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

Danny E. Miller1,2,3,4,5, Ana P. Dorador2,6, Kelley Van Vaerenbergh6,5,6, Angela Li5, Emily K. Grantham5, Stefan Cerbin6, Celeste Cummings5, Marilyn Barragan5, Rhonda R. Egidy1, Allison R. Scott1, Kate E. Hall1, Anoja Perera5, William D. Gilliland7, R. Scott Hawley1,2, Justin P. Blumenstiel5*

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

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
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, genomic 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 TE 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 read-out 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,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; \textit{Mps1}^{A15}/\textit{Mps1}^{A15}: 182/748; Revertant.131/Revertant.131: 1/557) was in close proximity to the original \textit{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 \textit{Mps1}^{A15} and revertant alleles and no differences in coverage that would be expected from gene duplication (S1 Fig), but did identify a new \textit{Hobo} insertion within the flanking gene \textit{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 \textit{Mps1}^{A15}.\textit{rev}.

PCR and Sanger sequencing confirmed the nature of the \textit{Doc} and \textit{Hobo} insertions (Fig 1 and S1 File). The original \textit{Mps1}^{A15} \textit{Doc} insertion is 5' truncated and lacks the first 827 nucleotides containing the promoter [55]. The \textit{Doc} is inserted within the fourth of six exons in the sense orientation with respect to \textit{alt}, thus placing a target for germline antisense \textit{Doc} piRNAs within the \textit{alt} transcript. The \textit{Hobo} insertion contains the 5' and 3' ends of the consensus \textit{Hobo} element, but is internally deleted. Similar to the \textit{Doc} insertion, it is in the sense orientation with respect to the \textit{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 \textit{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 \textit{Mps1}^{A15} allele, \textit{Mps1} and the neighboring gene \textit{alt} have no expression (Fig 2). Interestingly, the silencing effect of the \textit{Doc} insertion spreads beyond \textit{Mps1} to also cause silencing of \textit{CG7524}, which is divergently transcribed with respect to \textit{alt}, while additional genes are not affected. Thus, the \textit{Doc} insertion into \textit{alt} leads to the germline silencing of \textit{alt} and two other genes. polyA mRNA-seq on embryos from \textit{Mps1}^{A15}.\textit{rev} homozygous females indicates that the \textit{Hobo} insertion near the 5' end of \textit{alt} restored the expression of \textit{Mps1} and the flanking gene \textit{CG7524}. However, the silencing of \textit{alt} persists. The \textit{Hobo} insertion does not cause additional silencing of other flanking genes, such as \textit{CG7655}, which is divergently transcribed with respect to \textit{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 \textit{Doc} or \textit{Hobo} insertion, +/+ ovaries indicate a modest population of piRNAs derived from the site of convergent transcription between \textit{Mps1} and \textit{alt} (Fig 3A). Of note, these piRNAs are essentially derived from only one strand, in the sense orientation with respect to \textit{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

Mps1\textsuperscript{A15}/Mps1\textsuperscript{A15}

Mps1\textsuperscript{A15,rev}/Mps1\textsuperscript{A15,rev}
Fig 2. A Hobo insertion triggers germline de-silencing of two genes silenced by the Mps1\textsuperscript{A15} Doc insertion. Germine polyA mRNA TPM values from 0–2 hour embryos laid by homozygous mothers. \textit{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\textsuperscript{A15} Doc insertion, CG7524, Mps1 and \textit{alt} are silenced in the germline of Mps1\textsuperscript{A15} homozygous mothers. In Mps1\textsuperscript{A15.rev} homozygous mothers, the Mps1\textsuperscript{A15.rev} Hobo insertion restores germline expression of CG7524 and Mps1 in the presence of the Doc insertion, but expression of \textit{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 \textit{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 \textit{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\textsuperscript{A15} allele triggers dual-strand piRNA biogenesis across multiple genes. C) The Hobo insertion of the Mps1\textsuperscript{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\textsuperscript{A15.rev} compared to Mps1\textsuperscript{A15} but maintained at similar levels in both from \textit{alt}.

https://doi.org/10.1371/journal.pgen.1010598.g003
The insertion of the DOC element in Mps1<sup>A15</sup> 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<sup>A15</sup> and Mps1<sup>A15</sup>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<sup>A15</sup> allele is restored in deadlock<sup>10N56</sup>/deadlock<sup>4</sup> 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<sup>A15</sup> 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.
Transposon mediated de-silencing by altered piRNA biogenesis

A. $Mps1^{A15}/Mps1^{A15}$

![Diagram showing gene expression and silencing](image)

B. $del^{+/del^{*}}; Mps1^{+}/Mps1^{A15}$

![Bar graphs showing gene expression](image)

![Sequence analysis](image)
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 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 flanking the 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, Mps1A15 and Mps1A15rev heterozygotes. For each of the three pairs of crosses, reciprocal females did not show differences in the expression of Mps1 (Fig 6). Thus,
Transposon mediated de-silencing by altered piRNA biogenesis

A

Downstream 2

\[\text{direction of transcription}\]

\[3015 \text{ bp}\]

Hobo

\begin{align*}
  & \text{Mps1}^{A15}/\text{Mps1}^{A15} \\
  & \text{Mps1}^{A15.\text{rev}}/\text{Mps1}^{A15.\text{rev}}
\end{align*}

B

\begin{align*}
  & \text{alt} \\
  & \text{Mps1}
\end{align*}

termination

RDC

\[\text{H3K9me3}\]

\[\text{piwi}\]

dual-strand cluster/piRNA biogenesis

re-establishment of

\begin{align*}
  & \text{canonical Mps1 mRNA processing} \\
  & \text{transcriptional termination}
\end{align*}

or promoter blocking

\begin{align*}
  & \text{Doc} \\
  & \text{Hobo}
\end{align*}
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, though maternal Doc and Hobo piRNAs may certainly play an important role in establishing the silencing in the first place. Additionally, +/Mps1A15.rev genotypes have similar

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

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 Mps1A15/+ and Mps1A15/A15 are intermediate between Mps1A15/A15 homozygotes and +/+ (or Mps1A15.rev/Mps1A15.rev) homozygotes, indicating a strict dose effect of Mps1A15 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
+Mps1A15 and Mps1A15 rev/Mps1A15 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 Mps1A15 revertant allele
From an EMS screen [54] to identify suppressors of the Mps1A15 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 Mps1A15 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 Mps1A15 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 Mps1A15 homozygotes and rescue in Mps1A15Rev 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, Mps1A15/Mps1A15 and Mps1A15Rev/Mps1A15Rev females. polyA mRNA-seq was also performed on ovaries of females with the following genotypes: del+/del; Mps1+/Mps1A15, del+/del; Mps1A15/Mps1A15, delHNS6/del; Mps1+/Mps1A15, delHNS6/del; Mps1A15/Mps1A15. 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, re quantified 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<sup>A15</sup>/Mps1<sup>A15</sup> and Mps1<sup>A15,rev</sup>/Mps1<sup>A15,rev</sup> 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 25 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<sup>HIN56</sup>/del<sup>HIN56</sup>; Mps1<sup>A15</sup>/Mps1<sup>A15</sup> and del<sup>+</sup>/del<sup>-</sup>; Mps1<sup>A15</sup>/Mps1<sup>A15</sup> 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 <i>Drosophila melanogaster</i> 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 <i>Hobo</i> 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 <i>alt</i> and <i>rp49</i> (<i>UpstreamF: AGC CTT TAT GAG TCA CTC CA</i>, <i>Downstream1F: TAA TGA ATG AGT GCG AGT AC</i>, <i>UpstreamR: CAA CAT GCA ATG CTG CTT TA</i>, <i>Downstream1F: TAA TGA ATG AGT GCG AGT AC</i>, <i>PLOS GENETICS</i> Transposon mediated de-silencing by altered piRNA biogenesis
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 vis-
ualization 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 replica-
tion across mothers of a given genotype for a total of 9 samples per genotype/mother combina-
tion. 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 sis-
ters. 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 Mps1A15 and revertant alleles.
(PDF)

S2 Fig. Biogenesis signatures.
(PDF)

S3 Fig. Abundance of Doc, Hobo and TE mapping piRNAs in Mps1A15 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**

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]


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


