Adaptive *P*-element insertions in a long non-coding RNA are associated with unique silencing properties

Savana Hadjipanteli[1,2], Natalie Copeland[1], Nancy Marmolejo Bustamante[1], Maureen Nweke[1], Basabdatta Adhikari[1] and Erin S. Kelleher[1*]

[1] Department of Biology and Biochemistry, University of Houston, Houston, TX 77204
[2] Penn Epigenetics Institute, Perelman School of Medicine, Philadelphia, PA 19104
[*] To whom correspondence should be addressed: eskelleher@uh.edu

ABSTRACT

Transposable elements are genetic parasites whose mobilization throughout the genome is a major source of deleterious mutations. However, some TE insertions are beneficial because they improve host fitness. Adaptive TE insertions sometimes alter the function of adjacent genes by positively and negatively impacting their expression, or by altering their encoding proteins. Alternatively, individual TE insertions can also be adaptive because they occur in piRNA clusters and lead to piRNA-mediated silencing of transposition. In a recent laboratory evolution experiment, we discovered that the long non-coding RNA *CR43651* is an adaptive insertion hot spot for *P*-element DNA transposons in *Drosophila melanogaster*. The functional effects of these insertions on *P*-element repression and *CR43651* function were unknown. In this study, we examined the effects of *CR43651* insertions on *P*-element transcriptional regulation, piRNA biogenesis, and viability. We determined that although *CR43651* is not a canonical piRNA cluster, chromosomes containing antisense *P*-element insertions in *CR43651* exhibit enhanced piRNA-like silencing in the stage of oogenesis when *P*-elements transpose, potentially explaining their adaptive benefit. We also discovered that the fitness benefit provided by *P*-element repression is offset by recessive viability effects of insertion chromosomes, potentially due to disrupted production of *mir14,* a miRNA produced from *CR43651.*

INTRODUCTION

Transposable elements (TEs) are mobile genetic parasites that replicate in germline cells. Through transposition, they cause DNA damage and produce new mutations, both of which are harmful to host fitness. In eukaryotes, piRNA clusters play a key role in defending host genomes from these fitness costs by producing the Piwi-interacting RNAs (piRNAs, 23-32 nt) that regulate TE activity and genomic spread. These specialized loci often contain insertions from multiple TE families, which are transcribed and processed into diverse piRNA species [1,2,reviewed in 3]. piRNA then enact co-transcriptional and post-transcriptional regulation of genomic targets, including homologous TE loci and TE-derived mRNA [4,5].

Transposition and the regulation of TEs by piRNAs is particularly well characterized in the ovary of the genetic model *Drosophila melanogaster* [reviewed in 6]. The *D. melanogaster* genome contains hundreds of TE families with diverse strategies for gaining access to the oocyte nucleus at different stages of oogenesis. For example, non-LTR retrotransposons hi-jack host RNA transport machinery to localize to the oocyte nucleus late in oogenesis (*I*-element [7,8]) or even later to the primordial germline of the resulting embryos (*Tahre* [9]). By contrast, DNA transposons generally mobilize in the earliest stages of oogenesis, during mitotic and meiotic S-phase [*P*-element; 10 Figure 1A]. Similar to their TE targets, piRNA cluster transcription is also dynamic over the course of oogenesis. Large numbers of piRNA clusters are transcribed in germline stem cells (GSCs), however, many of these clusters become inactivated as oogenesis progresses [11]. Furthermore, piRNA silencing for many germline TEs is transiently released in the germaria, the specialized structure within the ovary where the mitotic divisions and meiotic S-phase occurs [Figure 1A; 12]. This reduction in silencing may pose a particular challenge to the regulation of DNA transposons, because it coincides with their transposition.

P-elements are particularly well-characterized DNA transposons of *D. melanogaster*. In natural populations, *P*-elements are frequently regulated by individual *P*-element insertions in piRNA clusters located in the telomeric associated sequences (TAS) [13–15]. In a recent laboratory evolution experiment where *P*-elements were introduced into naive genomes, we discovered that piRNA silencing evolved rapidly, and was coincident with insertions into previously described TAS piRNA clusters [16]. However, we also observed a large number of insertions into a specific long non-coding RNA locus: *CR43651*. These insertion alleles were often segregating at elevated frequencies in experimental populations, indicating they confer a selective advantage. Furthermore, they lead to modest piRNA production from flanking genomic DNA in *CR43651*, suggesting they cause *CR43651* to act as a piRNA source locus in at least a subset of ovarian cells.

If *P*-element insertions in *CR43651* establish piRNA mediated silencing, especially during stages of oogenesis where *P*-elements transpose, this could explain their selective advantage. However, *CR43651* also represents a precursor transcript for *miR14*, which is involved in regulating numerous developmental and life history traits [17–19]. In this study we examined the viability effects and silencing properties of *P*-element insertions into *CR43651*. We discovered that some chromosomes carrying *CR43651* insertions exhibit enhanced piRNA production and silencing, especially in the earliest stages of oogenesis. However, *CR43651*

does not represent a canonical piRNA cluster, in that insertion into this locus alone is not sufficient to establish silencing.

RESULTS

piRNA-like silencing in association with *CR43651* extraction chromosomes.

To isolate chromosomes containing *P*-element insertions in *CR43651*, we generated second chromosome extraction lines from two experimental populations from Wang et al [16], (populations 10 and 25), in a repressive *TP5* background [14]. *TP5* is a *P*-element insertion on the X-chromosome, located in the telomere associated sequence (TAS) piRNA cluster, that establishes piRNA mediated silencing of *P*-elements. By generating chromosome extractions in a *TP5* background, we ensured that *P*-elements in balanced extraction lines (*TP5/FM6;insertion chromosome/CyO; +)* were transpositionally stable and new insertions were not produced. We identified extraction chromosomes containing *CR43651* insertions via PCR, and obtained three such chromosomes in a balanced *TP5* background: *S10-3, S10-7* and *S25-3. S10-3* and *S10-7* contain the same antisense insertion (A4 2R:9279721-) whereas *S25-3* contains a different sense insertion (A4:9279848+).

We evaluated the silencing capacity of insertion chromosomes by determining their effects on the expression of *P{lacZ}* reporters regulated by the *P*-element promoter and 5' UTR [20]. These reporter genes share 503 bp of expressed homologous sequence with native *P*-elements, therefore, their expression is sensitive to the production of *P*-element derived piRNA [21]. Furthermore, piRNA-mediated silencing of these reporters requires a repressive piRNA cluster insertion in both the maternal and zygotic genotype [22], allowing us to separate the zygotic effects of the insertion chromosomes from *TP5* by crossing *TP5/FM6; extraction chromosome/CyO* females to males with a *P{lacZ}* reporter (Figure 1B). We initially tested the *BQ16* reporter, which is expressed predominantly in germline cells in the absence of piRNAs [21,23]. We then evaluated *P{lacZ}* expression in different genotype classes of offspring, considering both the early germline cells where *P*-elements transpose (contained within the germaria) and the later stage developing egg chambers (Figure 1A).

Female offspring inheriting *TP5* but not extraction chromosomes (*w/TP5;+/CyO;BQ16/*+), exhibit moderate LacZ activity in germaria, but no LacZ activity in developing egg chambers (Figure 1C, 1D). This aligns with previous studies suggesting that insertions of *P*-element derived transgenes into TAS piRNA clusters exhibit incomplete silencing in germaria [12]. By contrast, female offspring that inherited neither *TP5* nor insertion chromosomes (*w/FM6; +/CyO; BQ16/*+) exhibited robust LacZ activity throughout the female germline, confirming the zygotic silencing effects associated with *TP5.*

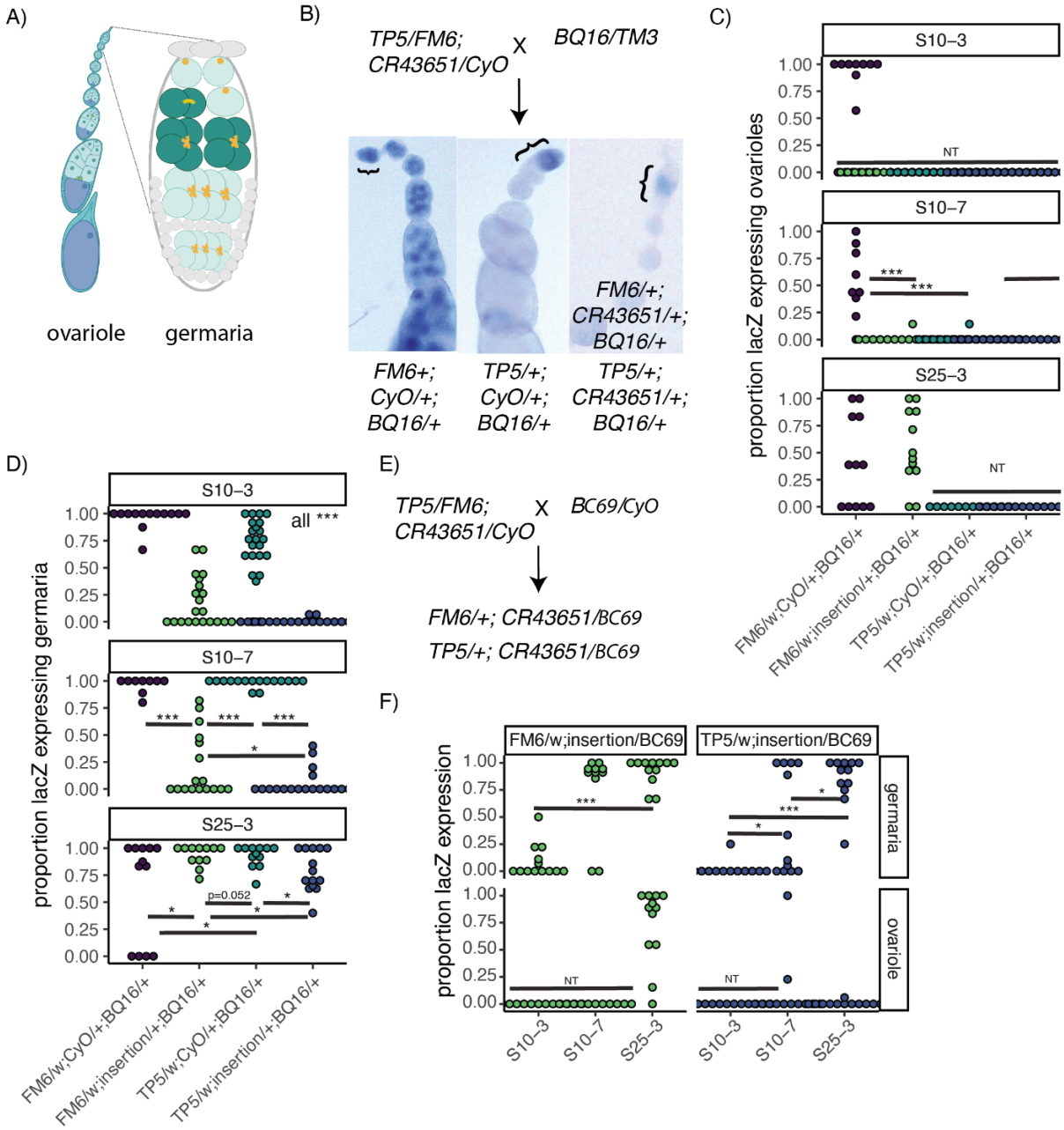

Figure 1. Silencing of *lacZ* reporters in germaria and ovarioles. A) Schematic of an ovariole and a germaria. Mitotically dividing cells in which *P*-elements transpose and piRNA silencing is reduced are highlighted in dark green. B) Crossing scheme to compare piRNA-like repression of the *BQ16 lacZ* reporter between chromosomes containing *CR43651* insertions and the *TP5* chromosome containing a single *P*-element insertion in the *X-TAS* piRNA cluster. Representative images of different offspring genotype classes are also shown. Germaria, where piRNA silencing is incomplete and more robust in the presence of *CR43651* insertion chromosomes are indicated with brackets. The proportion of ovarioles (C) and germaria (D) with strong LacZ activity is compared between zygotic genotypes of the same mother. E) Crossing scheme to compare piRNA-like repression of the *BC69 lacZ* reporter among chromosomes containing *CR43651* with and without the *TP5* chromosome. F) The proportion of germaria and ovarioles with robust LacZ activity in the offspring from E. Genotypes were compared in panels C, D, and F through one-way permutation tests. NT indicated that the mean could not be compared to other groups due to absence of variance. Significant comparisons are indicated as follows: * P > 0.05, ** P > 0.01 *** P > 0.001. Data anderlying panels, D, E, and F are provided in S1 and S2 Data. Panel A was created with BioRender.com

Surprisingly, we observed that extraction chromosomes *S10-3* and *S10-7* also exhibit zygotic silencing effects, as flies that inherited these extraction chromosomes, but not *TP5* (*FM6/w; insertion chromosome/CyO*), exhibited strong repression of *P{lacZ}* (Figure 1B-1D). Furthermore, *S10-3* and *S10-7* were associated with significantly stronger silencing in germaria than *TP5*, with the majority of germaria exhibiting weak or no LacZ activity (Figure 1B,1C). By contrast, *S25-3* exhibited no silencing in germaria or ovarioles, as LacZ activity in these females did not differ from *w/FM6; +/CyO; BQ16/+* controls. Together our data suggest that *S10-3* and *S10-7* exhibit enhanced silencing in germaria when compared to chromosomes containing insertions in canonical piRNA clusters (*TP5*). Given that *P*-elements mobilize predominantly in pre-meiotic and meiotic stages of oogenesis, which occur in germaria, this effect would likely decrease *P*-element transposition rates, conferring a selective advantage.

Previous studies suggest that the *P{lacZ}* silencing capacity of *P*-element insertions in *TAS* piRNA clusters is amplified by additional *P*-element copies throughout the genome [24,25], most likely because they feed forward piRNA biogenesis through ping-pong amplification [26–28]. Consistent with this model, the most robust silencing is observed in flies inheriting both *TP5* and an insertion chromosome (*TP5/w; insertion chromosome/CyO*). Critically however, this effect could only be detected in germaria. Furthermore the additional silencing that occurs is most robust when *TP5* is combined with *S10-3* and *S10-7*. *S25-3* is associated with minimal silencing of *BQ16* in germaria, even in combination with *TP5.*

To establish whether silencing associated with *S10-3* and *S10-7* is general to germline expressed *P{lacZ}* reporters, we compared the ability of insertion chromosomes to silence the expression of a second reporter, *BC69. BC69* exhibits higher germline expression levels than *BQ16* and is less sensitive to piRNA-like silencing [22,23]. Because *BC69* is homozygous sterile and balanced by *CyO*, we could only compare silencing differences associated with the 3 insertion chromosomes in the presence and absence of *TP5* (Figure 1E,2F). Similar to our results with *BQ16, S10-3* exhibited strong silencing in germaria regardless of whether TP5 was present or absent, whereas *S25-3* exhibited minimal silencing in germaria. *S10-7* exhibited an intermediate phenotype, in which germaria silencing did not differ from *S25-3* in the absence of *TP5*, however, silencing in germaria was stronger when *S10-7* was combined with *TP5.* These observations support our conclusion that *S10-3* and *S10-7*, but not *S25-3* confer piRNA like silencing. They further suggest that *S10-3* may be a stronger silencer in germaria than *S10-7*.

Germline silencing of *lacZ* reporters is associated with enhanced *P*-element derived piRNA production.

To determine whether enhanced silencing of *P{lacZ}* reporters was associated with increased piRNA production, we compared the piRNA pools of *TP5* and *TP5;insertion chromosome* homozygotes (Figure 2). Consistent with their *P{lacZ}* silencing behaviors, *P*-element derived piRNA were more abundant in genotypes including insertion chromosomes, with the greatest piRNA abundance occurring in the *S10-3* and *S10-7* containing genotypes (Figure 2A,B). Amplified piRNA production in the presence of additional *P*-element copies is predicted to occur regardless of whether those copies are contained within piRNA clusters, because sense *P*-element mRNAs feed forward piRNA production through the ping-pong

amplification loop [29,30]. Consistent with this, we observed that the fraction of piRNAs arising from the antisense strand decreased in the presence of all three insertion chromosomes, while the ping-pong fraction increased (Figure 2B,2C).

Increased ping-pong alone does not appear to explain the strong silencing of *S10-3* and *S10-7* in germaria (Figure 1). The ping-pong fractions and antisense fractions are similar for all three insertion chromosomes, while overall piRNA abundance increases a lot more in the presence of *S10-3* and *S10-7* (FIgure 2A-2C). This suggests that *S10-3* and *S10-7* must have additional effects on piRNA biogenesis beyond providing additional *P*-element copies that feed forward ping-pong amplification.

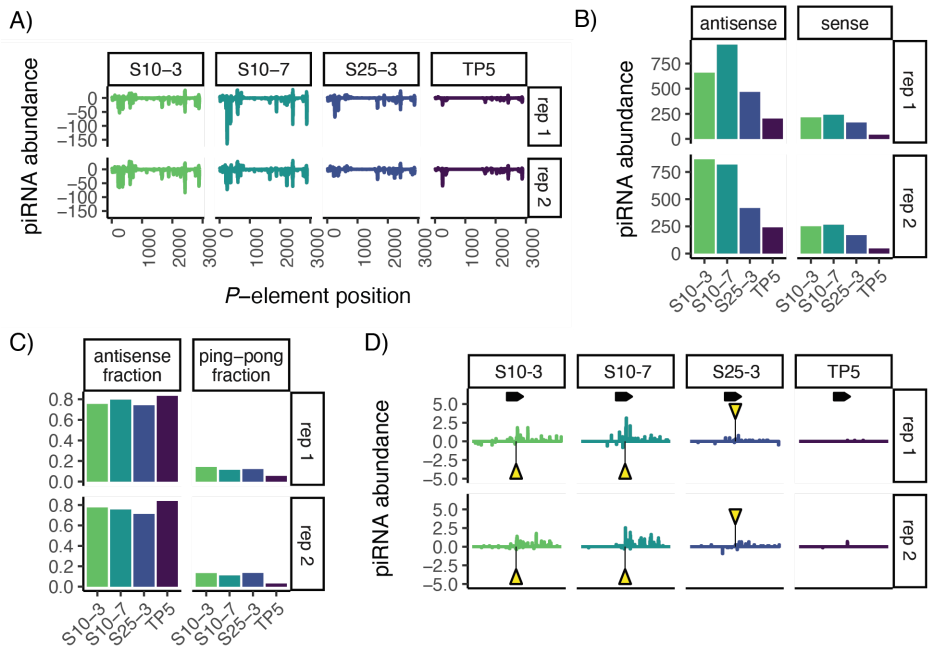

Figure 2. Insertion chromosomes amplify *P*-element derived piRNA abundance. A) Sliding window analysis of piRNAs (23-32 nt) mapping to *P*-elements. B) Abundance of piRNAs from sense and antisense strand. C) Antisense and ping-pong fractions of *P*-element derived piRNAs. All insertion chromosomes are in a balanced TP5 background (*TP5/FM6;insertion chromosome/CyO*). D) Sliding window analysis of piRNAs (23-32 nt) surrounding the insertion site in *CR43651*. Yellow triangles denote insertion positions and strands.Two biological replicates are shown for each genotype. small RNA sequencing underlying all panels are in PRJNA1292797.

CR43651 is not a piRNA cluster since in the absence of *P*-element insertions into this locus, no piRNA are produced from this region [Figure 2D; 16]. However, euchromatic TEs are often converted to piRNA source loci in the presence of piRNA silencing, leading to piRNA production from DNA adjacent to the insertion site [31]. We therefore compared the capacity for *P*-element insertions in *S10-3, S10-7* and *S25-3* to establish piRNA production in *CR43651*. We observed that piRNAs are produced from *CR43651* surrounding the *P*-element insertions in all three extraction chromosomes (Figure 2D). However, more piRNAs are produced surrounding the antisense insertion in *S10-3* and *S10-7,* consistent with the overall increase in *P*-element piRNAs and piRNA-like silencing in these genotypes.

Position-dependent piRNA production.

Our observations that *S10-3* and *S10-7* act as strong zygotic *P{lacZ}* silencers, and strongly enhance piRNA production, while *S25-3* does not exhibit zygotic silencing and only modestly enhances piRNA production (Figures 1,2) could suggest the impact of *P*-element insertions in *CR43651* on repression are position-dependent. In particular, the *P*-element in

S10-3 and *S10-7* occurs antisense to the *CR43651* promoter, potentially driving convergent transcription that facilitates piRNA biogenesis [32,33]. However, it is also possible that the extracted second chromosomes contain other *P*-element insertions, outside of *CR43651,* which explain their unique silencing effects. We therefore performed deep sequencing of *S10-3*, *S10-7* and *S25-3* genotypes to uncover additional insertions that may contribute to or explain enhanced silencing.

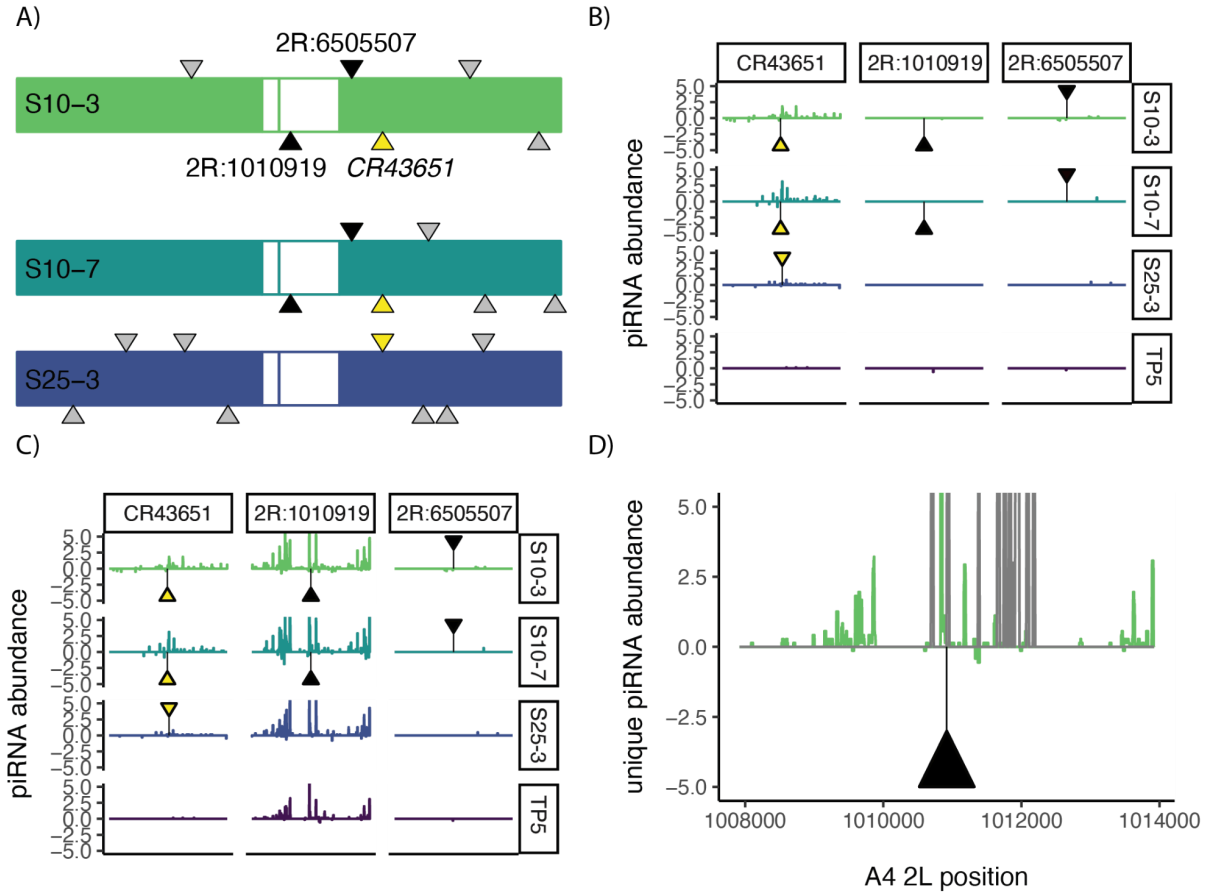

Figure 3. A) Second chromosome maps of *S10-3*, *S10-7* and *S25-3*. Location of *P*-element insertions are indicated by triangles. Yellow triangles highlight insertions in *CR43651* while black triangles indicate two additional insertions shared between *S10-3* and *S10-7*. A single insertion on 3L in *S10-3* is not shown. B and C) Sliding window analysis of uniquely mapping (B) and all (C) piRNAs aligning to the 6 Kb window surrounding the shared insertions. D) All piRNA from *S10-3* (green) are compared to theoretical uniquely mapping piRNA (grey). For B,C one biological replicate is shown, the other is shown in S1 Figure. Genome and small RNA sequencing underlying all panels are in PRJNA1292797.

We determined that silencing differences between insertion chromosomes are not explained by *P*-element dosage, as *S10-3*, *S10-7* and *S25-3* genotypes contain 6,5, and 7 additional *P*-element insertions, along with their known insertions in *CR43651* and *X-TAS* (supplementary table 1, Figure 3A). Consistent with their origins as second chromosome extraction lines into *TP5*, all insertions occurred on the second chromosome with the exception of the *TP5* insertions on the X chromosome and a single 3L insertion in *TP5/FM6;S10-3/CyO* (Figure 3A). We also discovered that *S10-3* and *S10-7* share two additional *P*-element

insertions on 2L, along with the antisense *P*-element in *CR43651*. We therefore critically evaluated whether these other shared insertions, and not the antisense insertion in *CR43651* were potentially responsible for silencing.

No insertions on any of the three extraction chromosomes occur in major piRNA clusters, however, 8 insertions occurred in regions previously identified as modest piRNA source loci in some genetic backgrounds [11,32,Supplementary table 1; 34,35]. Three of these insertions correspond to those that are shared by *S10-3* and *S10-7* (*CR43651,* 2R:1010919, and 2R:6505507). Of these three, only *CR43651* exhibits insertion-dependent piRNA production, consistent with its conversion into a piRNA cluster (3B-C). In contrast, the other two insertions exhibited no uniquely mapping piRNA regardless of the presence or absence of a *P*-element insertion (Figure 3B). Furthermore, although we observe abundant non-uniquely mapping piRNA in the 6 Kb window surrounding the insertion 2R:1010919-, the absence of uniquely mapping piRNAs that align to this region suggest these non-unique piRNA are likely derived from elsewhere in the genome (Figure 3B,3C). Indeed, by comparing all theoretical 25 mers from this 6 Kb window to the A4 assembly, we identified numerous potential unique piRNAs, which would be produced if this locus were piRNA source locus Figure (3D). Taken together, our observations suggest that of the three insertions shared between *S10-3* and *S10-7,* only the *CR43651* insertion is associated with piRNA production.

Kipferl-independent Rhino affinity for *CR43651*.

Dual-strand piRNA clusters are defined by their association with the HP1 homolog Rhino, which promotes bi-directional transcription and piRNA production [36]. Consistent with our understanding that *CR43651* is not a piRNA cluster in the traditional sense, previously published ChIP-seq data do not suggest that Rhino is frequently bound to *CR43651* in wild-type ovaries [Figure 4; 37]. However, Rhino-specified piRNA producing loci are proposed to be developmentally dynamic, with some piRNA producing sites being active only at the earliest stages of oogenesis [11,37]. piRNA producing sites in early oogenesis may be particularly significant to *P*-element control, as *P*-elements mobilize during S-phase in mitotic and meiotic germline cells.

The Zinc-finger protein Kipferl accumulates starting at the two cell stage in the pre-meiotic female germline, and is proposed to stabilize Rhino at a subset of piRNA clusters in developing egg chambers [37]. Therefore, piRNA clusters that are active in the earliest stages of oogenesis are revealed in *kipferl* mutants and germline knock-downs. Consistent with a model in which *CR43651* corresponds to a piRNA cluster that is active in early oogenesis, we observed that in published ChIP data from *kipferl* trans-heterozygous mutants, and to a lesser degree *kipferl* germline knockdowns, Rhino protein is is enriched at *CR43651* [Figure 4; 37]. These observations suggest that Rhino may be recruited to *CR43651* in early oogenesis, but is not stably maintained at this locus in later stages after Kipferl is activated.

***P-element* transgenes in *CR43651* do not exhibit piRNA-like silencing.**

Canonical piRNA clusters are characterized by a capacity to silence reporter genes *in trans* if DNA sequence homologous to the reporter is introduced into the piRNA cluster [21,23]. This

silencing reflects the production of complementary piRNA from the cluster, which target homologous RNA for co-transcriptional and post-transcriptional silencing [38]. For example, *P*-element derived transgenes located in TAS piRNA clusters are sufficient to establish silencing of *P{lacZ}* reporters in euchromatic regions [21,39].

To determine if *CR43651* acts like a canonical piRNA cluster, regulating any sequences homologous to those inserted into this locus, we examined the capacity of 6 *P*-element derived transgenes into *CR43651* to silence *BQ16* in the ovary (F1 offspring of insertion females x *BQ16* males, Figure 5A, S2 Table). Surprisingly, we observed that neither sense nor antisense *P*-element transgenic insertions in *CR43651* establish ovarian silencing of *BQ16* (Figure 5B-D). Since *P*-element derived transgenes and native *P*-elements exhibit equivalent sequence homologous to *P{lacZ}* (positions 84-586 of the native *P*-element), or more homologous sequence if the transgene also contains *lacZ*, these observations cannot be attributed to reduced homologous piRNA from the transgenes themselves. Rather they suggest that antisense transcription from a *P*-promoter in *CR43651* alone is not sufficient to establish piRNA silencing of euchromatic targets like *BQ16*.

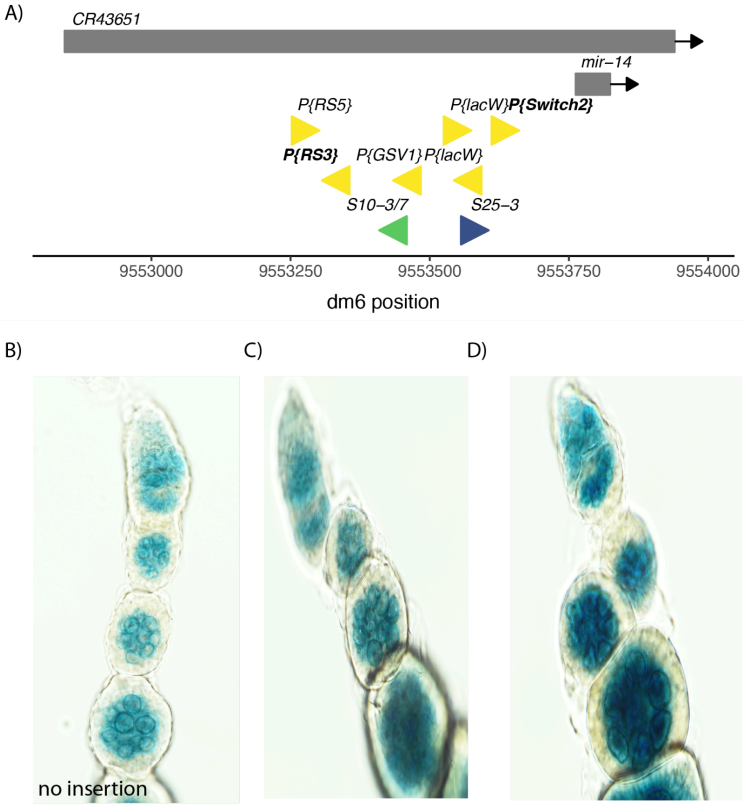

Figure 5. A) location and orientation of *P*-element insertions and transgenes within the *CR43651* locus. Insertions shown in C and D are indicated in bold. B-D) lacZ expression from the *BQ16* reporter in F1 ovaries from B) Oregon-RC (no insertion), C) *w[1118]; P{w[+mW.Scer\FRT.hs]=RS3}CB-5570-3,* and D) *y[1] w[*]; P{w[+mW.hs]=Switch2}GSG3687-1/CyO* females crossed to *BQ16/TM3* males.

Fitness costs associated with *CR43651* insertion chromosomes

In addition to establishing self-regulation through piRNA, TE insertions can also impact fitness, both positively and negatively, by altering the expression of adjacent host genes. The miRNA *miR14* is produced from *CR43651* transcripts. To determine whether *P*-element insertions in *CR43651* reduced *miR14* expression, we examined the abundance of *miR14* in *TP5* as compared to *TP5; insertion chromosome* homozygotes. We observed that for all 3 insertion chromosomes, mature *mir14* (*mir14-RA*) is reduced ~2-fold as compared to the *TP5* background (Figure 6A). Interestingly, the passenger strand miRNA (*mir14-RB*) was less affected in the insertion chromosomes, and did not change systematically from *TP5* abundance levels.

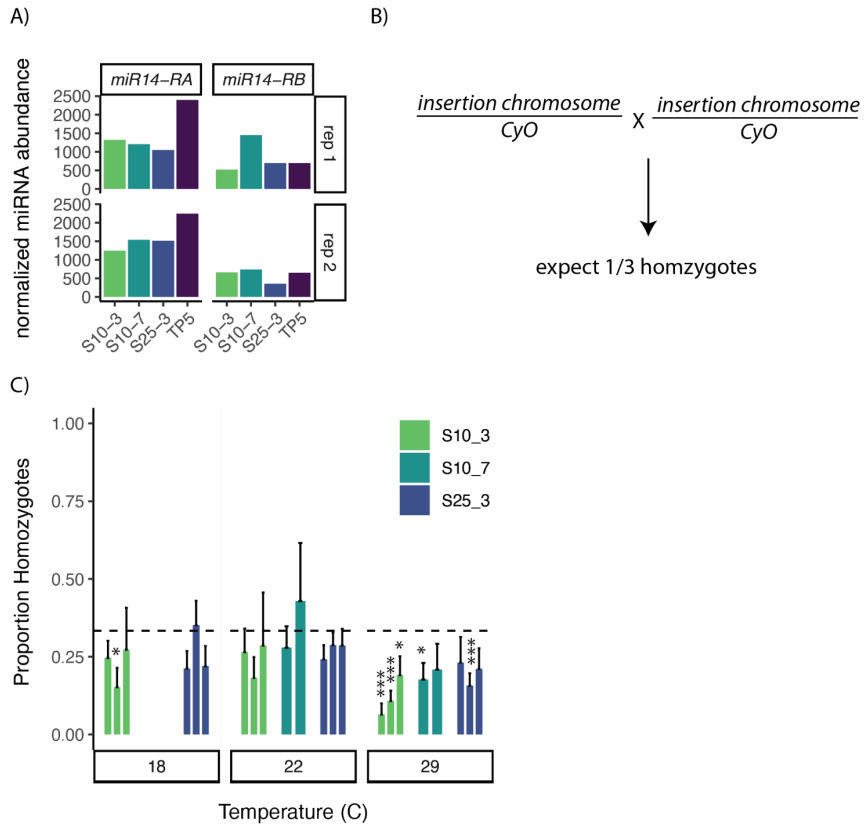

Figure 6. Chromosomes carrying insertions chromosomes show reduced homozygous viability. A) Abundance of mature *mir14* RNAs from insertion chromosomes as compared to the second chromosome in the *TP5* background. Abundance was normalized per million sequenced miRNAs. B) Crossing scheme to detect homozygous viability effects of insertion chromosomes. Assuming no effect on viability 1/3 of offpsring are expected to be homozygotes due to recessive lethality of the *CyO* balancer chromosome. B) Proportion homozygotes observed for each extraction chromosomes at 3 different temperatures. S10_7 could only be replicated at higher temperatures since it was difficult to maintain the balancer chromosome in this stock. All insertion chromosomes are in a balanced *TP5* background (*TP5/FM6;insertion chromosome/CyO*). P-values were obtained from a Chi-Squared Goodness-of-fit test given an expected probaility of 1/3. *,*** denotes $P < 0.05$, 0.001. Data unerlying panel C are provided in S3 Data. small RNA sequencing underlying panel A are in PRJNA1292797.

mir14 has a number of known functions in cell death, metabolism and development, particularly in larval stages [17–19]. *mir14* mutants furthermore generally die at the pupal stage. To determine whether homozygosity for *P*-element insertions into *CR43651* impacted egg to adult viability, we compared the eclosion rates of insertion chromosome homozygotes (*TP5;insertion chromosome*) to their heterozygous siblings (*TP5;insertion chromosome/CyO*) at 3 different temperatures (18°C,22°C,29°C). All insertions have significant homozygous fitness costs at 29°C, with fewer than the expected ⅓ of the eclosing adults corresponding to homozygotes (balancer chromosome homozygotes are not viable, Figure 6B,6C). One replicate of *S10-3* also exhibited fitness costs at 18°C. While we cannot rule out that these effects are impacted or exacerbated by homozygosity for other *P*-element insertions on these same chromosomes, the only insertion site shared between all three chromosomes is *CR43651*. This strongly suggests that homozygosity for *P*-element insertions in *CR43651* has deleterious effects on egg to adult viability, consistent with the reduction in *mir-14* expression we observe.

DISCUSSION

What is unique about the *CR43651* locus, and why are *P*-element insertions in this locus adaptive in laboratory populations? Our data show that chromosomes carrying antisense

insertions in *CR43651*, but not sense insertions, show strong piRNA-like silencing. Furthermore, piRNA-like silencing established by chromosomes containing antisense *CR43651* insertions is notably stronger in the germaria than silencing from the *bona fide* piRNA cluster *TP5 (*Figure 1). Germaria house the S-phase cells where *P*-elements transpose [10], making these structures of profound significance for *P*-element control. Furthermore, piRNA silencing established by insertions in canonical TAS piRNA clusters, like the *TP5* insertion, is diminished in this compartment [12]. We propose that therefore that the adaptive benefit of *CR43651* insertions may relate to their ability to establish silencing in a refractory genomic compartment.

Consistent with this model, we discovered that *CR43651* may have an innate affinity for Rhino protein, which specifies piRNA precursor transcription from dual-stranded piRNA clusters [32,40]. We observed that there is Rhino-binding to *CR43651* only in the absence of Kipeferl, a Rhino-interacting protein that is proposed to stabilize Rhino on chromatin, thereby defining many piRNA clusters (Figure 4)[37]. Kipferl is not expressed in the earliest stages of oogenesis but exhibits increasing expression as germline development progresses through germaria [37]. Therefore, Rhino may occupy *CR43651* at the earliest stages of oogenesis, where *P*-elements transpose.

In addition to early Rhino occupancy, other features of *CR43651* insertion alleles are also suggestive of a dual-stranded piRNA cluster. First, piRNA-like silencing was associated with antisense insertions, where transcription from the *P*-element promoter would be convergent with *CR43651*, thereby producing double-stranded RNAs (dsRNA). This architecture is reminiscent of some dual-stranded piRNA clusters, whose piRNA biogenesis is dependent on the promoters of adjacent protein coding genes [32]. It is also consistent with a recent study suggesting that convergent transcription from TE and host gene promoters can fuel the emergence of new piRNA clusters over successive generations [33].

One puzzle that remains is why chromosomes containing *P*-element derived transgenes inserted into *CR43651* do not establish *P{lacZ}* silencing in the ovary, even when they occur antisense to the gene promoter (Figure 5). These transgenes are regulated by the *P*-promoter, and further express the first two exons of *P*-transposase. Therefore, they have the same potential to produce homologous piRNA to a *P{lacZ}* reporter as native *P*-elements. Indeed, we tested one *P{lacW}* transgene that would have more homologous sequence with *BQ16* than a native *P*-element, since the two transgenes share the *lacZ* coding sequence. At least two scenarios are possible. One is that, although the antisense insertions may be zygotically competent for piRNA silencing, they may still depend on maternally deposited piRNAs from canonical piRNA cluster insertions to activate their piRNA biogenesis. A second scenario is that additional *P*-element copies may be required in the genetic background to feedforward piRNA production from *CR43651*. These models present exciting avenues for future work.

Another persistent puzzle is why *CR43651* insertion alleles are not more common among the wild-derived inbred lines of the *Drosophila* Genetic Reference Panel (DGRP), which contain abundant genomic *P*-elements [15,41,42]. Only one of >200 DGRP lines carry insertions in *CR43641,* despite the potential benefits of these insertions in *P*-element control. We propose that the absence of such insertions in inbred lines reflect the homozygous viability effects of insertion alleles on *mir14* expression and function. Interestingly, Harwich, which famously contains many *P*-elements and is not inbred, contains an insertion in *CR43651* [43]. Surveys of non-inbred wild-derived samples would be an interesting way to address this question.

MATERIALS AND METHODS

Generation of chromosome extraction lines and quantification of viability effects.
To isolate individual second chromosomes containing *P*-element insertions in *CR43651*, we extracted them using *w[1118]; wg[Sp-1]/CyO; sens[Ly-1]/TM6B, Tb[1]* (S2 Table). We then used a *TP5:WgSp-1/CyO* strains to move second chromosomes into a *TP5* background. We then screened for individual chromosome extraction lines containing insertions into *CR43651* using two PCR primers that flank the insertion hotspot region to look for insertion lines with unexpectedly large amplicons [16]. We then characterized whether the insertions were on the plus or minus strand by combining primers in the *CR43651* with primers in *P*-element (S3 Table). Stocks were maintained as *TP5;P{}CR43651/CyO*. To estimate viability effects, flies were intercrossed in temperature controlled incubators, and the number of heterozygous *TP5;P{}CR43651/CyO* and homozygous *TP5;P{}CR43651* offspring were quantified.

***PLacZ* silencing assays.**
BC69 and *BQ16* are *P{lacZ}* reporter genes on the second and third chromosomes that are expressed in the female germline. They contain in-frame fusions between the second exon of *P-transposase* and *Escheria coli lacZ* [20], rendering their LacZ production sensitive to the presence of piRNA [21]. *TP5;P{}CR43651/CyO* females were crossed to *TP5;wg^{Sp-1}/CyO* males to isolate double-balanced females: *TP5/FM6;wg^{Sp-1}/CyO.* These females were then crossed to *BC69/CyO* and *BQ16* males, and F1 females were collected and reared on yeast paste for 2-4 days. For crosses with *w;BQ16/Sb* compared all possible combination of first and second chromosomes, but for *w;BC69/CyO*, we compared only *TP5/w;P{}CR43651/BC69* with *FM6/w;P{}CR43651/BC69,* because we could not differentiate between *P{}CR43651/CyO* and *BC69* genotypes.

To ascertain silencing from *P*-element derived transgenes in *CR43651*, we crossed homozygous and balanced transgenic stocks (S2 Table) to *w;BQ16/Sb* males only and examined LacZ activity in offspring containing both the *CR43651* transgene and *w;BQ16/Sb*.

Dissected ovaries were stained for LacZ activity as in Lemaitre *et al.* [23]. Briefly, ovaries were dissected in 1X PBS and fixed in 0.5% glutaraldehyde for 4-6 minutes. They were then washed in 1X PBS twice, then incubated in 0.2% X-Gal overnight and 37°C. Ovaries were then washed and mounted in glycerol. Stained ovaries were visualized on a Nikon SZX10 under 20X magnification, or on a Nikon Ti2 under 60X magnification. LacZ expression was scored separately for germaria and developing egg chambers. For statistical analysis of LacZ activity, the proportion of LacZ expressing germaria or ovarioles was compared between genotypes, based on ~9 ovarioles from ~16 females for each genotype for BQ16 crosses, and based on ~13 ovarioles from ~12 females for BC69. The proportion of lacZ expressing ovarioles was compared between genotypes using a one-way permutation test implemented in the R-package “coin”.

Small RNA sequencing and analysis. Two biological replicates including ovaries from 20 3-5 day old females maintained with yeast paste were dissected for each of the extraction chromosomes in the *TP5* background (*TP5;P{}CR43651*), as well as in the *TP5* background

without the transgene. Ovaries were homogenized in Trizol and RNA was extracted according to manufacturer instructions, except that RNA was precipitated overnight with an excess of ice-cold ethanol. Depletion of 2S rRNA was completed according to [44], and small RNA library prep and single-end 50 nt sequencing on the Novaseq X plus was completed by Novogene (Carlsbad, CA).

Adapters were removed from small RNA sequencing reads, and they were separated into 2 size classes with trim-galore, 18-22 nt, and 23-32 nt. Read alignments were performed using BWA [45]. We first removed ribosomal RNAs (rRNAs) by aligning and removing reads that aligned to any annotated rRNA, including 2sRNA, from Release 6 of the *D. melanogaster* genome (dm6) [46]. miRNAs were then identified based on their alignment to all annotated miRNAs in dm6 (up to 1 mismatch). 23-32 nt RNAs did not correspond to miRNAs but aligned to the A4 assembly, a repbase TE consensus sequence for *D. melanogaster [47],* or the *P*-element consensus (0 mismatches), were considered piRNAs. For sliding window analyses and *P*-element derived piRNA abundance, raw read counts were normalized as reads per million mapped miRNAs from the same library.

Sliding window analysis of piRNA abundance over *P*-elements or their insertion sites were calculated using bedtools genomecov [48]. Ping-pong abundance was identified using an in-house script from Kelleher *et al* [49].

Genome Sequencing and Identification of *P*-element insertion sites.

Illumina genomic libraries were generated with the Nextera CD Index kit (Illumina, San Diego, CA) according to manufacturer instructions. 150 nt Paired-End Sequencing was performed on a Next-Seq 500 by the University of Houston Seq ’n Edit core. Adapters were removed using trim-galore. *P*-element insertions in A4 were identified through the PIDFE pipeline [15,16,50], and are reported in Supplementary table 1.

ACKNOWLEDGEMENTS

We are grateful to Stéphane Ronsseray and Antoine Boivin for providing the *P{lacZ}* reporter strains *BQ16* and *BC69.* This work was supported by R35 GM138112 to ESK. NMB and MW were supported by NSF 2147073 to Richard Meisel.

REFERENCES

1. Mével-Ninio M, Pelisson A, Kinder J, Campos AR, Bucheton A. The flamenco locus controls the gypsy and ZAM retroviruses and is required for Drosophila oogenesis. Genetics. 2007;175: 1615–1624.

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

3. Le Thomas A, Tóth KF, Aravin AA. To be or not to be a piRNA: genomic origin and processing of piRNAs. Genome Biol. 2014;15: 204.

4. Senti K-A, Jurczak D, Sachidanandam R, Brennecke J. piRNA-guided slicing of transposon transcripts enforces their transcriptional silencing via specifying the nuclear piRNA repertoire. Genes Dev. 2015;29: 1747–1762.

5. Yu Y, Gu J, Jin Y, Luo Y, Preall JB, Ma J, et al. Panoramix enforces piRNA-dependent cotranscriptional silencing. Science. 2015;350: 339–342.

6. Senti K-A, Brennecke J. The piRNA pathway: a fly's perspective on the guardian of the genome. Trends Genet. 2010;26: 499–509.

7. Hamilton RS, Hartswood E, Vendra G, Jones C, Van De Bor V, Finnegan D, et al. A bioinformatics search pipeline, RNA2DSearch, identifies RNA localization elements in Drosophila retrotransposons. RNA. 2009;15: 200–207.

8. Wang L, Dou K, Moon S, Tan FJ, Zhang ZZ. Hijacking Oogenesis Enables Massive Propagation of LINE and Retroviral Transposons. Cell. 2018;174: 1082–1094.e12.

9. Tiwari B, Kurtz P, Jones AE, Wylie A, Amatruda JF, Boggupalli DP, et al. Retrotransposons Mimic Germ Plasm Determinants to Promote Transgenerational Inheritance. Curr Biol. 2017;27: 3010–3016.e3.

10. Spradling AC, Bellen HJ, Hoskins RA. Drosophila P elements preferentially transpose to replication origins. Proc Natl Acad Sci U S A. 2011;108: 15948–15953.

11. Story B, Ma X, Ishihara K, Li H, Hall K, Peak A, et al. Defining the expression of piRNA and transposable elements in Drosophila ovarian germline stem cells and somatic support cells. Life Sci Alliance. 2019;2. doi:10.26508/lsa.201800211

12. Dufourt J, Dennis C, Boivin A, Gueguen N, Théron E, Goriaux C, et al. Spatio-temporal requirements for transposable element piRNA-mediated silencing during Drosophila oogenesis. Nucleic Acids Res. 2014;42: 2512–2524.

13. Marin L, Lehmann M, Nouaud D, Izaabel H, Anxolabéhère D, Ronsseray S. P-Element repression in Drosophila melanogaster by a naturally occurring defective telomeric P copy. Genetics. 2000;155: 1841–1854.

14. Stuart JR, Haley KJ, Swedzinski D, Lockner S, Kocian PE, Merriman PJ, et al. Telomeric P elements associated with cytotype regulation of the P transposon family in Drosophila melanogaster. Genetics. 2002;162: 1641–1654.

15. Zhang S, Pointer B, Kelleher ES. Rapid evolution of piRNA-mediated silencing of an invading transposable element was driven by abundant de novo mutations. Genome Res. 2020;30: 566–575.

16. Wang L, Zhang S, Hadjipanteli S, Saiz L, Nguyen L, Silva E, et al. P-element invasion fuels molecular adaptation in laboratory populations of Drosophila melanogaster. Evolution. 2023;77: 980–994.

17. Xu P, Vernooy SY, Guo M, Hay BA. The Drosophila microRNA Mir-14 suppresses cell death and is required for normal fat metabolism. Curr Biol. 2003;13: 790–795.

18. Varghese J, Cohen SM. microRNA miR-14 acts to modulate a positive autoregulatory loop controlling steroid hormone signaling in Drosophila. Genes Dev. 2007;21: 2277–2282.

19. Nelson C, Ambros V, Baehrecke EH. miR-14 regulates autophagy during developmental cell death by targeting ip3-kinase 2. Mol Cell. 2014;56: 376–388.

20. O'Kane CJ, Gehring WJ. Detection in situ of genomic regulatory elements in Drosophila. Proc Natl Acad Sci U S A. 1987;84: 9123–9127.

21. Josse T, Teysset L, Todeschini A-L, Sidor CM, Anxolabéhère D, Ronsseray S. Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin formation. PLoS Genet. 2007;3: 1633–1643.

22. Josse T, Maurel-Zaffran C, de Vanssay A, Teysset L, Todeschini A-L, Delmarre V, et al. Telomeric trans-silencing in Drosophila melanogaster: tissue specificity, development and functional interactions between non-homologous telomeres. PLoS One. 2008;3: e3249.

23. Lemaitre B, Ronsseray S, Coen D. Maternal repression of the P element promoter in the germline of Drosophila melanogaster: a model for the P cytotype. Genetics. 1993;135: 149–160.

24. Ronsseray S, Marin L, Lehmann M, Anxolabéhère D. Repression of hybrid dysgenesis in Drosophila melanogaster by combinations of telomeric P-element reporters and naturally occurring P elements. Genetics. 1998;149: 1857–1866.

25. Simmons MJ, Ryzek D-F, Lamour C, Goodman JW, Kummer NE, Merriman PJ. Cytotype regulation by telomeric P elements in Drosophila melanogaster: evidence for involvement of an RNA interference gene. Genetics. 2007;176: 1945–1955.

26. Belinco C, Diprima SN, Wolff RE, Thorp MW, Buschette JT, Simmons MJ. Cytotype regulation in Drosophila melanogaster: synergism between telomeric and non-telomeric P elements. Genet Res . 2009;91: 383–394.

27. Simmons MJ, Ragatz LM, Sinclair IR, Thorp MW, Buschette JT, Grimes CD. Maternal enhancement of cytotype regulation in Drosophila melanogaster by genetic interactions between telomeric P elements and non-telomeric transgenic P elements. Genet Res . 2012;94: 339–351.

28. Simmons MJ, Meeks MW, Jessen E, Becker JR, Buschette JT, Thorp MW. Genetic interactions between P elements involved in piRNA-mediated repression of hybrid dysgenesis in Drosophila melanogaster. G3 . 2014;4: 1417–1427.

29. 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.

30. Kelleher ES, Barbash DA. Analysis of piRNA-mediated silencing of active TEs in Drosophila melanogaster suggests limits on the evolution of host genome defense. Mol Biol Evol. 2013;30. doi:10.1093/molbev/mst081

31. 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.

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

33. Luo Y, He P, Kanrar N, Fejes Toth K, Aravin AA. Maternally inherited siRNAs initiate piRNA cluster formation. Mol Cell. 2023;83: 3835–3851.e7.

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.

35. Hayashi R, Schnabl J, Handler D, Mohn F, Ameres SL, Brennecke J. Genetic and mechanistic diversity of piRNA 3'-end formation. Nature. 2016;539: 588–592.

36. 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.

37. 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. doi:10.7554/eLife.80067

38. Le Thomas A, Stuwe E, Li S, Du J, Marinov G, Rozhkov N, et al. Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. Genes Dev. 2014;28: 1667–1680.

39. Ronsseray S, Boivin A, Anxolabéhère D. P-element repression in Drosophila melanogaster by variegating clusters of P-lacZ-white transgenes. Genetics. 2001;159: 1631–1642.

40. Klattenhoff C, Xi H, Li C, 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.

41. Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, et al. The Drosophila melanogaster Genetic Reference Panel. Nature. 2012;482: 173–178.

42. Srivastav SP, Kelleher ES. Paternal Induction of Hybrid Dysgenesis in *Drosophila melanogaster* Is Weakly Correlated with Both *P* -Element and *hobo* Element Dosage. G3: Genes|Genomes|Genetics. 2017; g3.117.040634.

43. Khurana JS, Wang J, Xu J, Koppetsch BS, Thomson TC, Nowosielska A, et al. Adaptation to P element transposon invasion in Drosophila melanogaster. Cell. 2011;147: 1551–1563.

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

45. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25: 1754–1760.

46. Hoskins RA, Carlson JW, Wan KH, Park S, Mendez I, Galle SE, et al. The Release 6 reference sequence of the Drosophila melanogaster genome. Genome Res. 2015;25: 445–458.

47. Bao W, Kojima KK, Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob DNA. 2015;6: 11.

48. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26: 841–842.

49. Kelleher ES, Edelman NB, Barbash DA. Drosophila Interspecific Hybrids Phenocopy piRNA-Pathway Mutants. PLoS Biol. 2012;10. doi:10.1371/journal.pbio.1001428

50. Zhang S, Kelleher ES. Targeted identification of TE insertions in a Drosophila genome through hemi-specific PCR. Mob DNA. 2017;8: 10.