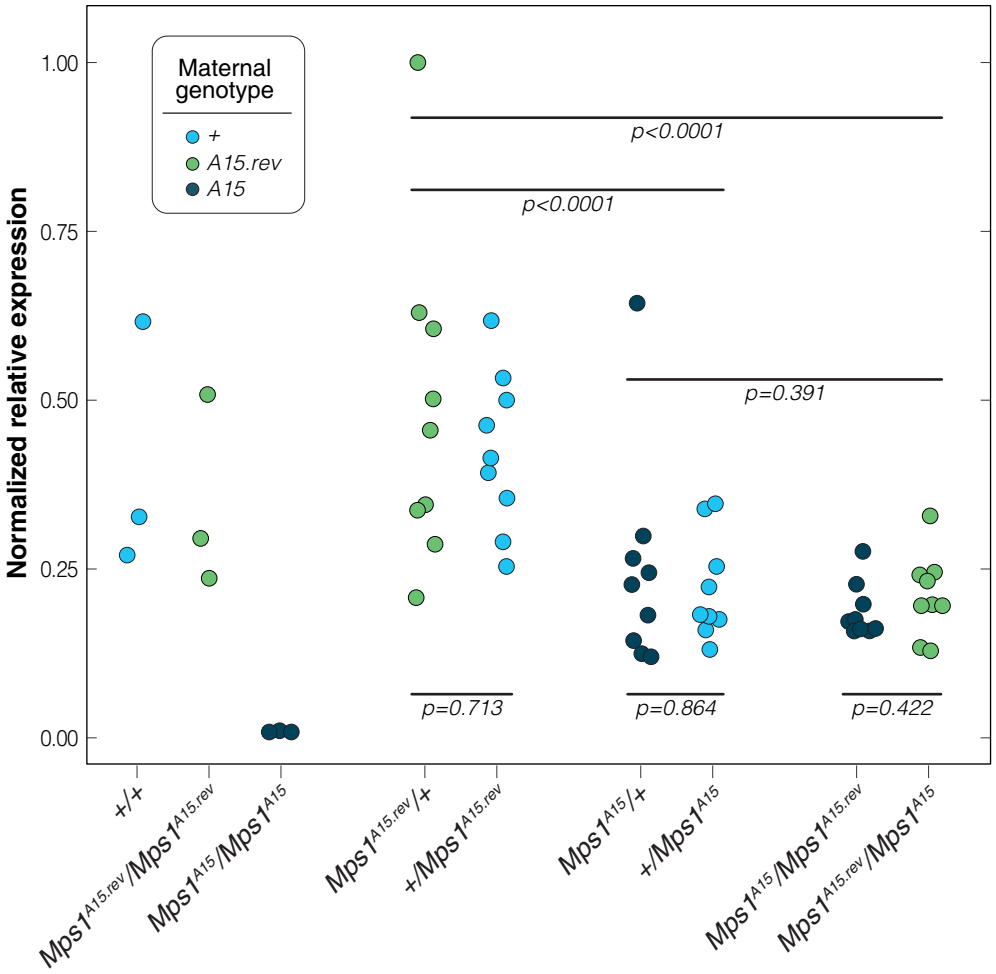



**Fig 6. Relative expression (plate normalized) of *Mps1* in ovaries is determined strictly by genotype and the dose of *A15* alleles, indicating a strict zygotic effect on gene expression.** Reciprocal progeny of all three pairwise crosses show similar expression levels, indicating no maternal effect. Expression levels of *Mps1* in $Mps1^{A15}$/+ and $Mps1^{A15}$/$Mps1^{A15.rev}$ are intermediate between $Mps1^{A15}$/$Mps1^{A15}$ homozygotes and +/+ (or $Mps1^{A15.rev}$/$Mps1^{A15.rev}$) homozygotes, indicating a strict dose effect of $Mps1^{A15}$ on *Mps1* expression. Relative expression levels are normalized to maximum 1.0.

https://doi.org/10.1371/journal.pgen.1010598.g006

expression to wildtype homozygotes and approximately twice the expression level of $+/Mps1^{A15}$ and $Mps1^{A15.rev}/Mps1^{A15}$ genotypes (Fig 6). Thus, the effect of these alleles on *Mps1* expression appears additive.

## Discussion

Systems of genome defense can maintain genome integrity through the repression of TEs, but off-target effects can lead to gene silencing. These off-target effects have been proposed to contribute to the burden of TEs themselves [5,14]. However, TE insertions near genes do not universally lead to flanking gene silencing. The underlying causes for variation in the effects of TE insertions are not well known. Here we report a pair of TE insertions where one insertion causes gene silencing and the other restores gene expression.

Silencing in *cis* of *Mps1* is dependent on the RDC complex component *deadlock*. We propose that the original gene silencing of *Mps1* caused by the *Doc* insertion in *alt* can be explained by conversion of the *Doc* element and three neighboring genes (*Mps1*, *alt* and *CG7524)* into a germline, standalone dual-strand cluster. Even though the truncated *Doc* insertion lacks a promoter, this conversion is likely explained by the insertion of a sense *Doc* target within a transcript driven by the *alt* promoter. The *Doc* fragment in sense orientation likely functions as a target for abundant endogenous *Doc* piRNAs. This conversion is perhaps enhanced by convergent transcription between *Mps1* and *alt* that produces sense piRNAs derived from the 3' end of *alt* [26]. Upon conversion to a standalone dual-strand cluster, silencing of functional transcripts is likely caused by transcripts being directed away from standard mRNA processing into a pathway of piRNA biogenesis [27,39,40]. In this case, flanking piRNA biogenesis from both strands can be considered a readout of genic co-transcriptional repression. However, it also suggests that the function of dual-strand clusters is not simply to generate a source of piRNAs but also to directly prevent individual TE insertions from producing functional transcripts. How does the *Hobo* insertion lead to de-silencing? The *Hobo* insertion still retains the *Hobo* promoter and TE insertions near gene TSS's have been shown to block PolII recruitment to genic promoters [29]. We showed increased expression of intron transcripts immediately upstream of the *Hobo* insertion, but reduced expression downstream of the insertion. Transcriptional repression or processivity of *alt* by the *Hobo* insertion thus likely precludes the *Doc* fragment from being a sufficient target for Piwi-piRNAs that can trigger conversion into a piRNA producing locus. We propose that even though transcription of dual-strand clusters frequently bypasses transcription termination signals [40], the de-silencing of *Mps1* is likely caused by transcriptional termination triggered by the *Hobo* insertion. Of note, *Hobo* piRNAs are less abundant than *Doc* piRNAs in both strains. If formation of a dual-strand cluster depends on the dose of piRNAs, the *Doc* element may have a sufficient abundance to trigger the properties of a cluster that inhibit capping, splicing and termination. The *Hobo* element, with fewer piRNAs, may retain the properties of transcriptional termination and thus abolish the *Doc* mediated cluster. Considering the cost of off-target gene silencing, some residual recognition of transcriptional termination signals within dual-strand clusters may be important to prevent excessive silencing of flanking genes. Importantly, while the *Hobo* insertion leads to a substantial reduction in the abundance of both sense and antisense *Mps1* piRNAs, flanking piRNA biogenesis is not completely blocked. In this case, one might expect that *Mps1* may be partially silenced. Nonetheless, both polyA mRNA-seq and RT-qPCR analysis reveal that the *Hobo* insertion completely restores the expression of *Mps1*. This supports the "all-or-nothing" model whereby euchromatic TEs can trigger either weak or strong, but not intermediate, silencing [24,26].

Strikingly, we found no evidence for maternal effects on the expression of *Mps1*, either for silencing alleles or de-silencing alleles. In *Drosophila*, maternal effects by piRNA play an important role in TE repression. This is revealed in syndromes of hybrid dysgenesis where paternal transmission of TEs causes excessive transposition if the mother lacks a corresponding pool of germline piRNAs. Maternal effects in TE regulation also reveal differences in how piRNA source loci depend on piRNAs for either their establishment or maintenance. Functional pericentric dual-strand clusters (such as 42AB) require maternal piRNAs but depletion of Piwi in adult ovaries does not lead to loss of cluster chromatin marks [36] Therefore, dual-strand cluster chromatin can be maintained in the absence of nuclear piRNAs. In contrast, standalone transgenes that trigger flanking piRNA biogenesis require piRNA production for maintenance of *Rhino*, *HP1* and *H3K9me3* chromatin [26]. Moreover, while maternal inheritance of *I-element* transgenes along with a substantial pool of *I-element* targeting piRNAs can trigger piRNA biogenesis from flanking regions in progeny, this mode of inheritance is not associated with altered chromatin signatures [26]. Overall, it is unclear why maternal effects and paramutation triggered by piRNAs can occur for some genes and not others [65–68].

The costs of gene silencing triggered by TEs have been proposed to shape the dynamics of TEs in populations [5,14]. However, TE insertions do not universally trigger flanking gene repression. In some cases, the expression of neighboring genes can be enhanced [69]. For example, an *Accord* LTR insertion in *Drosophila* melanogaster can enhance the expression of the cytochrome P450 gene *Cyp6g1* and provide resistance to DDT [70]. In *Drosophila simulans*, a *Doc* insertion near *Cyp6g1* has a similar effect and has been the target of positive selection [71]. Enhanced expression is attributed to the regulatory sequences carried by these elements. Here we present a case where TE insertions can alter the germline expression of a gene in opposing ways by altering the local profile of piRNA biogenesis. Formally, this represents a case of sign epistasis [47]. Since the *Hobo* insertion alone is predicted to be deleterious through *alt* silencing, but beneficial when combined with the *Doc* insertion, this satisfies the condition of sign epistasis. Theory has shown that TE dynamics within populations may not solely be influenced by their single effects, but also their epistatic interactions [72–74]. As genomic TE density increases, the likelihood for such interactions is expected to increase. In the face of a shifting landscape of genomic TE abundance, selection may fluctuate in the degree to which mechanisms of gene silencing are tolerated with respect to off-target gene silencing [75,76]. In addition to the proposed direct co-evolutionary arms races between specific TE families and mechanisms of genome defense [77,78], this global fluctuation of the TE landscape may further contribute to the patterns of positive selection observed in the piRNA machinery [79–88]. This may be especially true for components of the nuclear RDC complex where, unlike in the cytoplasm where an off-target small RNA may simply knockdown a single transcript, genomic autoimmunity can be very costly due to complete gene silencing. We propose that fluctuations in genomic TE density and abundance cause strong fluctuating selection on the processivity of non-canonical transcription and piRNA processing from within dual-strand clusters and outward to flanking genes. Likewise, the capacity for genic piRNAs to induce off-target cluster formation in *trans* may experience strong fluctuating selection. The degree to which piRNA abundance plays a role in triggering silencing, in both *cis* and *trans*, may thus be an important target of selection, when TE loads fluctuate, to avoid off-target effects that could dangerously amplify within a genome.

## Materials and methods

### Identification of the Hobo insertion Mps1$^{A15}$ revertant allele

From an EMS screen [54] to identify suppressors of the *Mps1*$^{A15}$ *Doc* insertion allele of *ald* seven stocks were identified that suppressed non-disjunction. Genetic mapping was then

performed to identify the position of the lesion. After one round of recombination with chromosomes carrying P-element *w+* insertions immediately flanking *Mps1* and *alt*, among 351 recombinant chromosomes tested, we were not able to segregate the suppressor lesion away from the $Mps1^{A15}$ lesion. This indicated that the suppressor lesion was very close to *Mps1* itself. Further analysis revealed that the *Doc* insertion was retained and no nucleotide variants were identified in this region. To identify the nearby suppressor lesion, we performed whole genome sequencing on the original $Mps1^{A15}$ stock and five of the revertant lines. DNA was prepared from homozygous males or females using the Qiagen DNeasy Blood and Tissue Kit. For each sample 500 ng of DNA was sheared to approximately 600-bp fragments using a Covaris S220 sonicator. KAPA HTP Library Prep Kit for Illumina and Bioo ScientificNEXTflex DNA barcodes were used to prepare libraries which were size selected to 500–700 bp. Libraries were quantified using a Bioanalyzer (Agilent Technologies) and a Qubit Fluorometer (Life Technologies). All libraries were pooled and sequenced as 150-bp paired-end samples on an Illumina NextSeq 500 in High-Output mode. Illumina Real Time Analysis version 2.4.11 was run to demultiplex reads and generate FASTQ files.

FASTQ files were aligned to release 6 of the *D*. *melanogaster* reference genome using bwa version 0.7.7-r441 [89]. SNPs and insertion/deletion polymorphisms were identified using SAMtools and BCFtools (version 0.1.19-44428cd) [90]. Transposable elements were identified as previously described [91]. Briefly, split and discordant read pairs were isolated using SAMBlaster [92] and individual reads were annotated using a BLAST search of the canonical *Drosophila melanogaster* transposable element database [93]. A position with multiple reads from a single TE was defined as a putative TE insertion site and was then manually analyzed. Using this approach, the *Hobo* insertion was identified. Further PCR and Sanger sequencing was used to confirm structure and insertion location within *alt*. Non-disjunction in $Mps1^{A15}$ homozygotes and rescue in $Mps1^{A15.Rev}$ homozygotes was also reconfirmed per [94].

## polyA mRNA-seq

To measure functional gene expression within the germline that is not shunted into piRNA biogenesis, we performed polyA mRNA-seq of total RNA collected from 0–2 hour old embryos laid by wildtype, $Mps1^{A15}/Mps1^{A15}$ and $Mps1^{A15.rev}/Mps1^{A15.rev}$ females. polyA mRNA-seq was also performed on ovaries of females with the following genotypes: $del^{+}/del^{*}$; $Mps1^{+}/Mps1^{A15}$, $del^{+}/del^{*}$; $Mps1^{A15}/Mps1^{A15}$, $del^{HN56}/del^{3}$; $Mps1^{+}/Mps1^{A15}$, $del^{HN56}/del^{3}$; $Mps1^{A15}/Mps1^{A15}$. Since zygotic gene expression does not begin until about two hours after egg deposition, RNA-seq from 0–2 hour embryos provides a measure of germline gene expression. In contrast, ovary RNA-seq provides expression from a mixture of germline and somatic compartments. However, this was necessary for analysis of *deadlock* mutants since such females are sterile. RNA was obtained from three different collections of pooled embryos or ovaries per genotype. polyA mRNA-seq libraries were generated from 100ng of high-quality total RNA, as assessed using the Bioanalyzer (Agilent). Libraries were made according to the manufacturer's directions for the TruSeq Stranded mRNA LT Sample Prep Kit–sets A and B (Illumina, Cat. No. RS-122-2101 and RS-122-2102). Resulting short fragment libraries were checked for quality and quantity using the Bioanalyzer (Agilent) and Qubit Fluorometer (Life Technologies). Libraries were pooled, requantified and sequenced as 75bp paired reads on a high-output flow cell using the Illumina NextSeq instrument. Following sequencing, Illumina Primary Analysis version RTA 2.4.11 and bcl2fastq2 v2.20 were run to demultiplex reads for all libraries and generate FASTQ files. TPM estimates were obtained using the CLC Genomics Workbench.

## Small RNA-seq

Small RNA-seq was performed using two approaches from RNA collected from whole ovaries of wildtype, $Mps1^{A15}/Mps1^{A15}$ and $Mps1^{A15.rev}/Mps1^{A15.rev}$ females and *deadlock* genotypes. One set of sequencing experiments were performed according to the manufacturer's directions for the TruSeq Small RNA Sample Preparation Kit (Illumina, RS-200-0012). The protocol was adapted to incorporate a 2S blocking DNA oligo for removal of prevalent Drosophila small ribosomal RNA from the sequencing library [95]. Libraries were amplified with 13 PCR cycles and resulting small RNA libraries were cut per the manufacturer's methods for 20–40 nt cDNA inserts. Short fragment libraries were checked for quality and quantity using the Bioanalyzer and Qubit Fluorometer (Life Technologies). Equal molar libraries were pooled, requantified and sequenced as 75 bp single read on the Illumina NextSeq 500 instrument using NextSeq Control Software 2.2.0.4. At least 6M reads were generated per library, and following sequencing, Illumina Primary Analysis version RTA 2.4.11 and bcl2fastq2 v2.20 were run to demultiplex reads for all libraries and generate FASTQ files. A slight modification was made for $del^{HN56}/del^{3}$; $Mps1^{A15}/Mps1^{A15}$ and $del^{+}/del^{*}$; $Mps1^{A15}/Mps1^{A15}$ libraries. Equal molar libraries were pooled, requantified and sequenced as 75 bp single read on the Illumina NextSeq 500 instrument using NextSeq Control Software 4.0.1. At least 6M reads were generated per library, and following sequencing, Illumina Primary Analysis version RTA 2.11.3.0 and bcl-convert-3.10.5 were run to demultiplex reads for all libraries and generate FASTQ files.

Additional small RNA-seq (Libraries 2 and 3 in Fig 3) was performed using an altered protocol. Pooled RNA samples were split and half the sample was oxidized [96]. RNA from each sample was ligated to 3' and 5' adapters using an rRNA blocking procedure [95] and subjected to direct reverse-transcription with unique barcoded RT primers. Barcoded RT products were pooled and size selected on a 10% acrylamide gel for the appropriate size of small RNA cDNAs (18–30 nt) appended to the additional sequence added by the adapters and RT primer. Size selection was facilitated by completing the same procedure in parallel on 18 and 30 nt RNA oligonucleotides. This procedure of pooled size selection allowed all cDNA samples to be extracted under identical conditions. Full protocol is provided in S2 File. Size selected RT products were extracted from acrylamide, subjected to 15 cycles (non-oxidized) and 18 cycles (oxidized) of PCR and sequenced.

For all small RNA sequencing experiments, reads were bioinformatically trimmed of adapters, unique molecular identifiers (6bp), size selected between 23 and 30 nt and miRNA/tRNA/rRNA depleted using the CLC Genomic Workbench or cutadapt [97] and bwa [89] and samtools [90]. Subsequent analysis revealed that the oxidation step did not significantly deplete miRNAs so these libraries are simply referred to as Libraries 2 and 3.

From small RNA analysis, 23 to 30 nt small RNA reads were mapped with bwa to release 6.33 of the *Drosophila melanogaster* reference, counts analyzed with BEDTools [98] and visualized with R. Nucleotide composition and biogenesis signatures were analyzed with the *unitas* package [99].

## Quantitative RT-PCR

RNA was collected from whole ovaries. For the analysis of the effects of the *Hobo* insertion, three independent pools of ovaries were collected for each genotype and the experiment was technically replicated for each RNA sample in triplicate. Total RNA was subjected to random hexamer reverse transcription (SuperScript IV) and qPCR performed (PowerUP SYBR Green Mastermix) on *alt* and *rp49* (*UpstreamF*: *AGC CTT TAT GAG TCA CTC CA*, *UpstreamR*: *CAA CAT GCA ATG CTG CTT TA; Downstream1F*:*TAA TGA ATG AGT GCG AGT AC,*

*Downstream1R*: *TCG CAG CAG CAT CAC TGA TAA; Downstream2F*:*AGG CTA AGC TTC GCG AAC TA*, *Downstream2R*: *GCG TCA ACT TGT CAT TCA GA; rp49F1*: *ATC GGT TAC GGA TCG AAC AA; rp49R1*: *GAC AAT CTC CTT GCG CTT CT)*. Statistical analysis and visualization was performed on averages of the three technical replicates per sample, with three samples having one of three technical replicates removed due to *rp49* failure. The linear model for expression (normalized to *rp49* expression and normalized by maximum value for each primer) in R tested for an interaction between *Hobo* insertion status and position relative to the insertion. For maternal effect analysis by RT-PCR, the experiment was performed in sets of three whereby three daughters were sampled for each of three mothers, thus providing replication across mothers of a given genotype for a total of 9 samples per genotype/mother combination. Total RNA was subjected to Oligo dT reverse transcription (NEB WarmStart RTx) and qPCR performed (NEB Luna qPCR MasterMix) on *Mps1* and *rp49* (*aldF2*: *CTG GGC TGC ATC CTT TAC CT; aldR2*: *TGG CCA TAT GAA CCA GCA TG; rp49F1*: *ATC GGT TAC GGA TCG AAC AA; rp49R1*: *GAC AAT CTC CTT GCG CTT CT)*. Each set of three daughters were analyzed on separate plates and statistical analysis was performed using a GLM model (family = gaussian) in R whereby the difference in *ald* and *rp49* Ct values were modeled as a function of plate effects, genotype effects and individual mother effects across cohorts of sisters. No significant effect of the individual mother was found, so this effect was removed from the model. However, a significant effect of plate was identified. Plate values of the difference between *ald* and *rp49* Ct were normalized based on the estimate of this plate effect and further testing was performed using a GLM model for specific comparisons of genotype and to test for maternal effects. Relative expression values are shown normalized to the maximum difference between *ald* and *rp49* Ct values, scaled according to qPCR amplification efficiency of 100%. Primer efficiency values were estimated in real-time and estimated between 98 and 102%

## Supporting information

**S1 Fig. Sequencing coverage plots of $Mps1^{A15}$ and revertant alleles.**
(PDF)

**S2 Fig. Biogenesis signatures.**
(PDF)

**S3 Fig. Abundance of *Doc*, *Hobo* and TE mapping piRNAs in $Mps1^{A15}$ and revertant strain.**
(PDF)

**S4 Fig. piRNA mappings in *deadlock* mutants.**
(PDF)

**S1 File. Sequence annotations of *Doc* and *Hobo* insertions.**
(DOCX)

**S2 File. Supplementary small RNA sequencing protocol.**
(PDF)

**S3 File. RNA-seq and qPCR figure data.**
(XLSX)

## Acknowledgments

Many thanks to Jenny Hackett of the KU Genome Sequencing Core for assistance. Thanks also to Trudi Schupbach and the Bloomington Stock center for *deadlock* alleles and Angela

Miller with help for figure preparation. Additional support was provided by the Stowers Institute for Medical Research and the University of Kansas.

## Author Contributions

**Conceptualization:** William D. Gilliland, R. Scott Hawley, Justin P. Blumenstiel.

**Data curation:** Justin P. Blumenstiel.

**Formal analysis:** Danny E. Miller, Justin P. Blumenstiel.

**Funding acquisition:** R. Scott Hawley, Justin P. Blumenstiel.

**Investigation:** Danny E. Miller, Ana P. Dorador, Kelley Van Vaerenberghe, Angela Li, Emily K. Grantham, Celeste Cummings, Marilyn Barragan, Rhonda R. Egidy, Allison R. Scott, Kate E. Hall, Anoja Perera, William D. Gilliland, Justin P. Blumenstiel.

**Methodology:** Danny E. Miller, Ana P. Dorador, Kelley Van Vaerenberghe, Justin P. Blumenstiel.

**Project administration:** Anoja Perera, Justin P. Blumenstiel.

**Resources:** R. Scott Hawley, Justin P. Blumenstiel.

**Supervision:** Anoja Perera, Justin P. Blumenstiel.

**Validation:** Ana P. Dorador, Justin P. Blumenstiel.

**Visualization:** Stefan Cerbin, Justin P. Blumenstiel.

**Writing – original draft:** Justin P. Blumenstiel.

**Writing – review & editing:** Danny E. Miller, Ana P. Dorador, Kelley Van Vaerenberghe, Angela Li, Emily K. Grantham, Stefan Cerbin, Celeste Cummings, Marilyn Barragan, Rhonda R. Egidy, Allison R. Scott, Kate E. Hall, Anoja Perera, William D. Gilliland, R. Scott Hawley, Justin P. Blumenstiel.

## References

1. Green MM. Mobile DNA elements and spontaneous gene mutation. Banbury Reports. Cold Spring Harbor Laboratory; 1988. pp. 41–50.
2. Hickey DA. Selfish DNA: A sexually-transmitted nuclear parasite. Genetics. 1982; 101: 519–531. https://doi.org/10.1093/genetics/101.3-4.519 PMID: 6293914
3. Ernst C, Odom DT, Kutter C. The emergence of piRNAs against transposon invasion to preserve mammalian genome integrity. Nat Commun. 2017; 8: 1411. https://doi.org/10.1038/s41467-017-01049-7 PMID: 29127279
4. Czech B, Hannon GJ. One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. Trends Biochem Sci. 2016; 41: 324–337. https://doi.org/10.1016/j.tibs.2015.12.008 PMID: 26810602
5. Hollister JD, Gaut BS. Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Res. 2009; 19: 1419–1428. https://doi.org/10.1101/gr.091678.109 PMID: 19478138
6. Wang X, Weigel D, Smith LM. Transposon variants and their effects on gene expression in Arabidopsis. PLoS Genet. 2013; 9: e1003255. https://doi.org/10.1371/journal.pgen.1003255 PMID: 23408902
7. Hollister JD, Smith LM, Guo Y-L, Ott F, Weigel D, Gaut BS. Transposable elements and small RNAs contribute to gene expression divergence between Arabidopsis thaliana and Arabidopsis lyrata. Proc Natl Acad Sci U S A. 2011; 108: 2322–2327. https://doi.org/10.1073/pnas.1018222108 PMID: 21252301
8. Liu J, He Y, Amasino R, Chen X. siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis. Genes Dev. 2004; 18: 2873–2878. https://doi.org/10.1101/gad.1217304 PMID: 15545622

9. Choi JY, Purugganan MD. Evolutionary Epigenomics of Retrotransposon-Mediated Methylation Spreading in Rice. Mol Biol Evol. 2018; 35: 365–382. https://doi.org/10.1093/molbev/msx284 PMID: 29126199

10. Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, et al. Role of transposable elements in heterochromatin and epigenetic control. Nature. 2004; 430: 471–476. https://doi.org/10.1038/nature02651 PMID: 15269773

11. Horvath R, Slotte T. The Role of Small RNA-Based Epigenetic Silencing for Purifying Selection on Transposable Elements in Capsella grandiflora. Genome Biol Evol. 2017; 9: 2911–2920. https://doi.org/10.1093/gbe/evx206 PMID: 29036316

12. Rebollo R, Karimi MM, Bilenky M, Gagnier L, Miceli-Royer K, Zhang Y, et al. Retrotransposon-induced heterochromatin spreading in the mouse revealed by insertional polymorphisms. PLoS Genet. 2011; 7: e1002301. https://doi.org/10.1371/journal.pgen.1002301 PMID: 21980304

13. Eichten SR, Ellis NA, Makarevitch I, Yeh C-T, Gent JI, Guo L, et al. Spreading of heterochromatin is limited to specific families of maize retrotransposons. PLoS Genet. 2012; 8: e1003127. https://doi.org/10.1371/journal.pgen.1003127 PMID: 23271981

14. Lee YC. The Role of piRNA-Mediated Epigenetic Silencing in the Population Dynamics of Transposable Elements in Drosophila melanogaster. PLoS Genet. 2015; 11: e1005269. https://doi.org/10.1371/journal.pgen.1005269 PMID: 26042931

15. Lee YCG, Karpen GH. Pervasive epigenetic effects of Drosophila euchromatic transposable elements impact their evolution. Elife. 2017;6. https://doi.org/10.7554/eLife.25762 PMID: 28695823

16. Haynes KA, Caudy AA, Collins L, Elgin SCR. Element 1360 and RNAi components contribute to HP1-dependent silencing of a pericentric reporter. Curr Biol. 2006; 16: 2222–2227. https://doi.org/10.1016/j.cub.2006.09.035 PMID: 17113386

17. Wang SH, Elgin SC. Drosophila Piwi functions downstream of piRNA production mediating a chromatin-based transposon silencing mechanism in female germ line. Proc Natl Acad Sci U S A. 2011; 108: 21164–21169. https://doi.org/10.1073/pnas.1107892109 PMID: 22160707

18. Gu T, Elgin SC. Maternal depletion of Piwi, a component of the RNAi system, impacts heterochromatin formation in Drosophila. PLoS Genet. 2013; 9: e1003780. https://doi.org/10.1371/journal.pgen.1003780 PMID: 24068954

19. Sentmanat MF, Elgin SCR. Ectopic assembly of heterochromatin in Drosophila melanogaster triggered by transposable elements. Proc Natl Acad Sci U S A. 2012; 109: 14104–14109. https://doi.org/10.1073/pnas.1207036109 PMID: 22891327

20. Pal-Bhadra M, Bhadra U, Birchler JA. RNAi related mechanisms affect both transcriptional and post-transcriptional transgene silencing in Drosophila. Mol Cell. 2002; 9: 315–327. https://doi.org/10.1016/s1097-2765(02)00440-9 PMID: 11864605

21. Robin S, Chambeyron S, Bucheton A, Busseau I. Gene silencing triggered by non-LTR retrotransposons in the female germline of Drosophila melanogaster. Genetics. 2003; 164: 521–531. https://doi.org/10.1093/genetics/164.2.521 PMID: 12807773

22. Huang Y, Shukla H, Lee YCG. Species-specific chromatin landscape determines how transposable elements shape genome evolution. bioRxiv. 2022. p. 2022.03.11.484033. https://doi.org/10.7554/eLife.81567 PMID: 35997258

23. Choi JY, Lee YCG. Double-edged sword: The evolutionary consequences of the epigenetic silencing of transposable elements. PLoS Genet. 2020; 16: e1008872. https://doi.org/10.1371/journal.pgen.1008872 PMID: 32673310

24. Olovnikov I, Ryazansky S, Shpiz S, Lavrov S, Abramov Y, Vaury C, et al. De novo piRNA cluster formation in the Drosophila germ line triggered by transgenes containing a transcribed transposon fragment. Nucleic Acids Res. 2013; 41: 5757–5768. https://doi.org/10.1093/nar/gkt310 PMID: 23620285

25. Shpiz S, Ryazansky S, Olovnikov I, Abramov Y, Kalmykova A. Euchromatic transposon insertions trigger production of novel Pi- and endo-siRNAs at the target sites in the drosophila germline. PLoS Genet. 2014; 10: e1004138. https://doi.org/10.1371/journal.pgen.1004138 PMID: 24516406

26. Komarov PA, Sokolova O, Akulenko N, Brasset E, Jensen S, Kalmykova A. Epigenetic Requirements for Triggering Heterochromatinization and Piwi-Interacting RNA Production from Transgenes in the Drosophila Germline. Cells. 2020;9. https://doi.org/10.3390/cells9040922 PMID: 32290057

27. Mohn F, Sienski G, Handler D, Brennecke J. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila. Cell. 2014; 157: 1364–1379. https://doi.org/10.1016/j.cell.2014.04.031 PMID: 24906153

28. Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, et al. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev. 2013; 27: 390–399. https://doi.org/10.1101/gad.209841.112 PMID: 23392610

29. Sienski G, Donertas D, Brennecke J. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell. 2012; 151: 964–980. https://doi.org/10.1016/j.cell.2012.10.040 PMID: 23159368

30. Rozhkov NV, Hammell M, Hannon GJ. Multiple roles for Piwi in silencing Drosophila transposons. Genes Dev. 2013; 27: 400–412. https://doi.org/10.1101/gad.209767.112 PMID: 23392609

31. Post C, Clark JP, Sytnikova YA, Chirn GW, Lau NC. The capacity of target silencing by Drosophila PIWI and piRNAs. RNA. 2014; 20: 1977–1986. https://doi.org/10.1261/rna.046300.114 PMID: 25336588

32. Sienski G, Batki J, Senti KA, Donertas D, Tirian L, Meixner K, et al. Silencio/CG9754 connects the Piwi-piRNA complex to the cellular heterochromatin machinery. Genes Dev. 2015; 29: 2258–2271. https://doi.org/10.1101/gad.271908.115 PMID: 26494711

33. Ninova M, Chen Y-CA, Godneeva B, Rogers AK, Luo Y, Fejes Tóth K, et al. Su(var)2-10 and the SUMO Pathway Link piRNA-Guided Target Recognition to Chromatin Silencing. Mol Cell. 2020; 77: 556–570. e6. https://doi.org/10.1016/j.molcel.2019.11.012 PMID: 31901446

34. Rangan P, Malone CD, Navarro C, Newbold SP, Hayes PS, Sachidanandam R, et al. piRNA production requires heterochromatin formation in Drosophila. Curr Biol. 2011; 21: 1373–1379. https://doi.org/10.1016/j.cub.2011.06.057 PMID: 21820311

35. Andreev VI, Yu C, Wang J, Schnabl J, Tirian L, Gehre M, et al. Panoramix SUMOylation on chromatin connects the piRNA pathway to the cellular heterochromatin machinery. Nat Struct Mol Biol. 2022; 29: 130–142. https://doi.org/10.1038/s41594-022-00721-x PMID: 35173350

36. Akkouche A, Mugat B, Barckmann B, Varela-Chavez C, Li B, Raffel R, et al. Piwi Is Required during Drosophila Embryogenesis to License Dual-Strand piRNA Clusters for Transposon Repression in Adult Ovaries. Mol Cell. 2017; 66: 411–419.e4. https://doi.org/10.1016/j.molcel.2017.03.017 PMID: 28457744

37. Zhang Z, Wang J, Schultz N, Zhang F, Parhad SS, Tu S, et al. The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. Cell. 2014; 157: 1353–1363. https://doi.org/10.1016/j.cell.2014.04.030 PMID: 24906152

38. Klattenhoff C, Xi HL, Li CJ, Lee S, Xu J, Khurana JS, et al. The Drosophila HP1 Homolog Rhino Is Required for Transposon Silencing and piRNA Production by Dual-Strand Clusters. Cell. 2009; 138: 1137–1149. https://doi.org/10.1016/j.cell.2009.07.014 PMID: 19732946

39. Andersen PR, Tirian L, Vunjak M, Brennecke J. A heterochromatin-dependent transcription machinery drives piRNA expression. Nature. 2017; 549: 54–59. https://doi.org/10.1038/nature23482 PMID: 28847004

40. Chen Y-CA, Stuwe E, Luo Y, Ninova M, Le Thomas A, Rozhavskaya E, et al. Cutoff Suppresses RNA Polymerase II Termination to Ensure Expression of piRNA Precursors. Mol Cell. 2016; 63: 97–109. https://doi.org/10.1016/j.molcel.2016.05.010 PMID: 27292797

41. Homolka D, Pandey RR, Goriaux C, Brasset E, Vaury C, Sachidanandam R, et al. PIWI Slicing and RNA Elements in Precursors Instruct Directional Primary piRNA Biogenesis. Cell Rep. 2015; 12: 418–428. https://doi.org/10.1016/j.celrep.2015.06.030 PMID: 26166577

42. Han BW, Wang W, Li C, Weng Z, Zamore PD. Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. Science. 2015; 348: 817–821. https://doi.org/10.1126/science.aaa1264 PMID: 25977554

43. Mohn F, Handler D, Brennecke J. Noncoding RNA. piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. Science. 2015; 348: 812–817. https://doi.org/10.1126/science.aaa1039 PMID: 25977553

44. Akulenko N, Ryazansky S, Morgunova V, Komarov PA, Olovnikov I, Vaury C, et al. Transcriptional and chromatin changes accompanying de novo formation of transgenic piRNA clusters. RNA. 2018; 24: 574–584. https://doi.org/10.1261/rna.062851.117 PMID: 29358235

45. Radion E, Sokolova O, Ryazansky S, Komarov PA, Abramov Y, Kalmykova A. The Integrity of piRNA Clusters is Abolished by Insulators in the Drosophila Germline. Genes. 2019;10. https://doi.org/10.3390/genes10030209 PMID: 30862119

46. Baumgartner L, Handler D, Platzer SW, Yu C, Duchek P, Brennecke J. The Drosophila ZAD zinc finger protein Kipferl guides Rhino to piRNA clusters. Elife. 2022;11. https://doi.org/10.7554/eLife.80067 PMID: 36193674

47. Weinreich DM, Watson RA, Chao L. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Evolution. 2005; 59: 1165–1174. PMID: 16050094

48. Fisk HA, Winey M. Spindle regulation: Mps1 flies into new areas. Curr Biol. 2004; 14: R1058–60. https://doi.org/10.1016/j.cub.2004.11.047 PMID: 15620641

49. Gilliland WD, Wayson SM, Hawley RS. The meiotic defects of mutants in the Drosophila mps1 gene reveal a critical role of Mps1 in the segregation of achiasmate homologs. Curr Biol. 2005; 15: 672–677. https://doi.org/10.1016/j.cub.2005.02.062 PMID: 15823541

50. Fischer MG, Heeger S, Häcker U, Lehner CF. The mitotic arrest in response to hypoxia and of polar bodies during early embryogenesis requires Drosophila Mps1. Curr Biol. 2004; 14: 2019–2024. https://doi.org/10.1016/j.cub.2004.11.008 PMID: 15556864

51. Winey M, Goetsch L, Baum P, Byers B. MPS1 and MPS2: novel yeast genes defining distinct steps of spindle pole body duplication. J Cell Biol. 1991; 114: 745–754. https://doi.org/10.1083/jcb.114.4.745 PMID: 1869587

52. Hawley RS, Gilliland WD. Sometimes The Result Is Not The Answer: The Truths and the Lies That Come From Using The Complementation Test. Genetics. 2006; 174: 5–15. https://doi.org/10.1534/genetics.106.064550 PMID: 16988106

53. Page SL, Nielsen RJ, Teeter K, Lake CM, Ong S, Wright KR, et al. A germline clone screen for meiotic mutants in Drosophila melanogaster. Fly. 2007; 1: 172–181. https://doi.org/10.4161/fly.4720 PMID: 18820465

54. Blumenstiel JP, Noll AC, Griffiths JA, Perera AG, Walton KN, Gilliland WD, et al. Identification of EMS-Induced Mutations in Drosophila melanogaster by Whole-Genome Sequencing. Genetics. 2009; 182: 25–32. https://doi.org/10.1534/genetics.109.101998 PMID: 19307605

55. Minchiotti G, Contursi C, Di Nocera PP. Multiple downstream promoter modules regulate the transcription of the Drosophila melanogaster I, Doc and F elements. J Mol Biol. 1997; 267: 37–46. https://doi.org/10.1006/jmbi.1996.0860 PMID: 9096205

56. Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, et al. Diversity and dynamics of the Drosophila transcriptome. Nature. 2014; 512: 393–399. https://doi.org/10.1038/nature12962 PMID: 24670639

57. Robine N, Lau NC, Balla S, Jin ZG, Okamura K, Kuramochi-Miyagawa S, et al. A Broadly Conserved Pathway Generates 3 ' UTR-Directed Primary piRNAs. Curr Biol. 2009; 19: 2066–2076. https://doi.org/10.1016/j.cub.2009.11.064 PMID: 20022248

58. Wehr K, Swan A, Schüpbach T. Deadlock, a novel protein of Drosophila, is required for germline maintenance, fusome morphogenesis and axial patterning in oogenesis and associates with centrosomes in the early embryo. Dev Biol. 2006; 294: 406–417. https://doi.org/10.1016/j.ydbio.2006.03.002 PMID: 16616913

59. Schüpbach T, Wieschaus E. Female sterile mutations on the second chromosome of Drosophila melanogaster. II. Mutations blocking oogenesis or altering egg morphology. Genetics. 1991; 129: 1119–1136. https://doi.org/10.1093/genetics/129.4.1119 PMID: 1783295

60. Czech B, Preall JB, McGinn J, Hannon GJ. A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway. Mol Cell. 2013; 50: 749–761. https://doi.org/10.1016/j.molcel.2013.04.007 PMID: 23665227

61. Kidwell MG, Kidwell JF, Sved JA. Hybrid Dysgenesis in *Drosophila melanogaster*: A Syndrome of Aberrant Traits Including Mutation, Sterility and Male Recombination. Genetics. 1977; 86: 813–833.

62. Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. An Epigenetic Role for Maternally Inherited piRNAs in Transposon Silencing. Science. 2008; 322: 1387–1392. https://doi.org/10.1126/science.1165171 PMID: 19039138

63. Ronsseray S, Lemaitre B, Coen D. Maternal inheritance of P-cytotype in *Drosophila melanogaster*—A pre-P cytotype is strictly extra-chromosomally transmitted. Mol Gen Genet. 1993; 241: 115–123.

64. Akkouche A, Grentzinger T, Fablet M, Armenise C, Burlet N, Braman V, et al. Maternally deposited germline piRNAs silence the tirant retrotransposon in somatic cells. EMBO Rep. 2013; 14: 458–464. https://doi.org/10.1038/embor.2013.38 PMID: 23559065

65. Le Thomas A, Marinov GK, Aravin AA. A transgenerational process defines piRNA biogenesis in Drosophila virilis. Cell Rep. 2014; 8: 1617–1623. https://doi.org/10.1016/j.celrep.2014.08.013 PMID: 25199836

66. Erwin AA, Galdos MA, Wickersheim ML, Harrison CC, Marr KD, Colicchio JM, et al. piRNAs Are Associated with Diverse Transgenerational Effects on Gene and Transposon Expression in a Hybrid Dysgenic Syndrome of D. virilis. PLoS Genet. 2015; 11: e1005332. https://doi.org/10.1371/journal.pgen.1005332 PMID: 26241928

67. de Vanssay A, Bouge AL, Boivin A, Hermant C, Teysset L, Delmarre V, et al. Paramutation in Drosophila linked to emergence of a piRNA-producing locus. Nature. 2012; 490: 112–115. https://doi.org/10.1038/nature11416 PMID: 22922650

68. Dorador AP, Dalikova M, Cerbin S, Stillman CM, Zych MG, Scott Hawley R, et al. Paramutation-like Epigenetic Conversion by piRNA at the Telomere of Drosophila virilis. Biology. 2022. p. 1480. https://doi.org/10.3390/biology11101480 PMID: 36290385

69. Chuong EB, Elde NC, Feschotte C. Regulatory activities of transposable elements: from conflicts to benefits. Nat Rev Genet. 2017; 18: 71–86. https://doi.org/10.1038/nrg.2016.139 PMID: 27867194

70. Chung H, Bogwitz MR, McCart C, Andrianopoulos A, Ffrench-Constant RH, Batterham P, et al. Cis-regulatory elements in the Accord retrotransposon result in tissue-specific expression of the Drosophila melanogaster insecticide resistance gene Cyp6g1. Genetics. 2007; 175: 1071–1077. https://doi.org/10.1534/genetics.106.066597 PMID: 17179088

71. Schlenke TA, Begun DJ. Strong selective sweep associated with a transposon insertion in Drosophila simulans. Proc Natl Acad Sci U S A. 2004; 101: 1626–1631. https://doi.org/10.1073/pnas.0303793101 PMID: 14745026

72. Petrov DA, Fiston-Lavier AS, Lipatov M, Lenkov K, Gonzalez J. Population Genomics of Transposable Elements in Drosophila melanogaster. Mol Biol Evol. 2011; 28: 1633–1644. https://doi.org/10.1093/molbev/msq337 PMID: 21172826

73. Charlesworth B, Langley CH. The population genetics of *Drosophila* transposable elements. Annu Rev Genet. 1989; 23: 251–287.

74. Lee YCG, Langley CH. Transposable elements in natural populations of *Drosophila melanogaster*. Philosophical Transactions of the Royal Society B-Biological Sciences. 2010; 365: 1219–1228.

75. Wang L, Barbash DA, Kelleher ES. Adaptive evolution among cytoplasmic piRNA proteins leads to decreased genomic auto-immunity. PLoS Genet. 2020; 16: e1008861. https://doi.org/10.1371/journal.pgen.1008861 PMID: 32525870

76. Kelleher ES. Protein–Protein Interactions Shape Genomic Autoimmunity in the Adaptively Evolving Rhino-Deadlock-Cutoff Complex. Genome Biol Evol. 2021;13: evab132. https://doi.org/10.1093/gbe/evab132 PMID: 34115120

77. Parhad SS, Yu T, Zhang G, Rice NP, Weng Z, Theurkauf WE. Adaptive Evolution Targets a piRNA Precursor Transcription Network. Cell Reports. 2020. pp. 2672–2685.e5. https://doi.org/10.1016/j.celrep.2020.01.109 PMID: 32101744

78. Parhad SS, Tu S, Weng Z, Theurkauf WE. Adaptive Evolution Leads to Cross-Species Incompatibility in the piRNA Transposon Silencing Machinery. Dev Cell. 2017; 43: 60–70.e5. https://doi.org/10.1016/j.devcel.2017.08.012 PMID: 28919205

79. Vermaak D, Henikoff S, Malik HS. Positive selection drives the evolution of *rhino*, a member of the Heterochromatin Protein 1 family in *Drosophila*. PLoS Genet. 2005; 1: 96–108.

80. Simkin A, Wong A, Poh YP, Theurkauf WE, Jensen JD. Recurrent and recent selective sweeps in the piRNA pathway. Evolution. 2013; 67: 1081–1090. https://doi.org/10.1111/evo.12011 PMID: 23550757

81. Lewis SH, Salmela H, Obbard DJ. Duplication and Diversification of Dipteran Argonaute Genes, and the Evolutionary Divergence of Piwi and Aubergine. Genome Biol Evol. 2016; 8: 507–518. https://doi.org/10.1093/gbe/evw018 PMID: 26868596

82. Heger A, Ponting CP. Evolutionary rate analyses of orthologs and paralogs from 12 *Drosophila* genomes. Genome Res. 2007; 17: 1837–1849.

83. Obbard DJ, Gordon KHJ, Buck AH, Jiggins FM. The evolution of RNAi as a defence against viruses and transposable elements. Philosophical Transactions of the Royal Society B-Biological Sciences. 2009; 364: 99–115. https://doi.org/10.1098/rstb.2008.0168 PMID: 18926973

84. Obbard DJ, Welch JJ, Kim K-W, Jiggins FM. Quantifying Adaptive Evolution in the *Drosophila* Immune System. PLoS Genet. 2009; 5: e1000698.

85. Kolaczkowski B, Hupalo DN, Kern AD. Recurrent Adaptation in RNA Interference Genes Across the *Drosophila* Phylogeny. Mol Biol Evol. 2010; 28: 1033–1042.

86. Yi M, Chen F, Luo M, Cheng Y, Zhao H, Cheng H, et al. Rapid evolution of piRNA pathway in the teleost fish: implication for an adaptation to transposon diversity. Genome Biol Evol. 2014; 6: 1393–1407. https://doi.org/10.1093/gbe/evu105 PMID: 24846630

87. Lee YC, Langley CH. Long-term and short-term evolutionary impacts of transposable elements on Drosophila. Genetics. 2012; 192: 1411–1432. https://doi.org/10.1534/genetics.112.145714 PMID: 22997235

88. Blumenstiel JP, Erwin AA, Hemmer LW. What Drives Positive Selection in the Drosophila piRNA Machinery? The Genomic Autoimmunity Hypothesis. Yale J Biol Med. 2016; 89: 499–512. PMID: 28018141

89. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009; 25: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324 PMID: 19451168

**90.** Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25: 2078–2079. https://doi.org/10.1093/bioinformatics/btp352 PMID: 19505943

**91.** Miller DE. Synaptonemal Complex-Deficient Drosophila melanogaster Females Exhibit Rare DSB Repair Events, Recurrent Copy-Number Variation, and an Increased Rate of de Novo Transposable Element Movement. G3: Genes, Genomes, Genetics. 2020; 10: 525–537. https://doi.org/10.1534/g3.119.400853 PMID: 31882405

**92.** Faust GG, Hall IM. SAMBLASTER: fast duplicate marking and structural variant read extraction. Bioinformatics. 2014; 30: 2503–2505. https://doi.org/10.1093/bioinformatics/btu314 PMID: 24812344

**93.** Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, et al. The Transposable Elements of the Drosophila melanogaster euchromatin: a genomics perspective. Genome Biol. 2002;3. https://doi.org/10.1186/gb-2002-3-12-research0084 PMID: 12537573

**94.** Zitron AE, Hawley RS. The genetic analysis of distributive segregation in Drosophila melanogaster. I. Isolation and characterization of Aberrant X segregation (Axs), a mutation defective in chromosome partner choice. Genetics. 1989. pp. 801–821. https://doi.org/10.1093/genetics/122.4.801 PMID: 2503421

**95.** Wickersheim ML, Blumenstiel JP. Terminator oligo blocking efficiently eliminates rRNA from Drosophila small RNA sequencing libraries. Biotechniques. 2013; 55: 269–272. https://doi.org/10.2144/000114102 PMID: 24215643

**96.** Williams Z, Morozov P, Mihailovic A, Lin C, Puvvula PK, Juranek S, et al. Discovery and Characterization of piRNAs in the Human Fetal Ovary. Cell Rep. 2015; 13: 854–863. https://doi.org/10.1016/j.celrep.2015.09.030 PMID: 26489470

**97.** Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011; 17: 10–12.

**98.** Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26: 841–842. https://doi.org/10.1093/bioinformatics/btq033 PMID: 20110278

**99.** Gebert D, Hewel C, Rosenkranz D. unitas: the universal tool for annotation of small RNAs. BMC Genomics. 2017; 18: 644. https://doi.org/10.1186/s12864-017-4031-9 PMID: 28830358