HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2025 May 21.

Published in final edited form as:

Cell Rep. 2025 April 22; 44(4): 115354. doi:10.1016/j.celrep.2025.115354.

Traffic Jam activates the *Flamenco* piRNA cluster locus and the Piwi pathway to ensure transposon silencing and *Drosophila* fertility

Austin J. Rivera[1], **Jou-Hsuan Roxie Lee**[1], **Shruti Gupta**[1], **Linda Yang**[1], **Raghuveera Kumar Goel**[1,2,4], **Joseph Zaia**[1,2], **Nelson C. Lau**[1,3,5,*]

[1]Department of Biochemistry and Cell Biology, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, USA

[2]Proteomics Service Center and Center for Network Systems Biology, Boston University, Boston, MA, USA

[3]Genome Science Institute, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, USA

[4]Present address: Division of Oncological Sciences, Knight Cancer Institute, Oregon Health and Science University (OHSU), Portland, OR, USA

[5]Lead contact

SUMMARY

Flamenco (*Flam*) is a prominent Piwi-interacting RNA (piRNA) locus expressed in *Drosophila* ovarian follicle cells that silences *gypsy/mdg4* transposons to ensure female fertility. Promoter-bashing reporter assays in ovarian somatic sheet (OSS) cells uncover compact enhancer sequences within *Flam*. We confirm the enhancer sequence relevance *in vivo* with *Drosophila Flam* deletion mutants that compromise *Flam* piRNA levels and female fertility. Proteomic analysis of proteins associated with *Flam* enhancer sequences discover the transcription factor Traffic Jam (TJ). *Tj* knockdown in OSS cells causes a decrease in *Flam* transcripts, *Flam* piRNAs, and multiple Piwi pathway genes. TJ chromatin immunoprecipitation sequencing (ChIP-seq) analysis confirms TJ binding at enhancer sequences deleted in our distinct *Flam* mutants. TJ also binds multiple Piwi pathway gene enhancers and long terminal repeats of transposons that decrease in expression after *Tj* knockdown. TJ plays an integral role in the ongoing arms race between selfish transposons and their suppression by the host Piwi pathway and *Flam* piRNA locus.

*Correspondence: nclau@bu.edu.

AUTHOR CONTRIBUTIONS

Conceptualization, N.C.L. and A.J.R.; methodology/investigation, N.C.L., A.J.R., S.G., J.-H.R.L., L.Y., and R.K.G.; formal analysis, N.C.L., A.J.R., S.G., J.-H.R.L., and R.K.G.; data curation/software, S.G. and J.-H.R.L.; writing – original draft, N.C.L.; writing – review & editing, N.C.L., S.G., A.J.R., and R.K.G.; funding acquisition, N.C.L. and J.Z.

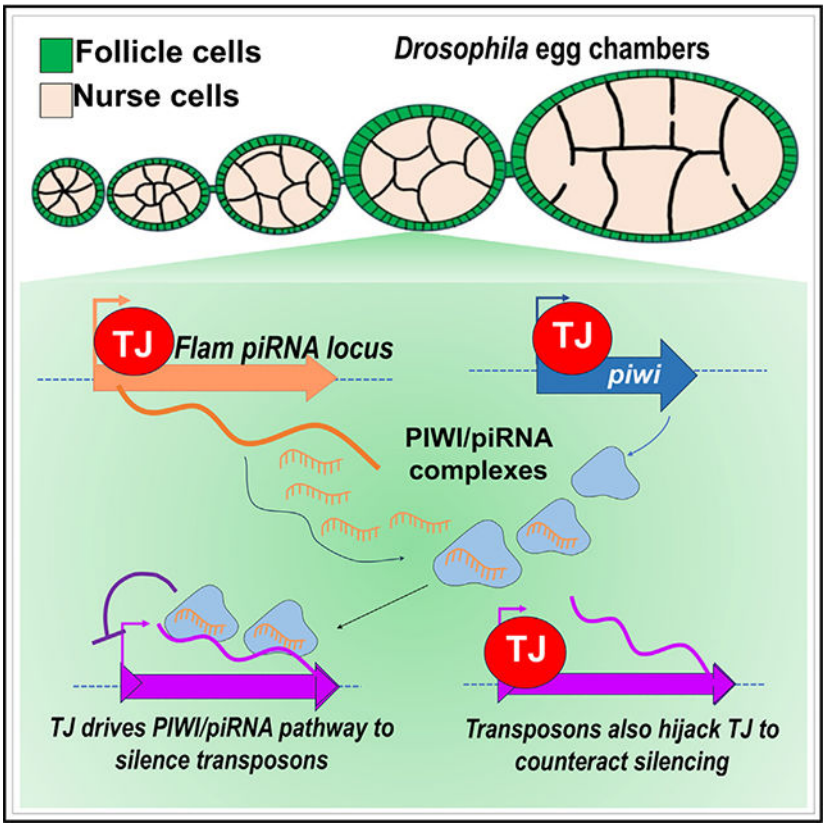
DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2025.115354.

Graphical abstract



In brief

Rivera et al. discover *Drosophila Flamenco* piRNA cluster enhancers bound and activated by the Traffic Jam transcription factor. Traffic Jam also binds many follicle cell genes including activating multiple Piwi pathway genes; and Traffic Jam binds and activates the LTRs of some retrotransposons.

INTRODUCTION

In the *Drosophila* ovary, the germline compartment of nurse cells and the oocyte are encompassed by the somatic compartment of follicle cells. The germline generates Piwi-interacting RNAs (piRNAs) from piRNA cluster loci that express transcripts from both genomic strands and are transcriptionally activated by the factors Moonshiner, TRF2, and the Rhino-Deadlock-Cutoff complex.[1,2] In contrast, follicle cells only express a distinct set of unistrand piRNA cluster loci, with the most prominent one named *Flamenco* (*Flam*) for the initial genetic discovery of mutants in this locus that enabled the mobilization of the *gypsy/mdg4* transposable elements (TEs).[3–5]

Although *Drosophila* mutants deleted of individual germline dual-strand piRNA cluster loci are nearly as fertile as wild-type (WT) flies with little TE activation,[6] *Flam* mutant females are severely subfertile because multiple retroviral TEs are activated and generate infectious

viral particles that are linked to massive ovarian collapse.[7–9] Two classic *Flam* mutations, KG00476 and BG02658, are referred to as $Flam^{KG}$ and $Flam^{BG}$ in this study and are both caused by *P element* insertions in the *Flam* promoter/enhancer regions that result in severe loss of *Flam* piRNAs.[5,7,10] A previous study proposed that the $Flam^{BG}$ mutation disrupted a binding site for the transcription factor *Cubitus interruptus* (*Ci*).[11] Because *Ci* is also broadly expressed in many other *Drosophila* cell types that lack *Flam* and piRNA expression,[12] we wondered if another factor specific to follicle cells could be the primary activator of *Flam*.

Ovarian somatic sheet (OSS) and the ovarian somatic cell (OSC) lines are derived from the follicle cell lineage and express the somatic PIWI pathway, including abundant genic piRNAs and *Flam* piRNAs.[13–15] These OSS/OSC lines have been important models for illuminating the somatic compartment of PIWI/piRNA biochemical silencing mechanisms and piRNA biogenesis factor roles,[16–19] but germline piRNA cluster loci and germ-line *Aub* and *Ago3* homologs of *piwi* are not expressed in these cells. Only the loss of the *l(3)mbt* and *lint-O* repressors in OSCs transformed these cells into expressing *Aub*, *Ago3*, and *vasa*, and yet the piRNAs in these modified OSCs still mainly resembled the original bulk of *Flam* piRNAs.[20,21]

Although the high-throughput self-transcribing active regulatory region sequencing (STARR-seq) method could define OSC-specific enhancer elements within the *Flam* locus,[22] neither the STARR-seq approach nor bioinformatics predictions[11,23] have determined which specific enhancer elements and transcription factor(s) fully drive *Flam* expression. Therefore, we approached this knowledge gap by re-designing a promoter-bashing reporter assay for *Flam* to reveal enhancer sequences and then utilized these enhancer sequences to discover the specific DNA-binding protein with proteomics. Our study discovers that the large musculoaponeurotic fibrosarcoma (MAF) transcription factor Traffic Jam (TJ) binds to *Flam* enhancers and directly regulates *Flam* expression. Additionally, we show that *Tj* has a broad influence on the follicle cell transcriptome by also serving as a nexus within the arms-race conflict between the *Drosophila* host genome defense pathways and the selfish TEs.

RESULTS

A revised promoter-bashing reporter assay pinpoints distinct *Flam* enhancer elements

A previously described *Flam* promoter luciferase reporter assay[11] first started with a construct that was 515 bp upstream and 101 bp downstream of the *Flam* transcription start site (TSS). This 616 bp *Flam* promoter element reporter tested deleting the TSS and a putative binding sequence for the *Ci* transcription factor. We also tested a very short fragment from *Flam* we designated as a negative control (segment i14) by first transfecting solely *Drosophila* OSS cells like in Goriaux et al.[11] At first, this original Goriaux et al. segment that we named “i4” appeared to promote gene activation compared to our negative control (Figure 1A).

However, to normalize out general transcription activation and focus on OSS cell-specific activation of the *Flam* promoter reporters, we included additional reporter transfections in S2 cells that completely lacked expression of the Piwi pathway and *Flam*.[13,24] We

then analyzed reporter expression levels by first calculating the ratio of OSS cellreporter readingsover S2cell readings and then normalizing these ratios of readings to the segment i14 negative control element. This analysis showed only a modest (≤2-fold) basal transcription activation difference between the Goriaux et al. element i4 and other similarly sized putative *Flam* enhancer constructs (Figure 1B). With this revised promoter-bashing reporter assay format, we then determined an "i11" element (~1.5 kb upstream and ~1.8 downstream of the *Flam* TSS) that displayed a >1,200-fold activation in OSS cells versus S2 cells.

Since this i11 sequence included original intronic elements proximal to the *Flam* TSS, we controlled for RNA splicing impact by adding a minimal synthetic intron just upstream of luciferase, and this completely mitigated splicing artifacts on reporters lacking this synthetic intron (Figure S1A). This revised promoter-bashing reporter assay reconfirmed the importance of the *Flam* TSS in the *Flam* promoter luciferase reporter but could not recapitulate an impact for the *Ci* binding site nor a role for *Ci* in *Flam* activation (Figure S1B). Our small interfering RNA (siRNA) knockdowns of *Ci* in OSS cells also did not affect *Flam* RNA levels nor raise TE expression (Figure S1C).

From these data, our goal was to deduce which discrete sequence elements activate the *Flam* promoter reporters most specifically in OSS cells (Figure 1C), to go beyond general transcription activation signals like a *Flam* TSS. The 80 bp around the *Flam* TSS that we named the *Flam* initiation transcription site (*FITS*) did behave as a positive control for the *Flam* reporter, but we did not see more direct OSS cell-specific activation by other TSS-adjacent sequences like i3-to-i8 elements upstream of the TSS or i22, i25, and i28 elements with sequences immediately downstream of the *Flam* TSS. We deduced instead two other impactful elements 426 and 743 bp downstream of the *Flam* TSS. We named these two ~139 and 80 bp sequences *Flam* promoter shadow enhancers #1 and~#2, respectively (*FPSE1* and *FPSE2*). The *FPSE1* element is a uniquely mapped sequence in the *Drosophila* genome, while the *FPSE2* element encompasses part of a DNAREP1 repeat and a putative start of an intron.

Distinct deletion mutants of *Flam* enhancer elements exhibit compromised fertility

To test the *in vivo* relevance of *FITS*, *FPSE1*, and *FPSE2*, we searchedfor CRISPR-Cas9 small guide RNA (sgRNA) targeting sequences that closely flanked these loci to generate deletion mutants in *Drosophila melanogaster*. We screened and obtained some precise deletion mutants arising from DNA repair directed by single-stranded oligonucleotides (ssODNs) after sgRNAs were injected into the Cas9-expressing fly line (Figure 2A). Possibly due to the DNAREP1 repeat in the *FPSE2* sequence, the two *FPSE2a* and *FPSE2b* mutations were less precisely mutated because additional sequences from the ssODNs were then duplicated into the sgRNA-guided deletion (Figure S2). We also generated two additional mutant alleles, *FPUN* and *FPRU*, to serve as controls for testing whether the putative *Ci* binding site upstream of $Flam^{11}$ and the $Flam^{BG}$ insertion site[5,7,10] were critical for *Flam* function. All mutants were confirmed by sequencing and subjected to 2–4 rounds of backcrossing to *w1118* before two final crosses to balance over the *FM7a* balancer chromosome.

Adult female homozygous mutants were compared to their heterozygous siblings in egg count assays and fecundity assays (Figure 2B). Although not completely sterile, all of the *FITS*, *FPSE1*, and *FPSE2a/b* homozygous mutants were subfertile because of the drastic reduction of mature eggs that lead to much lower fecundity. However, we did not observe any fertility impact between heterozygous and homozygous *FPUN* and *FPRU* mutants. This result suggests that the putative *Ci* binding site upstream of *Flam* may not actually impact *Flam* function,[11] and the *P element* inserted in this region in the $Flam^{BG}$ mutant[5,7,10] may have a more complicated perturbation of *Flam* regulation than simply disrupting a transcription factor binding site.

We decided to only examine further the mutants with fertility defects by extracting ovarian total RNA to measure by RT-qPCR the *Flam* RNA transcripts and two TE transcripts, *gypsy1* and *mdg1*, known to be elevated in $Flam^{KG}$ mutants.[10] *Flam* RNA was strongly reduced in homozygous *FITS* and *FPSE1* mutants at the same severity as $Flam^{BG}$ and $Flam^{KG}$ mutants, but *Flam* RNA was much less affected in *FPSE2a/b* mutants (Figure 2C). The strong reactivation of both *gypsy1* and *mdg1* was only observed in $Flam^{BG}$ and *FPSE1* homozygous mutants (Figure 2D). Although the *FPSE1* region is downstream of the *P element* insertion of $Flam^{BG}$ (Figure 2A), these data suggest the greatest phenotypic similarity between these two mutants.

The *FPSE1* mutation phenocopies the dramatic loss of *Flam* transcripts and piRNAs seen in $Flam^{BG}$ and $Flam^{KG}$ mutants

The *FPSE1* and *FPSE2a/b* mutant subfertile phenotypes validated the promoter-bashing reporter assay results and encouraged us to further characterize the long RNA and piRNA expression from the ovaries of these distinct *Flam* mutants. To expand upon the limited resolution of our RT-qPCR results, we then deeply sequenced matched long total RNA and small RNA libraries from the ovaries of homozygous *Flam* mutants, comparing relative reads per million (RPM) to ovaries from heterozygous siblings. Surprisingly, despite the deletion of the WT *Flam* TSS, the *FITS* ovaries displayed a shift of the distinct TSS to a site ~350 bp upstream of the original *Flam* TSS (Figure 3A). The *FITS* deletion impacted the RT-qPCR assay with primers directed at *Flam*'s exon 2, but much of the rest of the *Flam* long RNA transcript many kilobases downstream was still being efficiently expressed in *FITS* homozygous mutants (Figure 3C). As such, *Flam* piRNAs and TE silencing and TE small RNAs were also mainly unaffected by the *FITS* mutation (Figures 3B–3D and S3).

Although subfertile phenotypes in all these distinct mutant *Flam* females were reproducible (Figure 2B), there were also no significant differences in long *Flam* transcript levels nor TE expression differences between homozygous *FPSE2a* and *FPSE2b* mutants and their heterozygous siblings (Figures 3A, 3B, S3A, and S3B). We do not have an explanation for why fertility phenotypes persist despite insufficient effects on *Flam* piRNAs in the *FPSE2* and *FITS* mutants, but these results indicate plasticity in the regulation of the *Flam* piRNA cluster locus and suggest that the transcriptional activator binds elsewhere from the *FITS* and *FPSE2* regions.

The massive TE reactivation due to the loss of *Flam* long transcripts and piRNA expression first seen in $Flam^{BG}$ and $Flam^{KG}$ was recapitulated in the *FPSE1* homozygous mutants, such

as the activation of *ZAM*, *blood*, and *gypsy* TEs (Figure S3A). Although *Flam* piRNAs are greatly diminished in *FPSE1* mutants as well as in $Flam^{BG}$ and $Flam^{KB}$ mutants, the small RNA libraries from ovaries also showed elevated TE piRNAs that could have come from the germline compartment that now dominated the compromised somatic compartment in these mutants (Figure 3C).

Proteomics discovery of TJ as a prime candidate binding *Flam* enhancer elements

A goal of the promoter-bashing reporter assay was to define small DNA sequences of *Flam* enhancers to then perform an un-biased biochemical discovery of the transcription factor binding to these elements. We generated 280, 647, and 501 bp biotin-tagged PCR-amplified double-stranded DNAs (dsDNAs) representing *FITS*, *FPSE1*, and *FPSE2*, respectively, and then incubated these DNAs in OSS cell and S2 cell nuclear extracts before pulling the DNAs down with streptavidin-coupled Dynabeads. Pulled-down DNA-binding proteins were then visualized with silver-stained gels and subjected to on-bead trypsin digestion for Tandem Mass Tag (TMT) labeling in quantitative proteomic analysis (Figure S4A).

Three analyses of the proteomic profiles between the different DNA elements congruently pointed to the TJ transcription factor as the top enriched protein in *FPSE1* and *FPSE2* pull-downs (Figures S4B and S4C). The *Tj* gene is known to be a critical gene in *Drosophila* follicle cell development and ovary function, and *Tj*-null mutants are completely sterile from totally collapsed ovaries and testes.[25–27] The *Tj* mRNA is a precursor for abundant genic 3′ UTR piRNAs[13,28] while also being implicated in binding to the *piwi* gene promoter.[13] However, the full extent of TJ's regulation of the follicle cell transcriptome was still unknown.

We first tried to validate TJ's potential impact on *Flam* expression *in vivo* with follicle-cell-specific RNAi knockdown in the fly ovary, starting with the *Tj-GAL-4* driver coupled with a $GAL\text{-}80^{TS}$ repressor for temperature-controlled triggering of *Tj* knockdown. When *Tj* knockdown was triggered efficiently, ovaries collapsed to rudimentary stubs lacking tissue for analysis, whereas timed *Tj* knockdowns were insufficient when ovaries were allowed to develop enough follicle cells for analysis (Figure S5A). Additional later-stage follicle cell drivers like *CG13083-GAL4(FC0)* did not cause a sufficient decrease of *Tj* despite allowing for ovary development (Figure S5B). Given the very abundant expression and critical importance of *Tj* in follicle cell and ovary development, *in vivo* RNAi knockdown of *Tj* to study *Flam* expression remained a challenge.

To aid our validation efforts, we raised a distinct rabbit polyclonal antibody to the N-terminal peptide of TJ, and we determined the antibody's specificity to be equivalent to an original guinea pig-derived anti-TJ antibody[25] (Figure S5C). We also designed specific and efficient siRNAs to knock down *Tj* in OSS cells, which can survive several days and reveal TE activation upon *piwi* knockdown.[13,17–19,29,30] Our rabbit anti-TJ antibody recognized and immunoprecipitated TJ specifically and confirmed effective TJ protein loss during siTj siRNA knockdown in OSS cells (Figures S5C and S5D) but also confirmed that we were unable to transiently knock down TJ enough to affect *Flam* RNA while still allowing for ovary development (Figure S5E). Nevertheless, this anti-TJ antibody could interrogate mosaic clones with genetic loss of TJ expression in an accompanying study.[31]

Tj knockdown in OSS cells validates direct regulation of *Flam* expression

Taking advantage of the OSS cell system for siTj knockdowns, we conducted multiple validation approaches of *Tj*'s potential regulation of *Flam*. We first retested two minimal *Flam* promoter luciferase reporter constructs that either contained or lacked the *FPSE1* sequence in OSS cells after siTj knockdown (Figure 4A). Compared to siGFP and siPiwi negative controls, only the i29 promoter reporter containing the *FPSE1* sequence was less activated upon siTj knockdown (Figure 4B). We then applied the same *Flam* and TE RT-qPCR assays we had applied to the *FPSE1* mutant fly ovaries (Figures 2C and 2D) to the OSS cells after siTj knockdown. Indeed, *Flam* transcripts were reduced after siTj knockdown, followed by an increase of just the *gypsy1* TE but not the two other TEs that are targeted for silencing by PIWI and piRNAs in OSS cells (Figure 4C).

We then sequenced long RNAs and small RNAs from OSS cells transfected 48 h prior with siTj, siCi, and siPiwi siRNAs to knock down target genes and compared them to the siGFP control. The only piRNA effect of the siCi knockdown was a decrease in genic *Tj* piRNAs (Figure 4F). The *Flam* piRNAs were as severely depleted in siTj as in siPiwi, while the *Flam* long RNA was only reduced by siTj (Figures 4D, 4E, and S6A). The siTj knockdown also reduced piRNAs from the *20A* piRNA cluster locus without a significant decrease of *20A* long RNAs (Figure 4E). Many other TE-targeting and genic piRNAs were depleted as severely in the siTj knockdown as in the siPiwi knockdown (Figure 4F). However, in contrast to the expected observation of many elevated TE long RNAs observed in the siPiwi knockdown, a notable number of TE long RNAs had decreased in expression in the siTj knockdown (Figure 4G).

Turning to the protein-coding genes, many more genes were upregulated after siTj and siPiwi knockdown compared to the siGFP control (Figure S6C). We inspected more closely genes downregulated after siTj as potential direct targets of transcriptional activation by *Tj*, and we observed a ~75% decrease in *piwi* mRNA levels after siTj knockdown (Figure S6D), which agrees with the previously described binding of TJ to the *piwi* promoter.[13] Interestingly, several other piRNA biogenesis factor genes were also strongly downregulated by siTj knockdown, such as *armi*, *mael*, and *fs(1)Yb* (Figure S6E). We confirmed that the proteins from these genes, including PIWI, were decreased in OSS cells when two independent siRNAs knocked down *Tj* (Figure S5D). As siTj knockdown also reduces the expression of multiple Piwi pathway genes, this explains why all the OSS cell piRNAs are as severely depleted in siTj as in siPiwi knockdowns (Figure 4F).

TJ ChIP-seq reveals direct binding to *Flam FPSE1* region and Piwi pathway gene promoters

To determine how directly TJ binds to target gene promoters to influence transcription, we conducted TJ chromatin immunoprecipitation sequencing (ChIP-seq) in OSS cells as well as profiled the gene-activating H3K4me3 and gene-silencing H3K9me3 chromatin marks using Cut&Run in normal OSS cells and after siTj knockdown. We also integrated into our analysis a modEN-CODE ChIP-seq dataset of transgenically expressed GFP-Tj from whole flies,[32] which has the caveat of the anti-GFP antibody accessing the GFP-TJ fusion from tissues other than the ovary. Nevertheless, we could find TJ ChIP-seq peaks confidently

from both datasets with the irreproducibility discovery rate (IDR) approach, and then we annotated the peaks to genes, finding 1,053 genes with TJ peaks overlapping both datasets (Figure 5A; Tables S1 and S2).

We expected the TJ ChIP-seq data from OSS cells to more specifically reflect patterns in the ovarian follicle cells, but the GFP-TJ ChIP-seq from whole flies still provided excellent independent validation of shared binding patterns as a proxy of TJ in the fly. For instance, when both TJ peaks from OSS cells and GFP-TJ peaks from whole flies were analyzed with the MEME program,[33,34] both datasets yielded the same top *de novo* sequence motif "GCTGAC" as a core motif with a downstream "GC-tail" (Figure 5B). The newer program STREME[33] also confirmed this *de novo* sequence motif along with an upstream AT-rich sequence seen in half-MARE motifs for MAF transcription factors[35] (Figure S7A). The vast majority (>89%–90%) of the genes with either TJ or GFP-TJ IDR-called peaks all contained this TJ motif (Figure 5A, right Venn diagram). Remarkably, this TJ motif we determined in *Drosophila* cells matched exactly the TJ motif first predicted by the Fly Factor Survey that employed a bacterial 1-hybrid approach[36] (Figure S7B).

A metagene and functional annotation analysis of the distribution of TJ binding motifs shows a notable preference for TJ binding at gene TSSs and downstream of the TSS in mostly enhancer elements versus a lower proportion in upstream promoter elements (Figures S7C and S7D). However, the highly compact juxtaposition of *Drosophila* genes also biases this analysis because we could not extend the upstream region of one gene longer than ~400 bp before hitting the next gene. The majority (>70%) of the 1,505 genes in OSS cells have between 1 and 3 TJ binding motifs matching the TJ binding peaks (Figure S7E).

In *D. melanogaster*, eight protein-coding genes are configured within a ~27 kb proximity to the *Flam* TSS, and we could define seven reproducible H3K4me3 peaks demarcating a proxy for RNA polymerase II (RNA Pol II) residence at these genes' TSSs (Figure 5C). Six of these H3K4me3 peaks were all lost in the siTj knockdown in OSS cells, yet there were only two definitive IDR-called TJ peaks in this region: one peak flanking the end of the upstream protein-coding genes and the other TJ peak directly overlapping the *FPSE1* deleted region in the fly mutant (Figure 5C, bottom). In fact, the *FPSE1* deleted region contains the TJ sequence motif and overlaps with the main H3K4me3 peak associated with this *Flam* enhancer that lies ~450 bp downstream of the *Flam* TSS. The GFP-TJ IDR-called peak at this *FPSE1* region also matches the TJ peak from OSS cells.

The TJ binding peaks in GFP-TJ in whole flies and TJ in OSS cells were concordant across the Piwi pathway gene promoters that decreased in expression after siTj knockdown (Figure 5D). The TJ peak that our data define at the *piwi* promoter has much better resolution and is in perfect agreement with the previous ChIP-qPCR result from Saito et al.[13] Like *Flam*, the Piwi pathway genes' TJ motif was downstream of the genes' TSSs, and the most proximal H3K4me3 peak overlapping this TJ peak was lost after siTj knockdown. Lastly, these H3K4me3 peak changes at TJ-regulated genes were exquisitely tight, such as the examples of the unaffected H3K4me3 peaks remaining at *CG2701* and *ND-23* after siTj knockdown, despite their gene starts being less than ~500 bp from the starts of *fs(1)Yb* and *spn-E*, respectively.

TJ has been hijacked to activate TE expression in *Drosophila* ovary cells

Although one of TJ's original functions in the ovary is to promote transcriptomes required for follicle cell development,[25] our data also provide strong evidence for TJ in activating the Piwi/piRNA pathway that serves as the host genome defense against TE expression. Normal fly ovary cells are usually safeguarded by a functional Piwi/piRNA pathway, yet multiple retroviral TEs are poised to reanimate and generate infective viral-like particles at the moment the Piwi/piRNA pathway is genetically compromised.[37] This Senti et al. study posits the theme of an ongoing arms race and co-evolution between the selfish TEs and the somatic piRNA pathway that is exemplified by the enrichment of TE sequences in the *Flam* piRNA cluster locus to enable the most effective silencing of the most threatening and active TEs in *Drosophila* follicle cells.[37]

We considered this paradigm to explain why several TEs' long RNAs were diminished in OSS cells after siTj knockdown (Figures 6A and 4D). Since the ChIP-seq IDR peak-calling method cannot be applied to TE sequences, we mapped the ChIP-seq and Cut&Run reads to TE consensus sequences and looked for the highest intuitive peaks and differences in H3K9me3 and H3K4me3 levels. We observed pointed TJ and GFP-TJ peaks enriched at both of the long terminal repeats (LTRs) of many retroviral-like TEs, like *gypsy*, *gypsy5*, *HMS-Beagle2*, and *gtwin* (Figure 6B). Multiple other Metaviridae TEs, as well as some LINE and DNA TEs, exhibited these TJ peaks at LTRs and ends that would correspond to the transcriptional promoter regions for the TEs (Figure S8). Bolstering this result, TJ binding motifs were detected at the TJ ChIP-seq peaks at these TE ends.

After siTj knockdown, we observed clear increases in H3K9me3 levels at multiple TEs across much of the TE consensus sequences (Figures 6 and S8). Since these TEs are targeted by piRNAs, this increase in H3K9me3 after siTj knockdown may be directed by the longer residency of Piwi/piRNA complexes at TE nascent transcripts when there is not as much TJ to promote RNA Pol II transcription that would be counteracting the H3K9me3 levels. Similar to how PAF1 knockdown enhances Piwi/piRNA silencing of TE transcripts,[38] TJ knockdown may operate through a similar mechanism, even though the TEs' H3K4me3 profiles did not exhibit prototypical patterns seen at the other TJ-regulated genes.

To validate this mechanism, we cloned the WT LTRs of the *gypsy5* and *mdg1* TEs into our promoter luciferase reporters (Figure 6C). We also mutated these LTRs by deleting just the 6 bp core sequence of the TJ motif (5′-GTCAGC-3') found in these TE LTRs. The mutated LTRs that lacked only the TJ motif were diminished in promoting luciferase expression by 80% and 70% compared to the WT LTRs of *gypsy5* and *mdg1*, respectively (Figure 6C). The WT LTRs were more active in promoting reporter expression during siPiwi knockdown and failed to promote expression with siTj knockdown. In total, our data support the conclusion that multiple TEs have hijacked TJ to activate their expression in *Drosophila* cells.

DISCUSSION

The *Flam* locus is the most prominent piRNA cluster locus expressed in *Drosophila* ovarian follicle cells that is required for female fertility and silencing of *gypsy/mdg4* TEs. Although

two known *P element*-induced mutations in the *Flam* promoter region cause the loss of *Flam* piRNAs and female infertility, the *Flam* promoter and enhancer architecture was poorly understood. Our study combined promoter-bashing reporter assays, distinct *Flam* mutants deleting the enhancer elements, proteomics of the proteins binding these elements, and multiple transcriptomics and genomics approaches to validate that the TJ transcription factor binds *FPSE1* to directly drive *Flam* expression.

Our study also uncovers a complex interplay whereby the LTR TEs also hijack TJ for their own activation (Figures 6C and 6D). Given TJ's high abundance in *Drosophila* follicle cells and the selfish motivations of TEs to replicate and resist host silencing mechanisms, it may seem straightforward for TE promoter sequences in LTRs to evolve the TJ binding motif. To achieve détente in the TE-host genome arms race, the *Drosophila* Piwi pathway then responded by also co-evolving multiple TJ binding motifs at the protein-coding and piRNA-generating loci.

Perhaps TJ became an unwitting pawn in the TE-Piwi pathway chess game in *Drosophila* follicle cells. Now that we have defined the thousands of genes directly bound by this large MAF transcription factor (Tables S1 and S2), the main Gene Ontology terms linked to the TJ ChIP-seq are gonadal genesis, morphogenesis, and other development processes (Figures S7F–S7H). TJ is also expressed in the fly brain in a subset of neurons.[39,40] Although *Tj*-null adult mutants are completely sterile from gonad collapse, the larval gonads and neurons seem to tolerate the lack of TJ, including by still expressing *piwi* in *Tj*-null larval gonads.[13] This suggests that additional factors besides TJ in follicle cells regulate the Piwi pathway, like *ovo*, for example.[41]

Limitations of the study

Future studies will be needed to uncover the hidden roles of *FITS* and *FPSE2* that, in mutant flies, appear to impact fertility but mysteriously do not significantly affect *Flam* piRNA expression. Likewise, the *P element* insertions in $Flam^{BG}$ and $Flam^{KG}$ mutants are in regions distinct from *FPSE1*, which has the clearest binding sites for TJ, so future chromatin conformation interrogations and TJ ChIP-seq experiments will help illuminate how TJ binding at the *Flam* promoter region may be disrupted in these mutants. However, our study and the accompanying study[31] both confirm and expand upon the previous studies first linking TJ and the somatic piRNA pathway in flies[13,28] by showing that the TJ protein activates multiple somatic Piwi pathway genes, including the *Flam* piRNA cluster locus.

Will TJ's regulation of TEs and the Piwi/piRNA pathway be conserved in other dipterans? At evolutionary timescales, the *Flam* locus is evolving rapidly within drosophilids and is only conserved out to *D. erecta* and *D. yakuba*, which are ~10 million years diverged from *D. melanogaster*.[10,42] The configurations of the homologous *Flam* loci in *D. erecta* and *D. yakuba* are also quite distinct from *D. melanogaster*,[10,42,43] and in the more closely related *D. simulans* species, the massive genomic region containing *Flam* and adjacent protein-coding genes has been duplicated.[44] The *Tj* gene is conserved in dipterans like the mosquito genomes from *Anopheles* and *Aedes*, but much further study will be needed to assess whether true *Tj* orthologs are conserved in other insects and whether the TJ binding motifs are maintained and found in the other insects' TE sequences.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Nelson C. Lau (nclau@bu.edu).

Materials availability

All unique/stable reagents generated in this study are available from the lead contact. Material transfer agreements with Boston University may apply.

Data and code availability

- All sequencing data produced and generated by this study is available on Sequencing Read Archive (SRA) under BioProject PRJNA1133762 and processed data under GEO: GSE286025. See Table S4 for specific BioSample and SRA accessions. Other data accessed in this study are listed under GEO: GSE256732.
- The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium[45] via the PRIDE partner repository with the dataset identifier PXD052946 and https://doi.org/10.6019/PXD052946.
- The MSRG pipeline code can be found on the GitHub repository: https://github.com/laulabbumc/MosquitoSmallRNA.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

STAR★METHODS

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

OSS cells were authenticated by RNA sequencing while S2 cells were authenticated by morphology inspection under brightfield microscopy, but both were not tested for mycoplasma. Genetic modifications of *Drosophila melanogaster* strains are detailed in the other methods details. All strains were maintained on standard *Drosophila* food medium at 25°C in a 12h light/12h dark cycle. The origins of all fly and cell line strains are listed in the key resources table.

METHOD DETAILS

Flam promoter bashing and LTR promoters luciferase reporter assays: The pNanoLucTK plasmid (accession #KM359774) was re-engineered into pNN by replacing the TK promoter with a KpnI site for cloning various PCR-amplified genomic fragments of the *Flam* promoter/enhancer region from gDNA extracted from OSS cells. All the various *Flam* PCR fragments were generated with primers listed in Table S3, in combinations detailed in Figures 1 and S1. PCR fragments were cloned into pNN with the NEB HiFi Assembly Kit. PCR-based mutagenesis and HiFi Assembly generated additional deletions within the longer *Flam* promoter plasmids. To further engineer the artificial intron into the

Flam promoter reporter plasmid to make the base pNN flexible, the artificial intron was cut out from a *Renilla* luciferase plasmid that were the *Flam*-piRNA targeted luciferase reporter plasmids previously generated in Post et al.[17] This artificial intron was then cloned right in front of NanoLuc at the NcoI site.

To generate the Wild-type (WT) *gypsy5* and *mdg1* LTR luciferase reporters, PCR primers with KpnI sites were used to amplify the LTR elements from gDNA. The 6bp core sequence of the TJ motif (5′-GTCAGC-3') was deleted with mutagenesis primers listed in Table S3. The WT and mutated *gypsy5* and *mdg1* LTRs were then cloned into the promoter region in the pNN luciferase plasmid. All promoter plasmids were verified either by Sanger sequencing or Plasmidsaurus sequencing.

OSS and S2 cells were grown in T75 flasks at 25°C and harvested as pellets once fully confluent. Cells were prepared for nucleofection using SF cell line solution (Lonza). Cells resuspended in nucleofection solution were aliquoted into separate tubes with 3nmol of siRNA. Once mixed cell suspension was placed in Nucleocuvette vessel (Lonza) and nucleofection occurred in the X-unit of the 4D-Nucleofector Core Unit (Lonza) using zapping code EN-100. Nucleofected cells were placed immediately in 2X F.I.G. Media and plated in triplicates on 24-well plates at a 1:60 dilution.

Two days post-nucleofection of siRNAs, the reporter plasmid mix was prepared, per well, by diluting 0.25μg of pMALF firefly luciferase plasmid and 0.05μg of the *Flam* promoter Nano luciferase reporter plasmids in 0.15M NaCl. For *mdg1*-LTR reporter experiments we used 1ug pMALF and 0.01ug NN-mdg1 reporter per well. For *gypsy5*-LTR reporter experiments used 0.5ug pMALF and 0.5ug pNN-*gypsy5*-LTR reporter. For gypsy5-LTR luciferase assay we used per well 100uL of 1X Passive Lysis Buffer (PLB) instead of the standard 200uL amount.

FuGENE HD (Promega) was added to the reporter mix at 2.5μL per well and final mix was incubated for 15 minutes at room temperature. Final reporter mix was added dropwise to each well and incubated at 25°C over the next two days. Cells were lysed in 1X Passive Lysis Buffer (Promega) on a rocker for 20 minutes at room temperature. Luciferase activity was measured with the ONE-Glo Stop&Glo kit (Promega) using a BioTek Synergy HT plate reader and Gen 5 program, to give the reporter expression.

Generation of distinct flam Drosophila mutants and Drosophila RNAi crosses: For the lines *FITS* and *FPSE2A/B*, Cas9 small guide RNAs and ssODNs were designed based on the *Flam* promoter bashing reporter assay results to target regions located on or upstream of the first exon of *Flam*. These oligonucleotides or their sequences were sent to Rainbow Transgenic Flies (RTF) Inc to be injected into fly embryos of the *nos-Cas9 attp2* background. When flies reached adulthood, they were crossed to *w1118* and after successful mating parents were screened for mutations by gDNA extraction. Progeny from crosses containing the mutation were collected, based on our crossing scheme, and crossed to the balancer line *Adar/FM7a*. Balancer crosses occurred over three more generations and parents were assessed for mutation by gDNA PCR after each cross. For lines *FPSE1, FPUN and FPRU*, the sequences of guides and ssODNs were consulted with RTF Inc and

then sgRNA and ssODN injections, screening, and initial balancing were also conducted by RTF Inc. The *FPRU* mutant was an unplanned allele while trying to generate targeted *FPRU* mutant. To keep background consistent in all lines, *FPSE1, FPUN and FPRU* were rebalanced over FM7a once flies were received by RTF Inc.

Several *Drosophila* RNAi crosses were prepared with different conditions as well as different GAL4 driver lines. The first set of crosses were single fly crosses consisting of a female from the line +/+; Tj-GAL4,Gypsy-LacZ/CyO; P-Tub-GAL80TS/TM6-tb, created in the Lau Lab, and a male from one of the following RNAi lines: y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8] = TRiP.HMS01069}attP2 (RNAi-Tj#1), w; tjDf1/CyO; UAS-tjRNAi on chr 3 (RNAi-Tj#2, https://doi.org/10.1242/dev.089896), y[1] v[1]; P{y[+t7.7] v[+t1.8] = TRiP.JF02009}attP2 (RNAi-Tj#3), y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8] = TRiP.HMS00606}attP2 (RNAi-Piwi#1) and two negative control strains: (RNAi-White: y[1] v[1]; P{y[+t7.7] v[+t1.8] = TRiP.HMS00004}attP2/TM3, Sb[1]) and (RNAi-mCherry: Valium20 Control shRNA, BDSC#: 35785). Crosses were kept at 18°C until pupae began to form then they were raised to 28°C until progeny eclosed. The second set of fly crosses follow the same conditions except they were raised to 28°C during early adulthood. The final set of crosses were kept at room temperature and consist of single fly crosses between a female late-stage follicle cell GAL-4 driver, CG13083.89K-GAL4[M4]/Cyo, Dfd-eYFP (FC0), and a male from one of the previous RNAi lines. Female progeny from all crosses were collected according to the predicted knockdown phenotype and kept in vials for ~2 days to ensure ovary development. Ovaries were then dissected and inspected under a dissecting microscope to examine knockdown efficiency by ovary morphology.

Drosophila female fertility assays and RT-qPCR from flam mutant ovaries: *Flam* mutant lines were assessed through an ovary squash assay and fecundity. Single sibling crosses containing either wild-type or mutant females and mutant males were set up and mated for 3–4 days at 28°C. Upon confirming successful mating flies were removed and vials placed back at 28°C. Females were carefully squashed between a microscope slide and coverslip containing green food coloring. The number of eggs per female were counted and compared between wild-type and mutants. Following this assay viable progeny from each sibling cross were counted and compared between wild-type and mutants.

For each line 20 ovaries, from both wild-type and mutant females, were dissected in cold 1X PBS. Total RNA was extracted with Monarch Total RNA Miniprep Kit (NEB) following the manufacturer's instructions, including the DNase I treatment step. For RT-qPCR 1μg of total RNA was used for first strand cDNA preparation by using ProtoScript II Reverse Transcriptase kit (NEB) followed by LUNA Universal qPCR kit (NEB). RT-qPCR was performed in a Bio-Rad CFX Opus 96 Real-Time PCR System. Rp49 was used as the housekeeping gene to normalize downstream flamenco expression by 2-DDCt methods.

RNA profiling of OSS cells after siRNA knockdown: OSS cells were grown in T75 flasks at 25°C and harvested as pellets once fully confluent. Cells were prepared for nucleofection using SF cell line solution (Lonza). Cells resuspended in nucleofection solution were aliquoted into separate tubes with 3nmol of siRNA. The mixed cell suspension was placed in Nucleocuvette vessel (Lonza) and nucleofection occurred in the X-unit of the

4D-Nucleofector Core Unit (Lonza) using zapping code EN-100. Nucleofected cells were placed immediately in 2X F.I.G. Media and plated in triplicates on 6-well plates at a 1:60 dilution. Two days after nucleofection total cell RNA was harvested using Monarch Total RNA Miniprep Kit (NEB) following the manufacturer's instructions, including the DNase I treatment step. For RT-qPCR 1μg of total RNA was used for first strand cDNA preparation by using ProtoScript II Reverse Transcriptase kit (NEB) followed by LUNA Universal qPCR kit (NEB). RT-qPCR was performed in a Bio-Rad CFX Opus 96 Real-Time PCR System. Rp49 was used as the housekeeping gene to normalize downstream flamenco expression by 2-ΔΔCt methods.

Utilizing RNA from previous extraction from flamenco mutant ovaries and OSS knockdown cells to perform library preparations. Small RNA libraries were made using NEBNext Small RNA Library Prep Set (NEB #E7330) with up to 5μg of RNA input, and a series of terminator blocking oligos were added prior to linker ligation to suppress the 2S rRNA in the library preparation following the guidelines in.[66] Library samples were amplified up to 25 total PCR cycles. Total RNA libraries were made using the Zymo-Seq RiboFree Total RNA Library Kit using up to 250ng RNA input following manufacturer's protocol. All libraries were checked on an Agilent Bio-analyzer 2100 using either DNA 1000 or High Sensitivity DNA kit and sequenced at the BUMC Microarray and Sequencing Resource on an Illumina Next Seq 2000 using P3 flow cells for 50SE and 50PE reads for small RNA and total RNA libraries, respectively.

Proteomics analyses of OSS cells proteins binding to FITS, FPSE1 and FPSE2 DNA segments: OSS and S2 cells were grown at 25°C in T75 flasks and harvested as pellets once fully confluent. Pellets were resuspended in 1X CST lysis buffer (Cell Signaling Technologies) and sonicated five times at 30% amplitude for a cycle of 15 seconds on/60 seconds off using the Q125 sonicator (QSonica). Cell lysate was spun down at 13000rpm for 10 minutes at 4°C and transferred to a fresh tube. Biotinylated DNA probes for FITS, FPSE1, and FPSE2 were prepared by using one normal PCR primer and one biotinylated primer designed for these specific regions. M-280 Streptavidin Dynabeads (Invitrogen) and 20pmol of biotinylated DNA probe were combined and incubated for 1 hour at 4°C. Lysates were pre-cleared by incubating with 1μg of Streptavidin Magnetic Beads (NEB) on a rotator at room temperature for 30 minutes.

After pre-clearing, the lysates were incubated on ice with 200μg of Salmon sperm DNA (Sigma) for 5 minutes. This was followed by incubating lysates with biotinylated DNA probe for 1 hour at room temperature on a rotator. Once this final incubation was completed beads were washed three times in wash buffer followed by three washes in 1X PBS. Beads were resuspended in 100μL of 1X PBS and 80μL were sent off to the Center for Network Systems Biology (Boston University) for Mass Spectrometry while the remaining 20μL were analyzed by SDS-PAGE followed by silver staining with the Pierce Silver Stain for Mass Spectrometry kit (Thermo Scientific).

To prepare the DNA-pulled down proteins immobilized on streptavidin beads for mass spectrometry analysis, the samples were eluted by on-bead digestion in 40 μl digestion buffer (0.6M GuHCl, 4 mM Chloroacetamide, 1 mM TCEP). 0.75 μg trypsin was added to

each replicate and trypsinization was carried out overnight at 37°C. Samples were acidified with 0.1% TFA and the tryptic peptides were desalted before being quantified by BCA protein assay (Pierce). Peptides from each replicate were TMT-labelled (ThermoFisher Scientific) as per the manufacturer's protocol. Pooled TMT labeled peptides were desalted and fractionated offline via high pH reversed-phase chromatography to obtain five fractions using different concentrations of acetonitrile (5%, 10%, 15% 20% and 50%). The peptides were dried in a SpeedVac.

The dried TMT-labelled samples were resuspended in mobile phase A (0.1% Formic acid and 2% acetonitrile) for analysis on the ThermoFisher Exploris 480 instrument interfaced to the Easy nanoLC 1200 HPLC system (ThermoFisher Scientific). The Exploris 480 mass spectrometer was operated with the FAIMS Pro interface using three compensation voltages (CV = −50, - 57 and −64). The samples were loaded onto a reversed-phase C18 column (75 mm i.d. × 2 cm, Acclaim PepMap100 C18 3 mm, 100 Å, ThermoScientific) in mobile phase A and resolved over a 120-min gradient of mobile phase B (0.1% Formic acid, 80% acetonitrile) using an EASY-Spray column (ES803A, Thermo Scientific) maintained at 45° C. The instrument was operated in Data-Dependent mode (DDA) using an MS1 resolution of 60,000 and a scan range of 350–1400 m/z with a normalized AGC target of 300% (3e6 ions) and an injection time set to auto mode. MS2 scans were acquired at a resolution of 45,000 via HCD fragmentation with a normalized AGC target of 100% (1e5 ions) and an injection time of 80 ms. The duty cycle time was set to 0.8 s.

Bioinformatics analysis of the raw MS files started with splitting the channels into the respective individual CV files and analyzed by MaxQuant (Ver. 1.6.7.0) that integrates the Andromeda search engine. The raw data were searched against the *Drosophila melanogaster* database from UniProt (downloaded on 11/8/22). Enzyme specificity was set to Trypsin and up to 2 missed cleavages were allowed. Cysteine carbamidomethylation was specified as fixed modification whereas oxidation and N-terminal acetylation were set as variable modifications. Protein and peptide identifications were filtered at 1% FDR using the target-decoy strategy.[67] Peptide precursor ions were searched with a maximum mass deviation of 6 ppm and fragment ions with a maximum mass deviation of 20 ppm. The MaxQuant output file designated "ProteinGroups.txt" was used for normalization and further bioinformatic analyses. Subsequent bioinformatic analyses was performed in R. The protein group intensities identified were Log-transformed and normalized by Loess. Differential analysis was performed using the Limma R package.[68] Moderated t-tests were corrected with the Benjamini-Hochberg method for false discovery rate (FDR).

Western blots of traffic jam and other piwi-pathway proteins: OSS cells were nucleofected to include siRNAs and then plated on 6-well plates to grow at 25°C. Once confluent, cells were harvested as pellets and resuspended in 2X SDS Sample Buffer. They were then incubated at 100°C for 5 minutes. Denatured samples were then run on a 10% mini-PROTEAN TGX Gel (Biorad) at 120V for 1 hour and 5 minutes. Using the Trans-Blot Turbo Transfer System and RTA Transfer Kits (Biorad) the proteins were then transferred onto a PVDF membrane following the quick start guide provided and the mini-TGX protocol on the system. After transfer membranes were blocked with 5% non-fat milk for 1 hour at room temperature. Next primary antibodies were prepared at a 1:5000 dilution

in 5% non-fat milk and incubated with blots overnight at 4°C. The following day blots were washed for 5 minutes three times with cold 1X PBS. Next secondary antibodies were diluted 1:5000 and incubated with blots for 2 hours at room temperature. After incubation blots were washed for 5 minutes three times with cold 1X PBS. Blots were then imaged using the Odyssey CLx Imager (Li-Cor).

The distinct primary antibody specific to Traffic Jam (TJ) was generated in rabbits immunized against TJ peptide, KMEDPTINDTYVQEFD-Cys by Pacific Immunology Corp. The antibodies were then affinity purified by using SulfoLink Coupling Resin Columns (Thermo Scientific).

Cut&Run of H3K9me3 and H3K4m3 marks in OSS cells: OSS cells were grown in T75 flasks at 25°C and some OSS cells also were nucleofected with siRNAs to create siRNA Knockdown samples. Once cells were fully confluent, they were fixed with 1mM 3,3-Dithiodipropionic acid di(N-hydroxysuccinimide ester) or DSP (Sigma-Aldrich) for 30 minutes at room temperature. After fixation DSP was quenched with 110mM Glycine and cells were harvested. Samples then were prepared using the CUTANA ChIC/CUT&Run Kit Version 3 (Epicypher) with the following changes and additions: additional antibodies used up to 0.5µg were H3K9me3 Polyclonal (Invitrogen), TJ39/40, & TJ5 (Godt). During digest with pAG-Mnase samples were incubated for 1 hour at 4°C. After targeted chromatin digestion and release an additional reverse cross-linking step which calls for 0.1% SDS and 0.2 mg/mL Proteinase K to be added to samples and then incubated at 70°C for 10 minutes. After samples cross-links were reversed up to 5 ng was used to prepare libraries with the CUT&RUN Library Prep Kit Version 1.1 (Epicypher). Library preps were sequenced at the BUMC Microarray and Sequencing Resource on an Illumina Next Seq 2000 using 50×50bp paired end sequencing on a P2 flow cell for 100 cycles.

TJ ChIP-Seq from OSS cells: OSS cells were grown in T75 flasks at 25°C and fixed with 2% p-formaldehyde in 1X PBS for 10 minutes at room temperature. Cells were then quenched with final concentration 0.125M glycine and harvested using a cell scrapper. Cells were pelleted and washed two times with cold 1X PBS +1% BSA. Next the cell pellet was resuspended in 500µL of MNase Lysis Buffer supplemented with 1Mm DTT and 0.2% SDS. Cell lysate was then transferred to Micro Tube TPX Plastic tubes (Diagenode) and sonicated with medium wave-length for 10 cycles of 15 seconds on/30 seconds off using the Bioruptor UCD-200 (Diagenode). Samples were moved back to normal tubes and cell debris was pelleted. Supernatant was moved into a fresh tube and brought up to 2mL with ChIP-Dilution Buffer.

Supernatant was divided evenly and incubated at 4°C overnight with 2µg of antibody. The following day 40µL of Protein A magnetic beads (Dynabead) were prepped by washing once with 1X PBS and twice with ChIP Dilution Buffer. After washes beads were mixed with supernatant and incubated for 4 hours at 4°C. Once incubation was complete the following washes were performed: twice in ChIP Dilution Buffer, once in 1X TE, five times with ChIP-RIPA Buffer, twice with ChIP-RIPA/500 Buffer, twice with ChIP-LiCl Buffer, and last two washes with 1X TE. After all washes samples were eluted with 50µL 1X Elution Buffer and incubated overnight at 65°C to reverse cross-links.

On the third day elution was removed from beads and 20μg of RNase A was added to each sample then incubated at 37°C for 30 minutes. Next 100μg of Proteinase K was added to samples and incubated for 90 minutes at 37°C. After incubations samples were purified using the Monarch DNA/PCR Cleanup Kit (NEB), performing an extra wash during washing steps. Samples were eluted out in 50μL of nuclease-free water. Proceeded to DNA library prep using 25ng of input and following the standard protocol for the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB). During the PCR amplification, for target enrichment, all samples were run for only 9 cycles. All libraries were checked on an Agilent Bioanalyzer 2100 using either DNA 1000 or High Sensitivity DNA kit and sequenced at Novogene using Illumina Novaseq on a XPlus 10B flowcell.

Bioinformatics analysis of long RNA and small RNA differential expression: Small RNA libraries were sequenced as 50nt SE reads. Some libraries were sequenced multiple times to ensure sufficient sequencing depth of >5 million reads. Files from the same library were merged. Sequencing data was checked using FastQC.[69] Small RNA samples were run through the Mosquito Small RNA Genomics (MSRG) pipeline.[70,71] Outputs from MSRG include the alignment to genes and TEs. Total RNA libraries were sequenced as 50PE reads but were processed to look like small RNA reads so that they could also be run using the MSRG pipeline. All reads were trimmed to 35nt, R1 was reverse complemented to be in the same orientation as R2, and then the reverse complemented R1 was merged with R2. The resulting FASTQ file was then run through MSRG and processed just like the small RNA libraries.

Differential expression for OSS knockdowns were done in R using DESeq2[57] on gene counts from the MSRG pipeline. The “cook-sCutoff” and “independentFiltering” arguments in the results function were set to FALSE. The piRNA cluster counting was performed from the SAM file from MSRG was filtered based on region of interest. The following coordinates were used to extract reads mapping to piRNA clusters: flamTSS|chrX:21631207–21672456, DocStalk4|chrX:21672392–21686821, flamMDG1|chrX:21687094–21704920, flamMidEnd|chrX:21705960–22193174, flamTailEnd|chrX:22193175–22430696, 20A|chrX:21521000–21557000, piwi| chr2L:10982205–10987420, tj|chr2L:19464186–19467898, brat|chr2L:19164849–19180120. Reads per million (RPM) was determined in R by counting these reads and normalizing to library size from MSRG length distribution output.

Bioinformatics analysis of histone Cut&Run datasets and TJ ChIP-seq datasets: CUT&RUN data underwent the same quality control and alignment procedures as ChIP-seq data, using FastQC[69] and Bowtie2.[51] However, duplicate reads were retained due to the nature of CUT&RUN, which is reasonable to generate duplicate reads. Peaks were called against IgG controls, keeping all duplicates with MACS2. H3K4me3 peaks were called using narrow peak mode, while H3K9me3 peaks were called using broad peak mode. Unlike ChIP-seq, blacklist regions were not filtered because CUT&RUN data has different characteristics, and no specific blacklist file is available. Additionally, CUT&RUN data was mapped to the *E. coli* genome to verify spike-in proportions. The calculated scale factors from this mapping were used to normalize the BigWig files.

Raw sequencing reads from TJ OSS ChIP-seq were assessed for quality using FastQC. Adapter sequences and low-quality bases were trimmed using Trimmomatic with parameters: ILLUMINACLIP: TruSeq3-PE-2.fa:2:30:10:8, LEADING:3, TRAILING:3, SLIDING-WINDOW:4:15, MINLEN:36. The cleaned reads were then aligned to the dm6 reference genome with Bowtie2[51] using default settings.

The Whole Fly GFP-TJ ChIP-seq dataset was downloaded from the modENCODE repository under the accessing numbers BioProject PRJNA63463, Accession GSE256732. The source fly for this experiment is located at the Bloomington Drosophila Stock Center BDSC as strain #66391, and it has the genotype: w[1118]; PBac{y[+mDint2] w[+mC] = tj-GFP.FPTB}VK00033 (Chr 3, 65B2, 3L:6442676.). Raw FASTA files were trimmed using Trimmomatic with the same parameters as OSS ChIP-seq, but in single-end mode and single-end TruSeq adapters. Following trimming, the reads were aligned to the dm6 reference genome using BWA in aln mode.[53]

For both OSS and GFP-TJ data, the downstream analysis followed the same workflow. Sambamba[60] filtered out duplicate and unmapped reads with the criteria -F “[XS] = = null and not unmapped and not duplicate”. Samtools[72] was used to sort and index the uniquely mapped reads.

MACS2[58] identified peaks, using input samples as controls under default parameters. The consensus peak set for all three replicates was ensured by IDR,[61] evaluating peak sets across true replicates, pooled-pseudo-replicates, and pseudo-self-replicates to derive the final peak set with the highest number of peaks. Finally, blacklist regions were excluded after IDR peak calling using ENCODE’s blacklist version 2[73] with Bedtools.[55] Motifs were identified using MEME Suite[33] for both TJ OSS ChIP-seq and GFP-TJ ChIP-seq data. Each peak was trimmed to 500 base pairs in total, centered on the peak summit with 250 base pairs extending to each end. The most significant motif was then searched within these 500 bp peak sequences using FIMO[65] with a *p*-value threshold of < 5e-4 (default is *p*-value < 1e-4).

Peak annotation was conducted using ChIPseeker[62,74] with tssRegion = c(−400, 600) to annotate peaks at the gene level. The reference databases used were TxDb.Dmelanogaster.UCSC.dm6.ensGene (Team BC, Maintainer BP 2019 TxDb.Dmelanogaster. UCSC.dm6.ensGene: Annotation package for TxDb object(s) and org.Dm.e.g.,.db (org.Dm.e.g.,.db: Genome wide annotation for Fly). Motif annotation was based on the peak annotation; if a motif was located within a peak, it was assigned to the corresponding gene. Gene enrichment analysis was performed with DAVID[63] and visualized with tree plots using the rrvgo package.[64]

For siTj/siGFP H3K4me3 and H3K9me3 CUT&RUN data, reads were aligned to TE consensus sequences using Bowtie2. The aligned reads were sorted and indexed with Samtools. Scaling factors mentioned in cut and run processing section were applied for BigWig files visualization. TJ OSS ChIP-seq and TJ GFP-tag ChIP-seq reads were also aligned to TE consensus sequences using Bowtie2. The three replicates for each condition were merged into a single BAM file. These aligned reads were sorted and indexed using Samtools. siTj/siGFP RNA-seq reads were aligned to TE consensus sequences using

STAR.[52] TE counts for each sample were calculated using Samtools idxstats. TJ motif patterns and position weight matrices were downloaded from the Fly Factor Survey.[36] FIMO was used to search the motif through all TE sequences with a *p*-value threshold of < 5e-4.

QUANTIFICATION AND STATISTICAL ANALYSIS

All luciferase reporter assays were conducted in biological triplicates with technical triplicate readings for each distinct experimental session, and Wilcoxon tests were applied to determine statistical significance. Biological triplicate assays of egg counts and fecundity assays were conducted on all distinct *Flam*/*FITS*/*FPSE1*/*FPSE2* heterozygous controls and homozygous mutant siblings. Single measurements of small RNA and total RNA libraries were performed on >100 ovaries dissected from distinct *Flam*/*FITS*/*FPSE1*/*FPSE2* heterozygous controls and homozygous mutant siblings. Triplicates of OSS knockdowns were subjected to small RNA and total RNAseq for differential expression that was analyzed in R using DESeq2[57] on gene counts from the MSRG pipeline. A minimum of two replicates of Cut&Run and ChIP-seq experiments were applied to OSS cells. The quantification and statistical analysis of ChIP-seq peaks were determined with MACS2[58] using input samples as controls under default parameters. The consensus peak set for all three replicates was ensured by IDR[61] with the default statistical tests employed within the MEME Suite[33] and FIMO.[65]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Zeba Wunderlich for critical comments on this manuscript. We thank Michael Blower for assistance and use of the confocal microscope and Matthew Layne for the use of the plate reader. We acknowledge fly husbandry and OSS cell culture assistance from Jasmine Pierre, William Dorst, and Margarita Kyza-Karavioti. We thank Dorothea Godt for various fly mutants and an aliquot of the anti-TJ guinea pig antibody. We thank Kim McCall and Jianjun Sun for GAL4 driver flies and advice. We thank Ben Czech Nicholson and Emilie Brasset for sharing previews of their data on a complementary study to our study presented here and for pointing out the modENCODE GFP-TJ ChIP-seq dataset. N.C.L. is funded by NIH/NIGMS (grant R01GM135215 and equipment supplements), and a grant from the Boston University Institute of Sexual Medicine contributed support to N.C.L., R.K.G. and J.Z.

REFERENCES

1. Andersen PR, Tirian L, Vunjak M, and Brennecke J (2017). A hetero-chromatin-dependent transcription machinery drives piRNA expression. Nature 549, 54–59. 10.1038/nature23482. [PubMed: 28847004]
2. Mohn F, Sienski G, Handler D, and Brennecke J (2014). The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila. Cell 157, 1364–1379. 10.1016/j.cell.2014.04.031. [PubMed: 24906153]
3. Robert V, Prud'homme N, Kim A, Bucheton A, and Pélisson A (2001). Characterization of the flamenco region of the Drosophila melanogaster genome. Genetics 158, 701–713. [PubMed: 11404334]
4. Sarot E, Payen-Groschêne G, Bucheton A, and Pélisson A (2004). Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the Drosophila melanogaster flamenco gene. Genetics 166, 1313–1321. [PubMed: 15082550]

5. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, and Hannon GJ (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128, 1089–1103. 10.1016/j.cell.2007.01.043. [PubMed: 17346786]

6. Gebert D, Neubert LK, Lloyd C, Gui J, Lehmann R, and Teixeira FK (2021). Large Drosophila germline piRNA clusters are evolutionarily labile and dispensable for transposon regulation. Mol. Cell 81, 3965–3978.e5. 10.1016/j.molcel.2021.07.011. [PubMed: 34352205]

7. Mevel-Ninio M, Pelisson A, Kinder J, Campos AR, and Bucheton A (2007). The flamenco locus controls the gypsy and ZAM retroviruses and is required for Drosophila oogenesis. Genetics 175, 1615–1624. 10.1534/genetics.106.068106. [PubMed: 17277359]

8. Pelisson A, Song SU, Prud'homme N, Smith PA, Bucheton A, and Corces VG (1994). Gypsy transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the Drosophila flamenco gene. EMBO J. 13, 4401–4411. [PubMed: 7925283]

9. Prud'homme N, Gans M, Masson M, Terzian C, and Bucheton A (1995). Flamenco, a gene controlling the gypsy retrovirus of Drosophila melanogaster. Genetics 139, 697–711. [PubMed: 7713426]

10. Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, and Hannon GJ (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell 137, 522–535, S0092–8674(09)00377–8 [pii]. 10.1016/j.cell.2009.03.040. [PubMed: 19395010]

11. Goriaux C, Desset S, Renaud Y, Vaury C, and Brasset E (2014). Transcriptional properties and splicing of the flamenco piRNA cluster. EMBO Rep. 15, 411–418. 10.1002/embr.201337898. [PubMed: 24562610]

12. Wang QT, and Holmgren RA (1999). The subcellular localization and activity of Drosophila cubitus interruptus are regulated at multiple levels. Development 126, 5097–5106. 10.1242/dev.126.22.5097. [PubMed: 10529426]

13. Saito K, Inagaki S, Mituyama T, Kawamura Y, Ono Y, Sakota E, Kotani H, Asai K, Siomi H, and Siomi MC (2009). A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. Nature 461, 1296–1299, nature08501 [pii]. 10.1038/nature08501. [PubMed: 19812547]

14. Lau NC, Robine N, Martin R, Chung WJ, Niki Y, Berezikov E, and Lai EC (2009). Abundant primary piRNAs, endo-siRNAs, and microRNAs in a Drosophila ovary cell line. Genome Res. 19, 1776–1785, gr.094896. 109 [pii]. 10.1101/gr.094896.109. [PubMed: 19541914]

15. Niki Y, Yamaguchi T, and Mahowald AP (2006). Establishment of stable cell lines of Drosophila germ-line stem cells. Proc. Natl. Acad. Sci. USA 103, 16325–16330. 10.1073/pnas.0607435103. [PubMed: 17056713]

16. Sienski G, Dönertas D, and Brennecke J (2012). Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell 151, 964–980. 10.1016/j.cell.2012.10.040. [PubMed: 23159368]

17. Post C, Clark JP, Sytnikova YA, Chirn GW, and Lau NC (2014). The capacity of target silencing by Drosophila PIWI and piRNAs. Rna 20, 1977–1986. 10.1261/rna.046300.114. [PubMed: 25336588]

18. Saito K, Ishizu H, Komai M, Kotani H, Kawamura Y, Nishida KM, Siomi H, and Siomi MC (2010). Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in Drosophila. Genes Dev. 24, 2493–2498, gad.1989510 [pii]. 10.1101/gad.1989510. [PubMed: 20966047]

19. Haase AD, Fenoglio S, Muerdter F, Guzzardo PM, Czech B, Pappin DJ, Chen C, Gordon A, and Hannon GJ (2010). Probing the initiation and effector phases of the somatic piRNA pathway in Drosophila. Genes Dev. 24, 2499–2504, gad.1968110 [pii]. 10.1101/gad.1968110. [PubMed: 20966049]

20. Sumiyoshi T, Sato K, Yamamoto H, Iwasaki YW, Siomi H, and Siomi MC (2016). Loss of l(3)mbt leads to acquisition of the ping-pong cycle in Drosophila ovarian somatic cells. Genes Dev. 30, 1617–1622. 10.1101/gad.283929.116. [PubMed: 27474440]

21. Yamamoto-Matsuda H, Miyoshi K, Moritoh M, Yoshitane H, Fukada Y, Saito K, Yamanaka S, and Siomi MC (2022). Lint-O cooperates with L(3)mbt in target gene suppression to

maintain homeostasis in fly ovary and brain. EMBO Rep. 23, e53813. 10.15252/embr.202153813. [PubMed: 35993198]

22. Arnold CD, Gerlach D, Spies D, Matts JA, Sytnikova YA, Pagani M, Lau NC, and Stark A (2014). Quantitative genome-wide enhancer activity maps for five Drosophila species show functional enhancer conservation and turnover during cis-regulatory evolution. Nat. Genet 46, 685–692. 10.1038/ng.3009. [PubMed: 24908250]
23. Rivera J, Keränen SVE, Gallo SM, and Halfon MS (2019). REDfly: the transcriptional regulatory element database for Drosophila. Nucleic Acids Res. 47, D828–D834. 10.1093/nar/gky957. [PubMed: 30329093]
24. Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, and Siomi MC (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev. 20, 2214–2222. 10.1101/gad.1454806. [PubMed: 16882972]
25. Li MA, Alls JD, Avancini RM, Koo K, and Godt D (2003). The large Maf factor Traffic Jam controls gonad morphogenesis in Drosophila. Nat. Cell Biol 5, 994–1000. 10.1038/ncb1058. [PubMed: 14578908]
26. Panchal T, Chen X, Alchits E, Oh Y, Poon J, Kouptsova J, Laski FA, and Godt D (2017). Specification and spatial arrangement of cells in the germline stem cell niche of the Drosophila ovary depend on the Maf transcription factor Traffic jam. PLoS Genet. 13, e1006790. 10.1371/journal.pgen.1006790. [PubMed: 28542174]
27. Gunawan F, Arandjelovic M, and Godt D (2013). The Maf factor Traffic jam both enables and inhibits collective cell migration in Drosophila oogenesis. Development 140, 2808–2817. 10.1242/dev.089896. [PubMed: 23720044]
28. Robine N, Lau NC, Balla S, Jin Z, Okamura K, Kuramochi-Miyagawa S, Blower MD, and Lai EC (2009). A broadly conserved pathway generates 3'UTR-directed primary piRNAs. Curr. Biol 19, 2066–2076. 10.1016/j.cub.2009.11.064. [PubMed: 20022248]
29. Genzor P, Konstantinidou P, Stoyko D, Manzourolajdad A, Marlin Andrews C, Elchert AR, Stathopoulos C, and Haase AD (2021). Cellular abundance shapes function in piRNA-guided genome defense. Genome Res. 31, 2058–2068. 10.1101/gr.275478.121. [PubMed: 34667116]
30. Sytnikova YA, Rahman R, Chirn GW, Clark JP, and Lau NC (2014). Transposable element dynamics and PIWI regulation impacts lncRNA and gene expression diversity in Drosophila ovarian cell cultures. Genome Res. 24, 1977–1990. 10.1101/gr.178129.114. [PubMed: 25267525]
31. Alizada A, Martins A, Mouniee N, Rodriguez Suarez JV, Bertin B, Gueguen N, Mirouse V, Maupetit-Mehouas S, Rivera AJ, Lau NC, et al. (2025). The transcription factor Traffic jam orchestrates the somatic piRNA pathway in Drosophila ovaries. Cell Reports. 10.1016/j.celrep.2025.115453.
32. Kudron MM, Victorsen A, Gevirtzman L, Hillier LW, Fisher WW, Vafeados D, Kirkey M, Hammonds AS, Gersch J, Ammouri H, et al. (2018). The ModERN Resource: Genome-Wide Binding Profiles for Hundreds of Drosophila and Caenorhabditis elegans Transcription Factors. Genetics 208, 937–949. 10.1534/genetics.117.300657. [PubMed: 29284660]
33. Bailey TL, Johnson J, Grant CE, and Noble WS (2015). The MEME Suite. Nucleic Acids Res. 43, W39–W49. 10.1093/nar/gkv416. [PubMed: 25953851]
34. Bailey TL (2011). DREME: Motif discovery in transcription factor ChIP-seq data. Bioinformatics (Edam) 27, 1653. 10.1093/bioinformatics/btr261.
35. Kataoka K (2007). Multiple mechanisms and functions of maf transcription factors in the regulation of tissue-specific genes. J. Biochem 141, 775–781. 10.1093/jb/mvm105. [PubMed: 17569705]
36. Zhu LJ, Christensen RG, Kazemian M, Hull CJ, Enuameh MS, Basciotta MD, Brasefield JA, Zhu C, Asriyan Y, Lapointe DS, et al. (2011). FlyFactorSurvey: a database of Drosophila transcription factor binding specificities determined using the bacterial one-hybrid system. Nucleic Acids Res. 39, D111–D117. 10.1093/nar/gkq858. [PubMed: 21097781]
37. Senti K-A, Handler D, Rafanel B, Kosiol C, Schlötterer C, and Brennecke J (2023). Functional Adaptations of Endogenous Retroviruses to the Drosophila Host Underlie their Evolutionary Diversification. Preprint at bioRxiv. 10.1101/2023.08.03.551782.

38. Clark JP, Rahman R, Yang N, Yang LH, and Lau NC (2017). Drosophila PAF1 Modulates PIWI/piRNA Silencing Capacity. Curr. Biol 27, 2718–2726.e4. 10.1016/j.cub.2017.07.052. [PubMed: 28844648]
39. Jukam D, Xie B, Rister J, Terrell D, Charlton-Perkins M, Pistillo D, Gebelein B, Desplan C, and Cook T (2013). Opposite feedbacks in the Hippo pathway for growth control and neural fate. Science 342, 1238016. 10.1126/science.1238016. [PubMed: 23989952]
40. Konstantinides N, Kapuralin K, Fadil C, Barboza L, Satija R, and Desplan C (2018). Phenotypic Convergence: Distinct Transcription Factors Regulate Common Terminal Features. Cell 174, 622–635.e13. 10.1016/j.cell.2018.05.021. [PubMed: 29909983]
41. Alizada A, Hannon GJ, and Nicholson BC (2024). Ovo is a master regulator of the piRNA pathway in animal ovarian germ cells. Preprint at bioRxiv. 10.1101/2024.04.23.590802.
42. Chirn GW, Rahman R, Sytnikova YA, Matts JA, Zeng M, Gerlach D, Yu M, Berger B, Naramura M, Kile BT, and Lau NC (2015). Conserved piRNA Expression from a Distinct Set of piRNA Cluster Loci in Eutherian Mammals. PLoS Genet. 11, e1005652. 10.1371/journal.pgen.1005652. [PubMed: 26588211]
43. van Lopik J, Alizada A, Trapotsi MA, Hannon GJ, Bornelöv S, and Czech Nicholson B (2023). Unistrand piRNA clusters are an evolutionarily conserved mechanism to suppress endogenous retroviruses across the Drosophila genus. Nat. Commun 14, 7337. 10.1038/s41467-023-42787-1. [PubMed: 37957172]
44. Signor S, Vedanayagam J, Kim BY, Wierzbicki F, Kofler R, and Lai EC (2023). Rapid evolutionary diversification of the flamenco locus across simulans clade Drosophila species. PLoS Genet. 19, e1010914. 10.1371/journal.pgen.1010914. [PubMed: 37643184]
45. Deutsch EW, Bandeira N, Perez-Riverol Y, Sharma V, Carver JJ, Mendoza L, Kundu DJ, Wang S, Bandla C, Kamatchinathan S, et al. (2023). The ProteomeXchange consortium at 10 years: 2023 update. Nucleic Acids Res. 51, D1539–D1548. 10.1093/nar/gkac1040. [PubMed: 36370099]
46. Rivera A, Lee JHR, Gupta S, Yang L, Goel RK, Zaia J, and Lau NC (2024). Traffic Jam activates the Flamenco piRNA cluster locus and the Piwi pathway to ensure transposon silencing and Drosophila fertility. Preprint at bioRxiv, 2024.08.15.608167. 10.1101/2024.08.15.608167.
47. ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74. 10.1038/nature11247. [PubMed: 22955616]
48. Khodor YL, Rodriguez J, Abruzzi KC, Tang CHA, Marr MT 2nd, and Rosbash M (2011). Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in Drosophila. Genes Dev. 25, 2502–2512. 10.1101/gad.178962.111. [PubMed: 22156210]
49. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, and Li H (2021). Twelve years of SAMtools and BCFtools. GigaScience 10, giab008. 10.1093/gigascience/giab008. [PubMed: 33590861]
50. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, and Mesirov JP (2011). Integrative genomics viewer. Nat. Biotechnol 29, 24–26. 10.1038/nbt.1754. [PubMed: 21221095]
51. Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. 10.1038/nmeth.1923. [PubMed: 22388286]
52. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. 10.1093/bioinformatics/bts635. [PubMed: 23104886]
53. Li H, and Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. 10.1093/bioinformatics/btp324. [PubMed: 19451168]
54. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, and Manke T (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44, W160–W165. 10.1093/nar/gkw257. [PubMed: 27079975]
55. Quinlan AR, and Hall IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. 10.1093/bioinformatics/btq033. [PubMed: 20110278]
56. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, et al. (2019). Welcome to the Tidyverse. J. Open Source Softw 4, 1686.

57. Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. 10.1186/s13059-014-0550-8. [PubMed: 25516281]
58. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, and Liu XS (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137. 10.1186/gb-2008-9-9-r137. [PubMed: 18798982]
59. Bolger AM, Lohse M, and Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. 10.1093/bioinformatics/btu170. [PubMed: 24695404]
60. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, and Prins P (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics 31, 2032–2034. 10.1093/bioinformatics/btv098. [PubMed: 25697820]
61. Li Q, Brown JB, Huang H, and Bickel PJ (2011). Measuring reproducibility of high-throughput experiments. Ann. Appl. Stat 5, 1752–1779.
62. Yu G, Wang LG, and He QY (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics 31, 2382–2383. 10.1093/bioinformatics/btv145. [PubMed: 25765347]
63. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, Imamichi T, and Chang W (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. 50, W216–W221. 10.1093/nar/gkac194. [PubMed: 35325185]
64. Sayols S (2023). rrvgo: a Bioconductor package for interpreting lists of Gene Ontology terms. MicroPubl. Biol 2023. 10.17912/micropub.biology.000811.
65. Grant CE, Bailey TL, and Noble WS (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics 27, 1017–1018. 10.1093/bioinformatics/btr064. [PubMed: 21330290]
66. Wickersheim ML, and Blumenstiel JP (2013). Terminator oligo blocking efficiently eliminates rRNA from Drosophila small RNA sequencing libraries. Biotechniques 55, 269–272. 10.2144/000114102. [PubMed: 24215643]
67. Elias JE, and Gygi SP (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Methods 4, 207–214. 10.1038/nmeth1019. [PubMed: 17327847]
68. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, and Smyth GK (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47. 10.1093/nar/gkv007. [PubMed: 25605792]
69. Andrews S (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
70. Dayama G, Bulekova K, and Lau NC (2022). Extending and Running the Mosquito Small RNA Genomics Resource Pipeline. Methods Mol. Biol 2509, 341–352. 10.1007/978-1-0716-2380-0_20. [PubMed: 35796973]
71. Ma Q, Srivastav SP, Gamez S, Dayama G, Feitosa-Suntheimer F, Patterson EI, Johnson RM, Matson EM, Gold AS, Brackney DE, et al. (2021). A mosquito small RNA genomics resource reveals dynamic evolution and host responses to viruses and transposons. Genome Res. 31, 512–528. 10.1101/gr.265157.120. [PubMed: 33419731]
72. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, and Durbin R; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. 10.1093/bioinformatics/btp352. [PubMed: 19505943]
73. Amemiya HM, Kundaje A, and Boyle AP (2019). The ENCODE Blacklist: Identification of Problematic Regions of the Genome. Sci. Rep 9, 9354. 10.1038/s41598-019-45839-z. [PubMed: 31249361]
74. Wang Q, Li M, Wu T, Zhan L, Li L, Chen M, Xie W, Xie Z, Hu E, Xu S, and Yu G (2022). Exploring Epigenomic Datasets by ChIPseeker. Curr. Protoc 2, e585. 10.1002/cpz1.585. [PubMed: 36286622]

Highlights

- **Fly mutants of** *Drosophila Flamenco* (*Flam*) enhancers exhibit decreased female fertility
- Proteomics of proteins binding *Flam* enhancer identify Traffic Jam (TJ) transcription factor
- TJ knockdown downregulates multiple Piwi pathway genes and retrotransposons
- TJ binds sites in *Flam*, multiple Piwi pathway genes, and transposon long terminal repeats

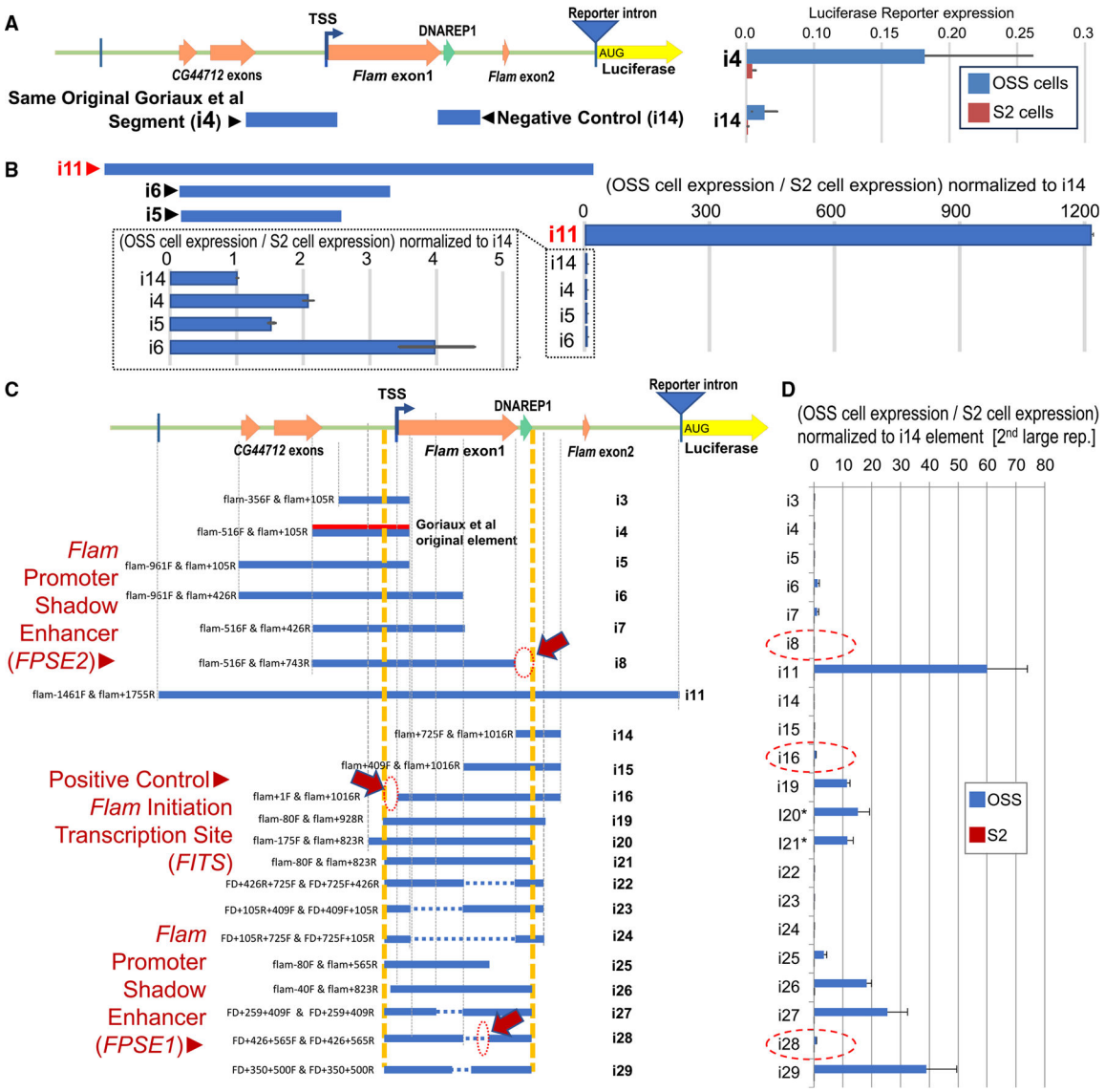


Figure 1. Promoter-bashing reporter assay discovers *Flam* enhancer elements

(A) Diagram of *Flam* promoter region targeted in a distinct format of a promoter-bashing assay that utilizes an intron upstream of the luciferase. Error bars are standard deviation of biological triplicate experiments.

(B) Snapshot of normalized reporter assay results comparing the Goriaux et al. i4 element to the i11 whole *Flam* promoter region. Error bars are propagated ratios of standard deviations from biological triplicate experiments.

(C) Complete diagram of bashing constructs highlighting discovery of *FITS*, *FPSE1*, and *FPSE2* regions.

(D) Corresponding normalized promoter-bashing assay results from the constructs depicted in (C). Error bars are propagated ratios of standard deviations from biological triplicate experiments.

Small text labels are the names of the primers used to generate the constructs. Deletions of interest are marked by dashed ovals. *The constructs i20* and i21* were the first indications that expression in OSS cells required the extra reporter intron. Constructs i3–i19 had similar expression results to constructs that lacked the reporter intron, and i22–i29 were only constructed with the reporter intron.

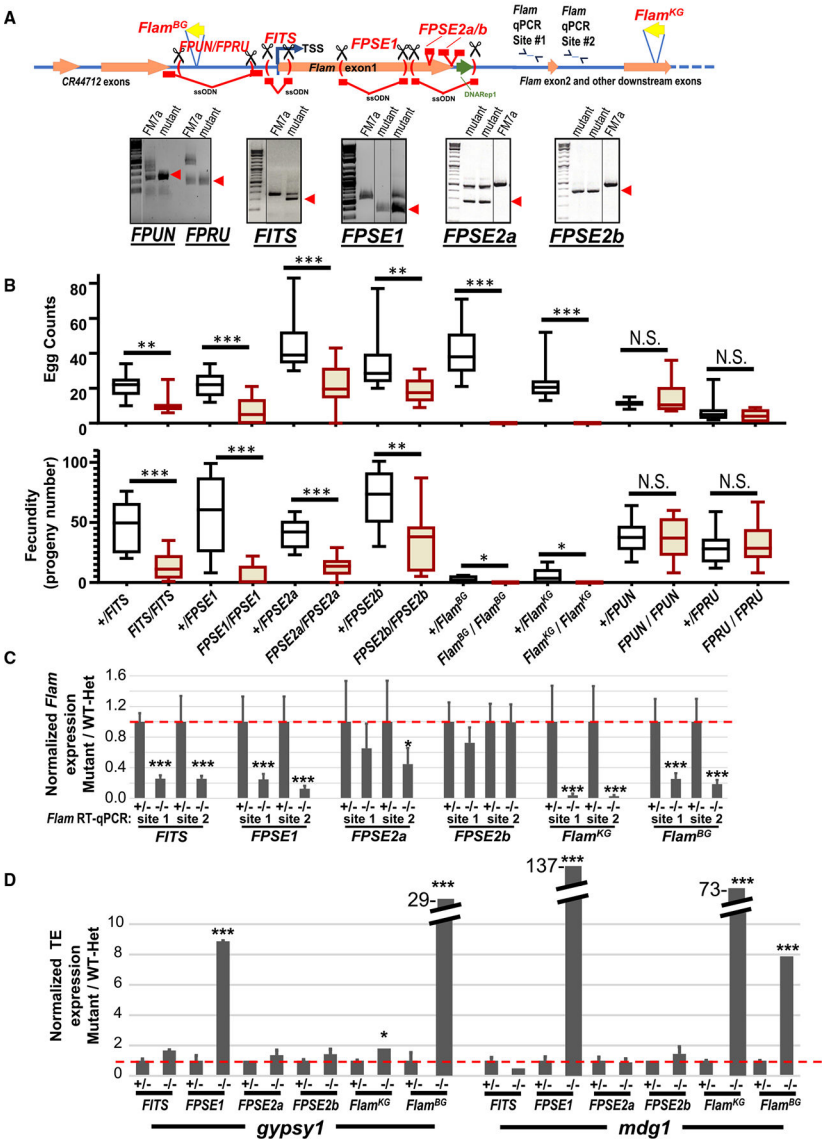


Figure 2. *Flam* mutants with deletions of *FITS*, *FPSE1*, and *FPSE2a*/*2b* regions have compromised female fertility

(A) Diagram of sgRNAs and ssODNs targeting *Flam* promoter/enhancer regions. Genomic PCR shows proper locus deletions in fly mutants compared to the *FM7a* balancer line.

(B) Lower female fertility and fecundity in these distinct *Flam* promoter/enhancer region mutants. $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$; t test. N.S. indicates no statistically significant difference in heterozygous versus homozygous *FPUN* and *FPRU* mutant deleting sequences upstream of the *Flam* TSS, including the region where the *P element* in $Flam^{BG}$ is inserted. Boxplots represent the minimum to maximum with a line at the mean.

(C and D) RT-qPCR of mutants versus heterozygous siblings show (C) *Flam* precursor RNA reduction and (D) *gypsy1* and *mdg1* RNA activation. $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$; Wilcoxon test. Error bars are propagated ratios of the standard deviations.

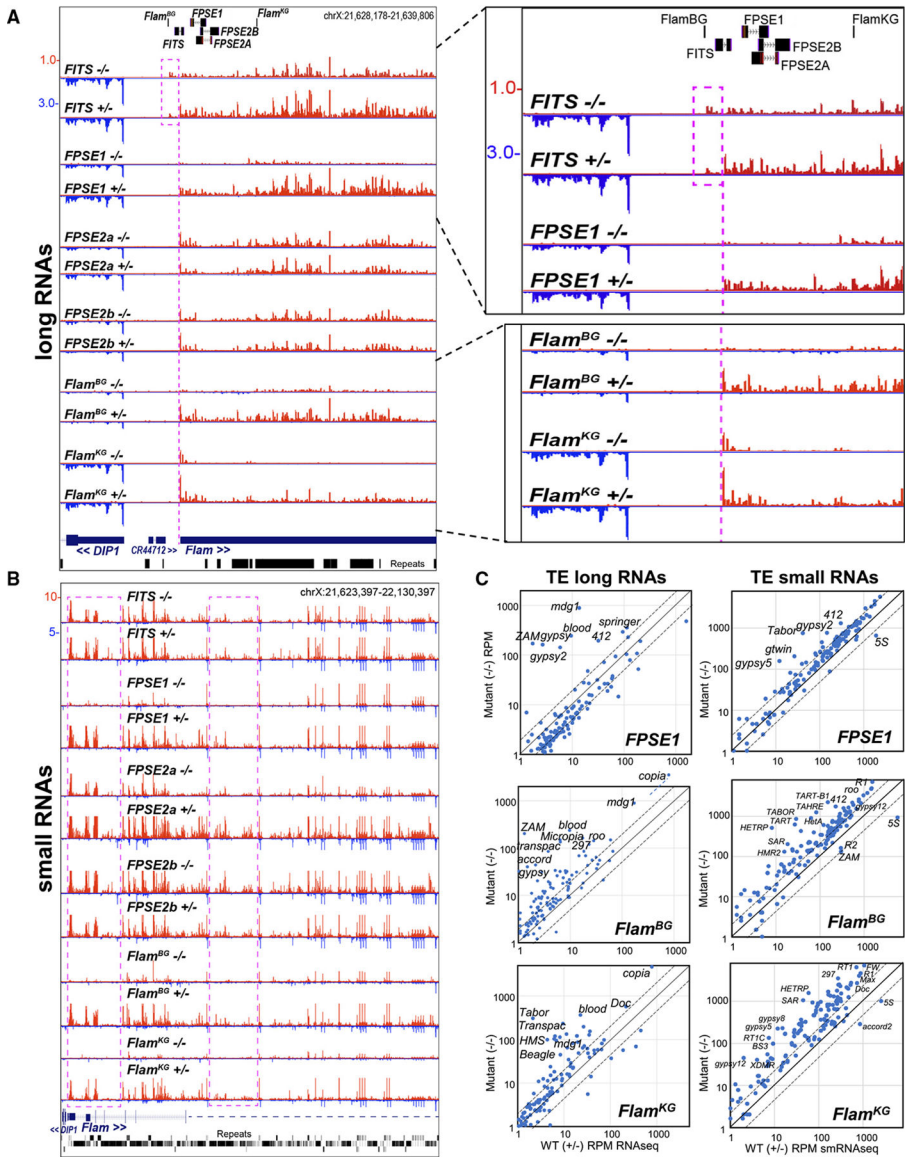


Figure 3. Long and small RNA-seq of ovaries from the *Flam* promoter/enhancer mutants of *FITS*, *FPSE1*, and *FPSE2a*/*2b* compared to the classic $Flam^{BG}$ and $Flam^{KG}$ mutants

(A) Genome browser snapshot of the ovary long RNAs proximal to the *Flam* transcription start site (TSS) in homozygous and heterozygous mutants. The dashed magenta line marks the main TSS of the *Flam* precursor transcript, while the magenta box highlights a distinct upstream TSS in *FITS* mutants that deleted the original TSS region. A zoomed inset is shown on the right.

(B) Genome browser snapshot of the ovary small RNAs throughout the ~500 kb region of *Flam* from homozygous and heterozygous mutants. Magenta boxes mark the *flamTSS* and *flamMidEnd* regions as areas of discernable small RNA accumulation differences, quantified in Figure S3A.

(C) Scatterplots of reads per million for TE consensus elements compared between homozygous and heterozygous mutants for long RNAs (left) and small RNAs (right) for the *FPSE1*, $Flam^{BG}$, and $Flam^{KG}$ mutant ovaries.

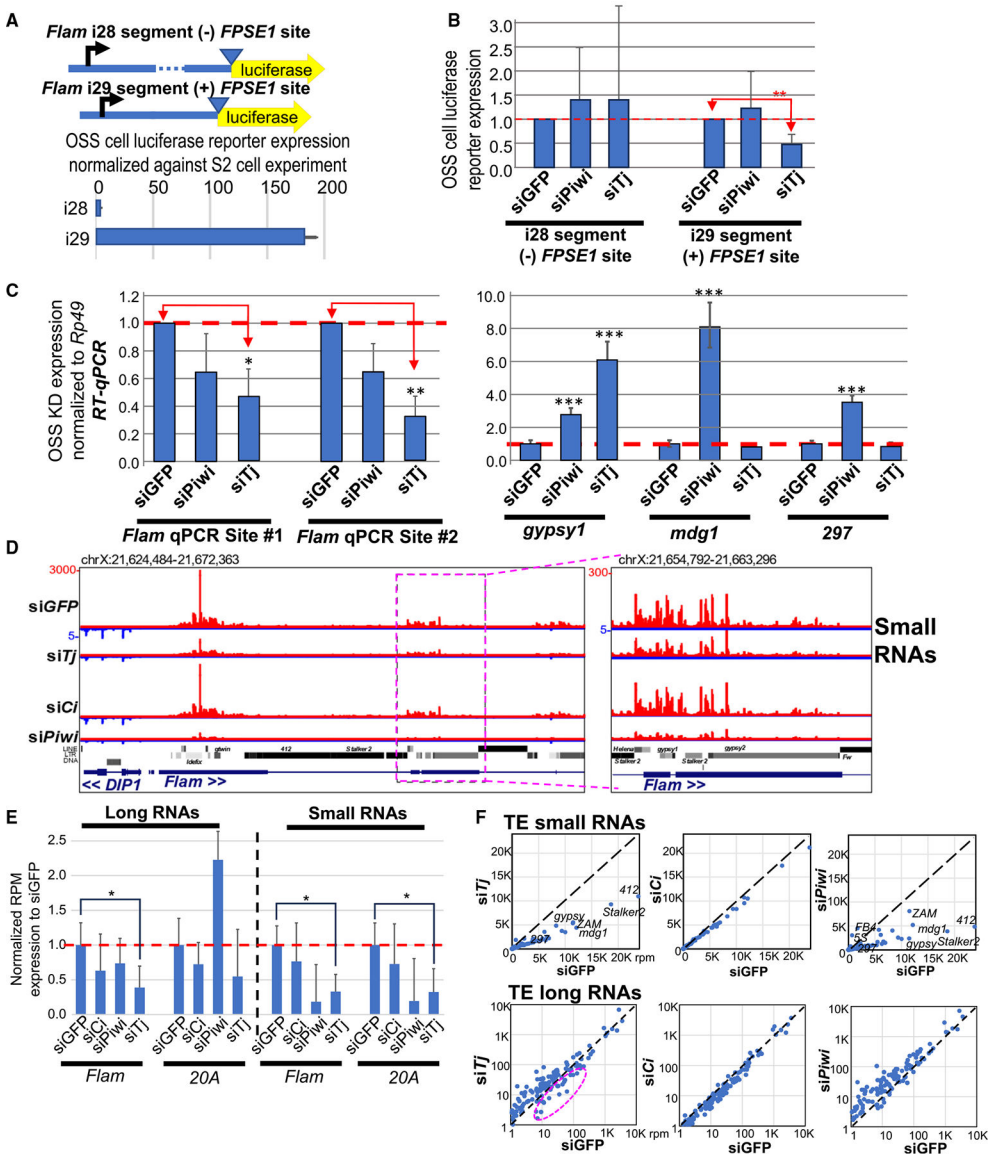


Figure 4. *Tj* knockdown in OSS cells decreases *Flam* promoter reporter expression, *Flam* transcripts, *Flam* piRNAs, and other piRNAs

(A) Diagram of *Flam* promoter reporters that contain or lack putative *TJ* binding region *FPSE1*, which is reflected by the respective luciferase expression levels in OSS cells relative to S2 cells.

(B) Reporter assay results show that only the i29 promoter reporter is negatively affected by the knockdown of *Tj*. $^{**}p<0.01$, Wilcoxon test. Error bars are propagated ratios of standard deviations from biological triplicates.

(C) RT-qPCR of *Flam* precursor RNAs and transposons in OSS cells.$^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$, t test. Error bars are propagated ratios of standard deviations from biological triplicates.

(D) Genome browser tracks showing great reduction of *Flam* piRNAs during siTj and siPiwi knockdowns (average of three biological replicates). All tracks have the same y axis scale as siGFP, with a zoom inset magnified on the right.

(E) Quantification of the total long and small RNAs from the *Flam* and *20A* piRNA clusters during knockdowns and through three biological replicates of deep sequencing. $*p < 0.05$, t test.

(F) Scatterplots comparing the changes in TE small RNAs levels versus TE long RNAs in OSS cells after siTj, siCi, and siPiwi knockdown compared to siGFP control. The dashed magenta oval notes a set of TEs with decreased long RNAs immediately after siTj knockdown.

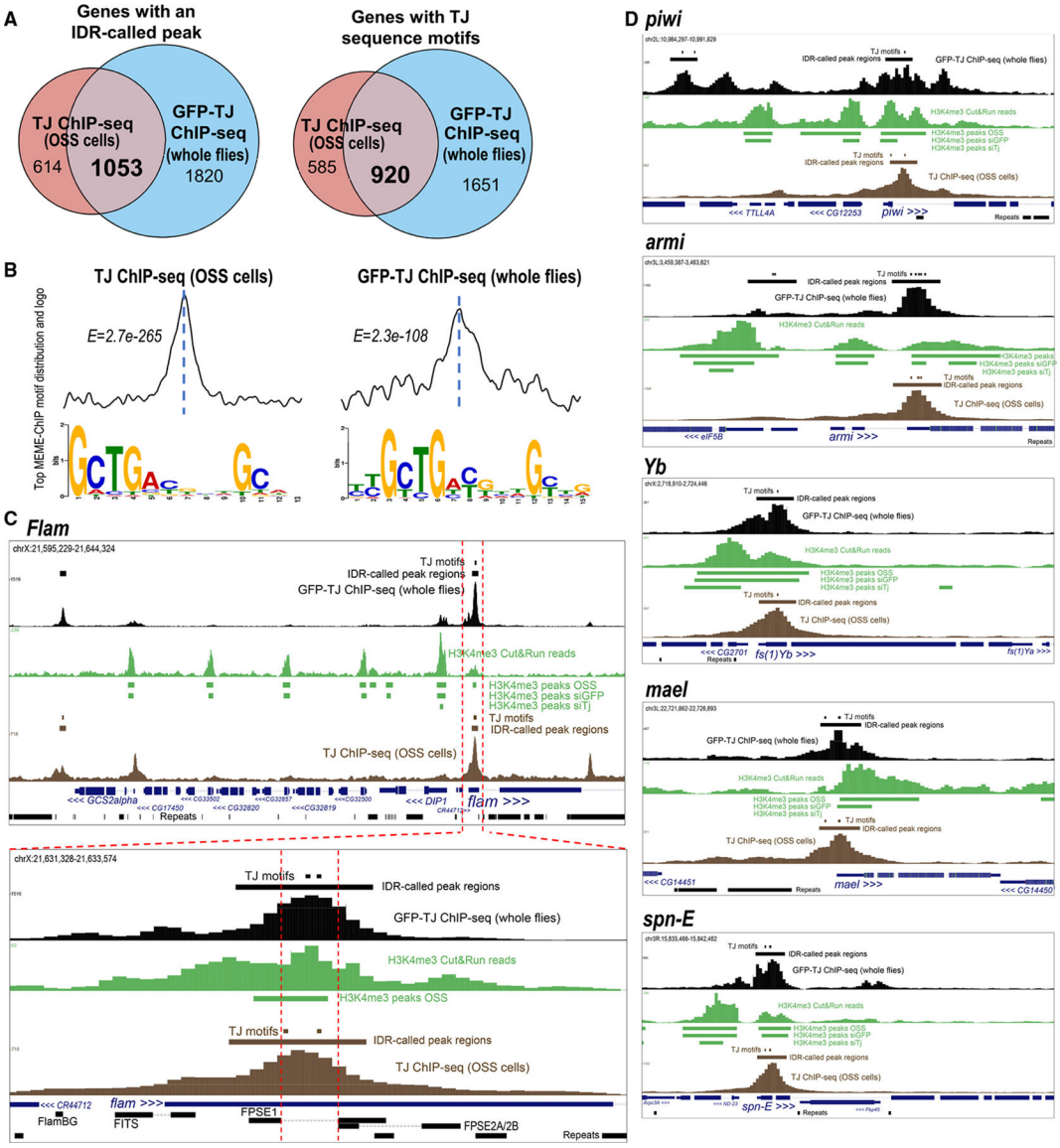


Figure 5. TJ binds the enhancers of *Flam* and multiple Piwi pathway genes to drive piRNA production

(A) Venn diagrams of the number of genes with IDR-called peaks from the whole fly GFP-TJ ChIP-seq and the OSS cell TJ ChIP-seq (left) and the genes with TJ motifs (right).

(B) MEME-ChIP analysis of the top enriched motif sequence from the input of the IDR-called peaks from the OSS cell TJ ChIP-seq and the whole fly GFP-TJ ChIP-seq datasets from MEME. Below each motif signature distribution at the center line is the sequence logo that is consistent with *in-vitro*-determined motifs of MAF transcription factors.

(C) Genome browser tracks of TJ ChIP-seq and Cut&Run profiling of H3K4me3 chromatin marks at the *Flam* piRNA cluster locus. The bottom image zooms in on the *Flam* TSS. Additional dashed red lines delineate the region deleted in the *FPSE1* fly mutant.

(D) Genome browser tracks at five Piwi pathway protein-coding genes with downregulated RNA expression after siTj knockdown in OSS cells.

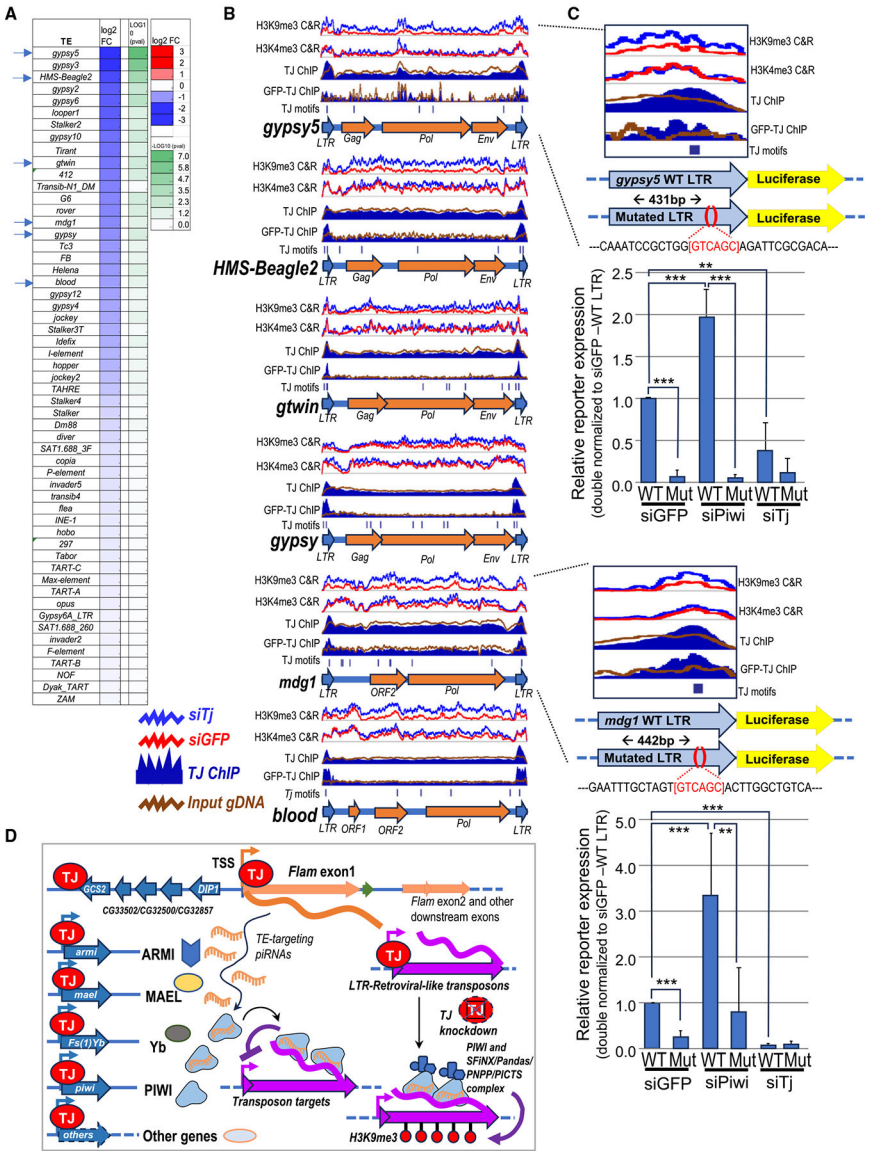


Figure 6. Some *Drosophila* TEs have hijacked TJ to activate their expression

(A) Heatmap showing lower RNA expression for certain retroviral LTR TEs after siTj knockdown. Arrows point to the TEs highlighted in the browser tracks in (B).

(B) Tracks of TJ-ChIP-seq and Cut&Run profiling of H3K4me3 and H3K9me3 chromatin marks at several retroviral LTR TEs actively expressed and silenced by the Piwi pathway in *Drosophila* cells and OSS cells. Additional TE tracks are shown in Figure S8C.

(C) Deeper inspection of the *gypsy5* and *mdg1* LTRs tested for their abilities to promote luciferase reporter expression and dependence on the TJ motif, highlighted in the red color bases in the sequence. Mutated LTRs delete just the 6 bp TJ motif. Graphs show the double-normalized results of 4 and 3 biological replicates for *gypsy5* and *mdg1* LTR reporter assays, respectively. ***$p < 0.001$ and **$p < 0.1$; Wilcoxon test.

(D) Model for TJ regulating both the Piwi pathway and TEs. During siTj knockdown, the TEs lose activation by TJ, which then allows for residual PIWI/piRNA complexes to push more H3K9me3 silencing marks onto TE chromatin before PIWI/piRNAs also decline.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-TJ (rabbit)	This study, Lau lab generated rabbit polycolonal, from Rivera et al.	From Rivera et al.[46]
Anti-TJ (guinea pig)	Godt lab generated guinea pig polycolonal from Li et al.[25]	From Li et al.[25]
Anti-Piwi	Lau lab generated rabbit polycolonal from Sytnikova et al.[30]	From Sytnikova et al.[30]
Anti-Armi (2F8A9)	National Institute of Genetics of Japan	Cat#1011Ab-1
Anti-Yb	National Institute of Genetics of Japan	Cat#1012Ab-1
Anti-Mael	National Institute of Genetics of Japan	Cat#1010Ab-1
Anti-β-tubulin	DSHB	Cat# E7; RRID: AB_528499
Anti-H3K9me3	Invitrogen	Cat# 720093; RRID: AB_2532803
Anti-H3K4me3	Epicypher	Cat# 13–0041
IRDye 680RD Goat anti-Rabbit IgG Secondary Antibody	Licor	Cat# 926–68071
IRDye 800CW Goat anti-Mouse IgG Secondary Antibody	Licor	Cat# 926–32210
Chemicals, peptides, and recombinant proteins		
DNase I	NEB	Cat# M0303S
Protoscript II Reverse Transcriptase	NEB	Cat# M0368
5X Protoscript II Buffer	NEB	Cat# M0368
0.1M DTT	NEB	Cat# M0368
10mM dNTP mix	NEB	Cat# N0447
RNase Inhibitor, Murine	NEB	Cat# M0314
Random Primer Mix	NEB	Cat# S1330
LUNA Universal qPCR Master Mix	NEB	Cat# M3003
10X CST Lysis Buffer	Cell Signaling Technologies	Cat# 9803
Dynabeads M-280 Streptavidin Magnetic Beads	Invitrogen	Cat# 11205D
Streptavidin Magnetic Beads	NEB	Cat# S1420
Salmon Sperm DNA	Sigma-Aldrich	Cat# D7656
FuGENE HD	Promega	Cat# E2311
5X Passive Lysis Buffer	Promega	Cat# E1941
10% mini-PROTEAN TGX Gel	BioRad	Cat# 4561034
DSP	Sigma-Aldrich	Cat# D3669
Proteinase K	NEB	Cat# T2001–1
p-formaldehyde	Electron Microscopy Sciences	Cat# 15710
BSA (Fraction V)	Fisher	Cat# BP1600
Mnase	NEB	Cat# M0247
Dynabeads Protein A Magnetic Beads	Invitrogen	Cat# 10002D
Rnase A	NEB	Cat# T3018
KMEDPTINDTYVQEFD-Cys	Pacific Immunology	Custom peptide sequence
Glycine	Fisher	Cat# BP381

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays and kits		
Monarch Total RNA Miniprep Kit	NEB	Cat# T2010
Pierce Silver Stain for Mass Spectrometry Kit	Thermo Scientific	Cat# 24600
Amaxa SF Cell Line 4D-Nucleofector X Kit L	Lonza	Cat# V4XC-2024
Nano-Glo Dual-Luciferase Reporter Assay System Kit	Promega	Cat# N1610
NEBNext Small RNA Library Prep Kit	NEB	Cat# E7560
Zymo-Seq RiboFree Total RNA Library Kit	Zymo	Cat# R3003
Agilent DNA 1000 Kit	Agilent	Cat# 5067–1504
Agilent High Sensitivity DNA Kit	Agilent	Cat# 5067–4626
Trans-Blot Turbo RTA Mini Transfer Kit	BioRad	Cat# 1704272
Trans-Blot Turbo RTA Midi Transfer Kit	BioRad	Cat# 1704273
CUTANA ChIC/CUT&Run Kit	Epicypher	Cat# 14–1048
CUT&Run Library Prep Kit Version 1.1	Epicypher	Cat# 14–1001
Monarch DNA/PCR Cleanup Kit	NEB	Cat# T1130
NEBNext Ultra II DNA Library Prep Kit for Illumina	NEB	Cat# E7645
SulfoLink Immobilization Kit for Peptides	Thermo Scientific	Cat# 44995
Deposited data		
Traffic Jam ChIP-seq from whole organism	ENCODE Project Consortium.[47]	GEO: GSE256732
Dmel OSS cell Traffic Jam ChIP-seq	This study.	PRJNA1133762, GEO: GSE286025
Dmel OSS cell H3K9me3 Cut&Run	This study.	PRJNA1133763, GEO: GSE286025
Dmel OSS cell H3K4me3 Cut&Run	This study.	PRJNA1133764, GEO: GSE286025
Dmel OSS cell small RNA	This study.	PRJNA1133765, GEO: GSE286025
Dmel OSS cell long RNA	This study.	PRJNA1133766, GEO: GSE286025
Proteomics of proteins binding to Flam enhancer DNA sequence pulldowns from OSS cell lysate.	This study.	ProteomeXchange Consortium[45]: PXD052946
Experimental models: Organisms/strains		
D. melanogaster (FITS)	This study, deletion mutant of Flam TSS	FITS/FM7A
D. melanogaster (FPSE1)	This study, deletion mutant of Flam Promoter/Shadow Enhancer 1	FPSE1/FM7A
D. melanogaster (FPSE2)	This study, deletion mutants of Flam Promoter/Shadow Enhancer 2 - Allele A & Allele B	FPSE2A/FM7A & FPSE2B/FM7A
D. melanogaster (FPUN)	This study, deletion mutant of Flam Promoter Upstream Negative control	FPUN/FM7A
D. melanogaster (FPRU)	This study, deletion mutant of Flam Promoter Region Unknown control	FPRU/FM7A
D. melanogaster (Flam KG)	Flam-KG mutant; y[1] P{y[+mDint2] w[BR.E.BR] = SUPor-P}lncRNA: flam[KG00476]/FM4	BDSC#: 16453
D. melanogaster (Flam BG)	Flam-BG mutant; w[1118] P{w[+mGT] = GT1}lncRNA: flam[BG02658]	BDSC#: 13912

REAGENT or RESOURCE	SOURCE	IDENTIFIER
D. melanogaster (TJ-GAL4 driver with GAL80TS)	Made in Lau lab during this study, +/+; Tj-GAL4,Gypsy-LacZ/CyO; P-Tub-GAL80TS/TM6-tb	Lau Lab line #113
D. melanogaster (FC0-GAL4 driver)	From Jianjun Sun Lab, CG13083.89K-GAL4[M4]/Cyo, Dfd-eYFP (FC0)	UCONN 7–92 CG13083.89K-Gal4[M4]/Cyo, Dfd-eYFP
D. melanogaster (RNAi-Tj#1)	y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8] = TRiP.HMS01069}attP2	BDSC#: 34595
D. melanogaster (RNAi-Tj#2)	From Godt lab: w; tjDf1/CyO; UAS-tjRNAi on chr 3	NIG-Fly Stock Center, Japan #10034R-2; Gunawan et al.[27]
D. melanogaster (RNAi-Tj#3)	y[1] v[1]; P{y[+t7.7] v[+t1.8] = TRiP.JF02009}attP2	BDSC#: 25987
D. melanogaster (RNAi-Piwi#1)	(UAS-dsRNAi of piwi, Chr3L) |y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8] = TRiP.HMS00606}attP2	BDSC#: 33724
D. melanogaster (RNAi-White)	y[1]v[1]; P{y[+t7.7]v[+t1.8] = TRiP.HMS00004}attP2/TM3, Sb[1]	BDSC#: 33613
D. melanogaster (RNAi-mCherry)	Valium20-mCherry Ctrl shRNA	BDSC#: 35785
Experimental models: Cell lines		
Ovary Somatic Sheet cells (OSS, *D. melanogaster)*	Lau lab, Post et al.[17]	From Post et al.[17]
S2 cells *(D. melanogaster)*	Gift from Michael Marr Lab.	From Khodor et al.[48]
Recombinant DNA		
pMALF	Firefly luciferase reporter plasmid with *Drosophila* mini-actin promoter, Post et al.[17]	From Post et al.[17]
pNN	This study, Nano-Luc reporter backbone with AmpR marker, multiple cloning site for cloning Flam promoter elements.	
pNN-i28	Nano-Luc reporter with Flam promoter element made with oligos FD+426 + 565F & FD+426 + 565R	See Figure 1 and Table S3
pNN-i29	Nano-Luc reporter with Flam promoter element made with oligos FD+350 + 500F & FD+350 + 500R	See Figure 1 and Table S3
All other pNN-i# and pNN-#Flam promoter Nano-Luc reporters	Various Nano-Luc reporter with Flam promoter element made with oligos detailed in Figures 1 and S1.	See Figures 1 and S1 and Table S3
pNN-gypsy5 LTR Fwd	This study, pNN with gypsy5 wild-type LTR promoter.	
pNN-gypsy5 Mut Fwd	This study, pNN with gypsy5 mutated LTR promoter.	
pNN-mdg1 LTR Fwd	This study, pNN with mdg1 wild-type LTR promoter.	
pNN-mdg1 Mut Fwd	This study, pNN with mdg1 mutated LTR promoter.	
Oligonucleotides		
See Table S3 of this study.	Sourced from Eton Biosciences, Genewiz, and Integrated DNA Technologies.	Table S3
Software and algorithms		
Samtools v1.12	Danecek et al.[49]	https://www.htslib.org/
FastQC v0.11.9	Andrews, S.	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

REAGENT or RESOURCE	SOURCE	IDENTIFIER
IGV v2.16.2	Robinson et al.[50]	https://igv.org/doc/desktop/
Bowtie2 v2.5.1	Langmead and Salzberg.[51]	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
STAR v2.7.10b	Dobin et al.[52]	https://github.com/alexdobin/STAR/
BWA v0.7.17	Li and Durbin.[53]	http://bio-bwa.sourceforge.net
R ≥ v4.0.2	https://www.r-project.org/	https://www.r-project.org/
Deeptools v3.5.1	Ramirez et al.[54]	https://github.com/deeptools/deepTools
Bedtools v2.31.0	Quinlan and Hall.[55]	http://bedtools.readthedocs.io
tidyverse v2.0.0	Wickham et al.[56]	https://tidyverse.tidyverse.org/
DESeq2 v1.38.3	Love et al.[57]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
MACS2 v2.2.7.1	Zhang et al.[58]	https://github.com/macs3-project/MACS
Trimmomatic v0.39	Bolger et al.[59]	https://github.com/usadellab/Trimmomatic
Sambamba v0.6.7	Tarasov et al.[60]	https://lomereiter.github.io/sambamba/
IDR v2.0.3	Li et al.[61]	https://github.com/nboley/idr
MEME Suite v5.3.3	Bailey et al.[33]	https://meme-suite.org/meme/
ChIPseeker v1.34.1	Yu et al.[62]	https://github.com/YuLab-SMU/ChIPseeker
DAVID 2021 (Dec. 2021)	Sherman et al.[63]	https://davidbioinformatics.nih.gov/home.jsp
rrvgo v1.2.0	Sayols.[64]	https://github.com/ssayols/rrvgo/
eulerr v7.0.2	Larsson.	https://github.com/jolars/eulerr/
FIMO	Grant et al.[65]	https://meme-suite.org/meme/doc/fimo.html
MSRG pipeline code https://github.com/laulabbumc/MosquitoSmallRNA		