The transcription factor Traffic jam orchestrates the somatic piRNA pathway in *Drosophila* ovaries

Azad Alizada[2,*], Aline Martins[1,*], Nolwenn Mouniée[1,*], Julia V. Rodriguez Suarez[2,*], Benjamin Bertin[1], Nathalie Gueguen[1], Vincent Mirouse[1], Anna-Maria Papameletiou[2], Austin J. Rivera[3], Nelson C. Lau[3], Abdou Akkouche[1], Stéphanie Maupetit-Mehouas[1], Gregory J. Hannon[2,#], Benjamin Czech Nicholson[2,#], Emilie Brasset[1,#]

[*] These authors contributed equally to this work.

[#] Corresponding authors:
GJH (greg.hannon@cruk.cam.ac.uk)
BCN (ben.nicholson@cruk.cam.ac.uk)
EB (emilie.brasset@uca.fr)

Affiliations:

[1] iGReD, Université Clermont Auvergne, CNRS, INSERM, Faculté de Médecine, 63000 Clermont-Ferrand, France

[2] Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Cambridge CB2 0RE, UK

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

Keywords
piRNA pathway, *Drosophila*, oogenesis, transposable elements, transcription factors

Highlights

- Traffic jam (Tj) acts as a master regulator of the somatic piRNA pathway in *Drosophila*.
- Tj regulates a network of piRNA pathway genes, mirroring the gene-regulatory mechanism of A-MYB in the mouse testis and Ovo in fly ovaries.
- *Cis*-regulatory elements with Tj motifs are present at the promoters of somatic piRNA pathway genes.
- The expression of the *flamenco* piRNA cluster is directly controlled by Tj.

Abstract

The PIWI-interacting RNA (piRNA) pathway is essential for transposable element (TE) silencing in animal gonads. While the transcriptional regulation of piRNA pathway components in germ cells has been documented in mice and flies, their control in somatic cells of *Drosophila* ovaries remains unresolved. Here, we demonstrate that Traffic jam (Tj), the *Drosophila* orthologue of large Maf transcription factors in mammals, is a master regulator of the somatic piRNA pathway. Tj binds to regulatory regions of somatic piRNA factors and the major piRNA cluster *flamenco*, which carries a Tj-bound enhancer downstream of its promoter. Depletion of Tj in somatic follicle cells causes downregulation of piRNA factors, loss of *flam* expression and de-repression of *gypsy*-family TEs. We propose that the arms race between the host and TEs led to the co-evolution of promoters in piRNA pathway genes as well as TE regulatory regions that both rely on a shared transcription factor.

Introduction

Transposable elements are DNA sequences with the ability to move or replicate to new positions within the host genome[1]. Succeeding waves of TE mobilization and repression have allowed TEs to accumulate in the genomes of nearly all organisms. Approximately 20% of the *Drosophila* genome and more than 50% of the human genome is composed of TE sequences[2]. While their ability to transpose makes them powerful facilitators of genome evolution, their mobilization can also have deleterious effects, such as mutations, gene disruptions, and chromosomal rearrangements[3,4]. The resulting selective pressure has driven the co-evolution of many mechanisms within the hosts to effectively counter the disruptive TE activity[5].

In animals, the piRNA pathway is one of the key defence mechanisms against active TEs. In *Drosophila*, the piRNA pathway is crucial for maintaining genome stability, particularly in gonadal tissues, which consist of germ cells surrounded by somatic follicle cells[6–8]. Previous work has uncovered the existence of distinct piRNA pathways in germ cells and somatic follicle cells[7–9]. Germ cells express a more sophisticated piRNA system that involves both post-transcriptional silencing of transposon transcripts via the ping-pong loop and co-transcriptional silencing of nascent TEs. The first mechanism requires Aubergine, Argonaute-3 and numerous Tudor and RNA binding proteins such as Qin, Tejas, Vasa, Krimper, Spn-E and BoYb, whereas co-transcriptional TE silencing relies on Piwi, its cofactor Arx and the PICTS complex composed of Panx, Nxf2, Nxt1, and Ctp[10–16]. Furthermore, piRNAs in germ cells are predominantly generated from precursor transcripts derived from discrete genomic loci called dual-strand piRNA clusters[6,17–19]. The target repertoires of these piRNA clusters are dictated by both full-length and truncated transposon copies and their transcription and export relies on non-canonical mechanisms, involving Rhino (Rhi), Deadlock (Del), Cutoff (Cuff), Moonshiner (Moon), Nxf3, and Bootlegger (Boot)[6].

To avoid the robust silencing machinery in germ cells, some TEs such as *gypsy* or ZAM have gained the ability to express exclusively in the somatic follicle cells, form virus-like

particles and invade germ cells such as the oocyte[20–22]. This infectious capacity enables their transposition into the germline genome and hence transmission of new TE copies to future generations. As a consequence of this host-transposon arms race, somatic follicle cells have adapted a simplified piRNA pathway version comprised of the factors required for piRNA biogenesis (e.g., Armi, Zuc, Mino, Gasz, Daed, SoYb, Vret, Shu and the soma-specific Yb) and co-transcriptional transposon silencing (e.g., Piwi, Arx, Panx, Nxf2, Nxt1, and Ctp)[8,13,14,23–27]. In addition, somatic piRNA clusters have emerged to control TE expression in the somatic follicle cells of the ovary. In contrast to the germline-active dual-strand piRNA clusters, the soma-expressed piRNA clusters are unistrand and give rise to piRNAs capable of targetting active TEs. The major somatic piRNA cluster in *Drosophila* ovaries is *flamenco* (*flam*)[6,28]. Contrary to the dual-strand piRNA clusters, *flam* is conserved across millions of years of Drosophilid evolution[29], and transcribed by a canonical mechanism that includes a promoter, splicing, and polyadenylation[30]. Furthermore, to-date, *flam* is the only known essential piRNA cluster, with *flam* mutants showing rudimentary ovaries and complete female sterility[18,31].

Despite the importance of the somatic piRNA pathway, including the *flam* cluster and the somatic piRNA factors, the gene-regulatory circuit controlling their transcription remains to be determined. A recent study has shown that the transcription factor (TF) Ovo plays a key role in coordinating the expression of germline-specific piRNA pathway genes in the ovaries[32]. However, it remains unclear whether one or multiple TF(s) drive the somatic piRNA pathway in the ovarian follicle cells. Here, we uncover Traffic jam (Tj), a large Maf TF, as a master regulator of the somatic piRNA pathway. Tj is critical for gonad development[33] and was previously implicated in the regulation of *piwi* expression in somatic cells[34]. In this study, we extend these findings by showing that Tj regulates the expression of many somatic piRNA pathway factors, as well as the key unistrand piRNA cluster, *flamenco*. We demonstrate that Tj directly binds to the regulatory regions of these genes, driving their soma-specific expression in ovarian follicle cells. Furthermore, we identify a specific enhancer region downstream of the *flam* TSS that is essential for its transcriptional activity in somatic cells and harbors Tj binding sites. Our findings shed light

on the intricate transcriptional control necessary for TE silencing in the ovarian soma, offering new insights into the regulation of the soma-specific piRNA pathway.

Results

The transcription factor Traffic jam controls somatic piRNA factor expression

To uncover the regulatory mechanisms underlying the expression of somatic piRNA pathway factors, we first performed differential gene expression analysis between germline and somatic cells derived from *Drosophila* ovaries. We dissected and dissociated ovaries of transgenic *D. melanogaster* flies expressing GFP specifically in germ cells (using the *vasa* promoter) and sorted the cells using FACS into germline (*vas*-GFP+) and somatic (*vas*-GFP-) populations[32]. We then extracted RNAs and generated RNA-seq libraries before carrying out a differential gene expression analysis between germline and somatic cells. This approach identified piRNA pathway genes with significant expression in the soma of the *Drosophila* ovary, namely*, fs(1)Yb*, *nxf2*, *panx*, and *armi* (**Figure 1A, left**). Among these factors, *fs(1)Yb* was reported to be exclusively expressed in somatic cells[35], while the other factors are known to also function in germ cells[36]**.** As expected these four soma-enriched piRNA factors also showed expression in ovarian somatic cells (OSCs) (**Figure 1A**), a cell line derived from ovary somatic follicle cells[34,37].

Next, to identify the potential TFs controlling the expression of these somatic-enriched piRNA factors (*fs(1)Yb, nxf2, panx*, and *armi*), we analysed ATAC-seq from wildtype OSCs[38] and performed motif scanning using the FIMO tool (match $p<0.001$; in MEME suite) to look for transcription factor motifs (the FlyFactorSurvey database) that were shared between their promoter ATAC-seq peaks (±500 bp of the transcription start sites) (**Figure 1B and Table S1**). Surprisingly, this analysis identified only a single transcription factor motif shared between the promoter ATAC-seq peaks of the four soma-enriched piRNA factors, Tj. We confirmed *in vivo* Tj binding to the promoters of *fs(1)Yb*, *nxf2*, *panx*, and *armi* (**Figure 1C**) by re-analysing publicly available Tj-GFP ChIP-seq (referred to as Tj ChIP-seq) data from whole female flies[39]. Moreover, Tj ChIP-seq peaks and Tj motifs were also present within promoters of other somatic piRNA factors that are expressed in both somatic and germline cells such as *piwi* but these were absent in promoters of the

germline-specific piRNA factors (**Figure 1A, right and Figure 1C**), suggesting that Tj could be specifically controlling somatic piRNA pathway genes.

To validate our hypothesis that Tj regulates soma-expressed piRNA factors, we performed *tj* knockdowns in OSCs, followed by RT-qPCR. Tj depletion resulted in downregulation of somatic piRNA factors (**Figure 1D**). Although *tj* levels were reduced by ~40% following 48 hours of knockdown, we observed similar levels of downregulation for *nxf2*, *piwi*, *armi*, and *fs(1)yb* (RT-qPCR, n=3 replicates from distinct samples, $p<0.01$) (**Figure 1D**). To corroborate these results, we performed RNA-seq following 48 hours and 96 hours of *tj* knockdown. These also showed significant downregulation of the soma-enriched piRNA factors *fs(1)Yb*, *nxf2*, *panx*, *soYb*, and *armi* (RNA-seq, n≥3 replicates from distinct samples, p.adj.<0.01) (**Figure 1E and Figure S1A**). The piRNA factors shared between the somatic and germline cells which showed Tj binding at their promoters *in vivo* (i.e., *piwi*, *vret*, *mino*, and *shu*), were similarly downregulated upon *tj* knockdowns in OSCs (**Figures 1E and S1A**). These results suggest that Tj is a soma-enriched TF responsible for the regulation of soma-expressed piRNA pathway genes in *Drosophila* ovarian somatic cells, likely via direct promoter binding.

Genome-wide Tj ChIP-seq analysis revealed 7,956 Tj binding sites overall, which were enriched at the promoters of soma-expressed genes (including the somatic piRNA pathway genes), and depleted from promoters of the germline-enriched genes, including germline-specific piRNA pathway genes (**Figures 1A and 1F**). Specifically, Tj peaks overlapped with 714 transcription start sites (within ±500 bp of TSS) of 419 soma-expressed genes (defined using RNA-seq $\log_2 FC<-1$, adjusted p.value<0.01, n=3 from FACS-sorted *vas*-GFP- cells), representing 62% of all soma-enriched gene TSSs (n=1,142) (**Table S1**). Motif analysis revealed that the top-scoring *de novo* motif within Tj ChIP-seq peaks had a consensus sequence of TGCTGAC (STREME e-value=2.0×10^{-08}). We noticed that the top-enriched motif within the peak centers had a consensus sequence of TGCTGA (CentriMo e-value=1.1×10^{-20}) and was often found nearby its reverse complement TCAGCA (as close as by 1 bp spacing), resulting in a pseudo-palindromic TGCTGA$(N)_n$TCAGCA motif (CentriMo e-value=1.3×10^{-16}; FIMO n=2,113 sites) (**Figures**

1G and S1B). Interestingly, this closely mirrors the Maf recognition element (MARE) motifs recognized by DNA-binding domains of the dimeric basic Leucine Zipper Maf family transcription factors[40,41], suggesting that Tj could bind some target sites as a dimer. Overall, our results indicate that Tj specifically directs transcriptional regulation of the somatic ovarian program, including somatic piRNA pathway components, by binding to TF motifs at the promoters of somatic target genes.

To test Tj's role in regulating somatic piRNA factors in an *in vivo* setting, we disrupted Tj activity and examined piRNA factors expression through immunostaining in the *Drosophila* adult ovaries (**Figure 1H**). As Tj plays a critical role in regulating the formation and function of the germline stem cell niche and its mutation leads to ovarian atrophy[42,43], we abolished Tj activity in mosaic clones of follicle cells carrying a genetic null allele of Tj, the tj^{eo2} allele[33]. This allele carries a point mutation resulting in a premature stop codon resulting in a truncated version of the Tj protein, which lacks critical DNA-binding domains, including the leucine zipper and nuclear localization signal[33]. As expected, in wild-type follicle cells, Armi, protein essential for piRNA biogenesis, accumulate in distinct foci representing Yb-bodies near the nuclear periphery[44]. However, in *tj* mutant clones, Armi proteins were completely absent (**Figure 1H**). This finding confirms that, as observed in OSCs, Tj regulates piRNA factor expression within follicle cells of the adult ovary.

Tj binds to conserved motifs in the regulatory region of the piRNA cluster *flam*

We next asked whether Tj also regulates somatic piRNA cluster transcription. The major piRNA cluster active in OSCs and somatic follicle cells is *flamenco*[6,28]. Analysis of ovary and OSC ATAC-seq data[29,38] revealed peaks upstream of *flam* that correspond to the nearby gene DIP1 (ATAC-seq peaks #1-4) as well as three conserved peaks 1 kb upstream and 5 kb downstream of the *flam* TSS in *D. melanogaster* (ATAC-seq peaks #5, #6 and #7) that were also present in the orthologous region in the *D. yakuba* genome (**Figure 2A**). Peak #5 harbored the motif of cubitus interruptus, a TF previously reported to contriunte to *flam* expression[45], while peak #6 contained the initiator element (Inr) of

the *flam* TSS[45]. Peak #8 overlapped an alternative, further downstream TSS that was previously identified by PRO-seq (the second *flam* PRO-seq transcription initiation peak[46]. We noted that the conserved peaks #5 and #7 had disproportionately stronger ATAC-seq signals in OSCs compared to the whole ovaries, suggesting their somatic enrichment (**Figure 2A**). To identify enhancer activity genome-wide, we re-analysed Self-Transcribing Active Regulatory Region sequencing (STARR-seq) data performed in OSCs using genomic DNA of both *D. melanogaster* and *D. yakuba*[47–49]. We found that enhancer activity signal overlapped with ATAC-seq peak #7 in *D. melanogaster* and its corresponding peak in the *D. yakuba* (**Figure 2A**). Motif scanning (FIMO) revealed 21 transcription factor motifs (the FlyFactorSurvey database) that were shared between peak #7 in *D. melanogaster* and its orthologous region in *D. yakuba* **(Figure 2B)**. Intriguingly, the motif for Tj was among the shared hits. Expression filtering using RNA-seq data from ovaries of *D. melanogaster* and *D.yakuba* revealed TF candidates expressed in this tissue in both species **(Figure 2B)**. Of these, four (*tj*, *crc*, *br*, and *vri*) showed soma-enrichment with *tj* being the top hit **(Figure 2C)**.

In the *D. melanogaster* genome*,* Tj motifs were found downstream of the *flam* TSS at positions +327 to +332 bp (chrX:21,632,218-21,632,223; dm6), +437 to +450 bp (mutated pseudo-palindrome; chrX:21,632,328-21,632,340; dm6), +517 to +523 bp (summit of Tj ChIP-seq peak; chrX:21,632,408-21,632,414; dm6), and +935 to +940 bp (upstream of the alternative *flam* TSS; chrX: 21,632,826-21,632,831; dm6) (**Figure 2A**). Through Tj ChIP-seq analysis, we confirmed a strong Tj binding peak (-336 bp to +1214 bp) ovarlapping these motifs. Interestingly, two of the Tj motifs (+437 to +450 bp and +517 to +523 bp) also overlapped the conserved ovary ATAC-seq peak #7 (**Figure 2A**). The summits of the Tj ChIP-seq peak, the ovary ATAC-seq peak and the OSC ATAC-seq peak coincided with the +517 to +523 bp Tj motif (**Figure 2A**). Taken together, our results suggest that, in addition to regulation of somatic piRNA genes, Tj may also control the piRNA cluster *flam*.

Expression of *flam* in follicle cells is dependent on a regulatory region located downstream of its TSS

Previous work identified the major TSS, and the minimal promoter of *flam* in OSS cells[45]. To confirm the *flam* regulatory regions, including putative Tj motifs, *in vivo*, we generated transgenic flies carrying transgene expressing the *tomato* reporter gene under the control of different *flam* promoter regions (starting at -1626 bp upstream and ending at +6334 bp downstream of the *flam* TSS) (**Figure 3A**). To compare the relative expression of these reporter constructs, we integrated all reporter variants into the identical genomic site at cytological position 53B2 and followed Tomato expression in fly ovaries by immunostaining and RT-qPCR (**Figures 3A-C**). The strongest Tomato expression in somatic follicle cells was observed for the -515; +1200 and the -515; +2086 constructs, with weaker signals found with the -515; +718 construct (**Figures 3B-C**). The increased Tomato expression observed in the -515; +1200 line prompted us to further explore this region. The presence of the fourth Tj motif identified at +935 to +940 bp (TCAGCA) together with the alternative *flam* TSS immediately downstream of this region at +977 to +1202 (ovary ATAC-seq peak #8 matching the second *flam* PRO-seq transcription initiation peak) potentially contributed to the enhanced Tomato expression observed when the region in constructs extending beyond +718.

Surprisingly almost no artificial tomato piRNAs was detected with the -515; +432 and -515; +718 constructs (**Figure 3D**). These results suggest that the first 718 bp of the *flam* transcript, which were previously described as the piRNA trigger sequence in OSS cells[50], are not sufficient to direct artificial piRNA processing from the *tomato* reporter *in vivo* (**Figure 3D**). In contrast, a longer -515; +6334 construct showed a weaker Tomato expression but exhibited a markedly stronger piRNA production relative to the -515; +1200 and -515; +2086 reporters (**Figures 3B-D**). Hence, including more *flam* sequence appeared to enhance piRNA processing efficiency and resulted in a concordant decrease of reporter mRNAs. This observation could be similar to the size-dependent dynamics observed in pachytene piRNA clusters in mice[51]. As expected, no Tomato expression was detected in the germline regardless of the constructs used (**Figure 3B**).

Neither the transgene carrying the -515; +30 *flam* regulatory region (which includes the minimal promoter), nor the reporter comprising -515; +432 (which includes the first exon) was sufficient to drive strong Tomato expression in somatic follicle cells (**Figure 3B-C**). These results suggest that while the Ci binding site (-394 to -366 bp from the *flam* TSS) is required for minimal *flam* expression in OSS cells, it is unable to induce strong expression in adult follicle cells. To confirm that the minimal *flam* promoter we have previously identified can efficiently drive transcription *in vivo,* we added UAS sequences in front of the -515; +432 *flam* promoter (**Figure 4A**). When combined with a follicle cell-specific *traffic jam* (*tj*)-Gal4 driver, we indeed observed strong Tomato expression in somatic follicle cells, (**Figures 4B-C**). Next, using CRISPR/Cas9, we deleted the minimal promoter region and tested its requirement for *flam* expression (**Figures 4A,D-F**). The resulting *flam* mutants, named *flam*Δ*prom*, were female sterile and exhibited fully atrophied ovaries, along with a loss of *flam* piRNAs, while the expression of the germline piRNA clusters remained unaffected (**Figure 4D-F**). These results suggest that the -515 and +432 region (relative to the *flam* TSS) contains the minimal promoter necessary for *flam* expression but not its full regulatory region. Furthermore, no Tomato expression was detected in ovaries of transgenic flies carrying the +770; +2086 *flam* region (**Figures S2A and S2B**). Overall, our findings suggest that the sequence between +432 and +1200 bp downstream of the *flam* TSS that overlaps Tj ChIP-seq peaks and contains Tj motifs, functions as an enhancer driving efficient soma-specific *flam* expression in the adult ovary.

We also investigated whether the sequence upstream of the *flam* TSS influences its expression by testing a larger *flam* promoter that includes the -1624 bp region that ranges from the *flam* TSS to the closest upstream gene, *DIP1* (**Figure 3A**). This extended promoter did not enhance Tomato expression in the somatic follicle cells of the -1624; +718 and -1624; +2086 transgenic lines (**Figures 3A-C**). These findings suggest the absence of additional enhancers upstream of the minimal *flam* promoter.

Tj is the master transcription factor for follicle cell-specific expression of *flam*

To test for a direct regulation of the *flam* piRNA cluster by Tj *in vivo*, we abolished Tj activity and monitored *flam* expression by smFISH in mosaic clones of follicle cells (**Figures 5A and S3**). We observed a complete absence of *flam* signal in *tj*-deficient mutant cells, whereas *flam* signal at the nuclear periphery was detected in neighboring control cells that were Tj positive, consistent with previous reports[52,53]. Furthermore, sequencing of small RNAs from ovaries with soma-specific knockdown (SKD) of Tj (using a *tj*-Gal4 driver), showed a complete loss of piRNAs from the *flam* locus. (**Figure 5B**). Since Tj has been previously reported to regulate Piwi [34] and to rule out indirect effects, we investigated *flam* transcription by smFISH in follicle cells upon SKD of *piwi*. In contrast to *tj* mutant follicle cells, which lack *flam* expression, the absence of Piwi had no effect on *flam* levels (**Figures 5C-D**). In conclusion, our findings strongly support a key role of Tj in directly regulating the transcription of *flam* piRNA precursors in the somatic follicle cells of the adult ovary.

To investigate whether additional somatic transcription factors could be involved in the follicle cell-specific expression of *flam in vivo*, we carried out a Translating Ribosome Affinity Purification (TRAP) experiment. This method utilizes the UAS/Gal4 system and allows the identification of cell-specific messenger RNA (mRNA) translation[54]. In adult *Drosophila* ovaries, we applied TRAP to the GFP-labeled follicle cell population using the *tj*-Gal4 driver (UAS-GFP::Rpl10AGFP; *tj*-Gal4) and to the germ cell population with *nos*-Gal4 driver (UAS-GFP::Rpl10AGFP; *nos*-Gal4). Transcriptional profiling of purified mRNA followed by bioinformatic analyses identified around 900 genes with significantly higher expression and translation in somatic cells compared to germline cells (FC>2, p-value <0.05). In support of this, genes enriched in Tj positive cells were associated with specific gene ontology biological processes related to the development of follicle somatic cells (**Figure S4**). Among the differentially enriched genes, 30 were DNA binding factors and this list included Tj (**Table S2**). Using *tj*-Gal4 mediated somatic knockdowns (SKD), we next performed an RNAi screen targeting each candidate, and retained only those that reproduced the *flam* mutant phenotype that is characterized by atrophied ovaries[31,55]. We found that depletion of *tj*, *Stat92E*, *schnurri*, *tango*, *chd64* and *mamo* caused atrophic ovary development similar to *flam*. We next checked whether these factors affect *flam*

expression by smFISH (**Figure S5**). Because SKD of *Stat92E*, *tango*, *chd64*, *mamo* and *schnurri* resulted in strongly atrophied ovaries, we instead transiently depleted these factors in adult follicle cells using a combination of *tj*-Gal4 the thermo-sensitive Tub-Gal80ts inhibitor[56]. A temperature shift from 18°C to 28°C enabled RNAi induction in late oogenesis stages, after ovaries had developed. In contrast to the atrophied ovary phenotype we observed with SKD, no impact on *flam* expression was detected following the depletion of *schnurri*, *tango*, *chd64*, and *mamo* using the thermo-sensitive RNAi approach (**Figure S5**). These results suggest that, while we cannot entirely rule out the contribution of multiple transcription factors, Tj appears to be the master TF responsible for soma-specific expression of *flam*.

Selective derepression of *gypsy*-family TEs following Tj depletion

Since *flam* and the somatic piRNA pathway factors control *gypsy*-family TEs in OSCs, we next asked whether Tj depletion resulted in transposon derepression. RT-qPCR revealed ~4-fold and ~20-fold upregulation of *gypsy* expression following both 48 h and 96 h of *tj* siRNA treatment in OSCs (**Figures S6A and S6B**). RNA-seq from OSCs further confirmed robust upregulation of *gypsy* and other *gypsy*-family TEs after 48 h and 96 h of *tj* knockdown (**Figures 6A and S6C**). RT-qPCR from ovaries with *tj*-SKD showed upregulation of *gypsy* and *ZAM* (**Figure S6D**). Upregulation of *gypsy* itself was validated using a established reporter (*tj*-Gal4; *gypsy*-LacZ) (**Figure 6B**). Surprisingly, although several *gypsy*-family TEs were derepressed in *flam* mutant ovaries (**Figure 6C**), only a subset were upregulated in ovaries or OSCs depleted of Tj. The TEs *mdg1* and *412*, which are upregulated in *flam* mutants ovaries (**Figure 6C**), remained repressed upon *tj* knockdown in OSCs or ovaries (**Figures 6A and S6C**).

As *tj*-SKD ovaries were strongly atrophied, we validated our results in *tj*-SKD-induced mosaic clones. Consistent with our RT-qPCR, smFISH detected de-repression of *gypsy* and *ZAM*, but not *mdg1* or *412*, in *tj*-SKD-induced mosaic clones, compared to neighboring control cells (**Figures 6D and S6E**). We hypothesized that the differences in TE expression might be caused by some TEs also being directly regulated by Tj. Indeed,

most of TEs deregulated in *flamΔprom* flies, such as *412* and *mdg1* showed strong Tj ChIP-seq peaks, particularly in the LTR region (**Figures 6E-F)**, which was distinct from H3K9me3 signal for the same TE. Co-opted regulation by Tj could potentially explain why certain TEs are not de-repressed following *tj* knockdown. In contrast, TEs such as *ZAM*, and *297,* which show little to no Tj binding in their promoter regions, are upregulated under both *flamΔprom* and *tj* knockdown conditions, consistent with their transcription being independent of the direct Tj binding. In addition, we noted that some TEs, such as *gypsy* display Tj binding yet remain de-repressed upon *tj*-SKD suggesting that other somatic transcription factors could contribute to their expression or that residual Tj protein could be sufficient for their activation. Of note, this is supported by the accompanying study by Rivera and colleagues which also showed that Tj binds and recognizes motifs within the LTRs of several TEs[57]. Overall, these results suggest co-evolution between Tj-dependent regulatory regions of TEs and piRNA pathway genes and clusters.

Discussion

Our study identifies Tj as a pivotal transcriptional regulator of the somatic piRNA pathway in *Drosophila* ovaries and adds a key layer to our understanding of TE repression mechanisms in *Drosophila*, particularly in non-germline tissues. Previous studies have demonstrated Tj's involvement in gonadal development and its regulation of *piwi* expression[33,34]. Our findings expand on these roles by demonstrating Tj's broader control over somatic piRNA pathway components in ovarian follicle cells. Our results reveal Tj as a regulator of somatic piRNA pathway components, including *fs(1)Yb*, *nxf2*, *panx*, and *armi*. ChIP-seq and motif analyses demonstrate that Tj binding at conserved promoter motifs is critical for the expression of these genes. Moreover, Tj regulates the soma-expressed piRNA cluster *flam*, which is essential for TE silencing. This dual regulation of piRNA factors and clusters mirrors the function of mammalian transcription factors such as A-MYB, TCFL5 and BTBD18 which coordinate piRNA pathway gene expression and pachytene piRNA cluster activation during spermatogenesis[58,59].

Beyond its role in piRNA regulation, Tj appears to be a master regulator of somatic cell identity with approximately 8,000 Tj binding sites detected at the promoters of somatically enriched genes. RNA-seq analyses of *tj* knockdowns in OSCs revealed a significant shift from somatic to germline gene expression, with upregulation of germline-specific genes such as *osk*, *ovo*, and *nos* (**Table S1**). This shift is accompanied by the downregulation of *l(3)mbt*, suggesting that Tj maintains the somatic expression program by repressing germline gene transcription. Interestingly, Tj and Ovo, a germline-specific transcription factor[38], exhibit opposing binding patterns at the promoters of somatic and germline genes (**Figure S7A**), underscoring their complementary roles in maintaining lineage-specific gene expression programmes within gonads (**Figure S7B**). However, while Ovo's role in regulating germline piRNA factors is conserved across species, it is not clear whether Tj's regulation of the somatic piRNA pathway is a *Drosophila*-specific adaptation or more broadly conserved. We speculate that this evolutionary innovation likely arose in response to the challenge posed by *gypsy*-family TEs, which can infect the germline via somatic expression[60,61]. These TEs have evolved distinct expression patterns in somatic

follicle cells[60,61]. When the piRNA pathway is disrupted, some TEs are broadly expressed, while others remain restricted to specific niches, *ZAM* in posterior follicle cells and *Idefix* in the germarium for example[60–62], indicating regulation by distinct transcription factors. Indeed, some TEs were reported to rely on specialized transcription factors, such as the Pointed for *ZAM*[63], while our results indicate that other TEs are likely co-regulated by Tj, allowing for broader or more robust somatic expression. Notably, only a subset of TEs derepressed in *flam* mutants are also affected by Tj loss, suggesting that some TEs are independent of Tj for their expression. Co-evolution between TEs and the piRNA pathway could suggest an ongoing evolutionary "arms race" between host regulatory mechanisms and TEs. The coordinated expression of *flam* and somatic piRNA pathway components likely reflects the selective pressures exerted by TEs that shaped their *cis*-regulatory sequences to respond to one key transcription factor in ovarian somatic cells allowing for efficient and coordinated silencing of both niche-specific and broadly expressed TEs in somatic cells.

Although primarily considered a transcriptional activator, Tj may also function as a repressor - either at the transcriptional level by silencing adhesion molecule genes or post-transcriptionally through its 3' UTR, which generates piRNAs that target and silence specific genes such as *FasIII*[33,34]. The precise mechanisms by which Tj regulates its targets remain to be fully elucidated, particularly the potential involvement of additional transcriptional regulators and co-factors. Other factors, such as Ci, Mirror or Stat92E, which are all DNA-binding proteins expressed in somatic follicle cells that were identified in our TRAP-based screening, could contribute to the regulation of piRNA genes and *flam*, as is often described in gene regulatory networks where multiple transcription factors work together to ensure robustness and precise control across development[64,65].

Tj's similarity to mammalian Maf proteins, c-Maf and MafB in particular, including its potential for dimerization, adds another layer of specificity to its regulatory function. Like other members of this TF family, Tj contains a leucine-zipper domain, a basic DNA-binding domain, and a Maf-specific extended homology domain in its C-terminal region.

The leucine-zipper domain, known for facilitating the dimerization of Maf proteins, either with other Maf or bZIP TFs (e.g., homo/heterodimeric AP-1 complexes), suggests that Tj could also function as a dimer[33]. The presence of pseudo-palindromic motifs within Tj binding sites, similar to the MARE motifs recognized by the dimeric Maf family proteins, further reinforces this idea[40,41]. Interestingly, the pseudo-palindromic Tj motifs at promoters of the somatic piRNA factors were often mutated (**Figure S1B**), which has been previously shown to confer DNA-binding specificity for homodimeric bZip transcription factor complexes over their heterodimeric versions in EMSA experiments[41].

In conclusion, our findings establish Tj as a master regulator of the somatic piRNA pathway and a key player in maintaining genome integrity through TE repression. By regulating both piRNA pathway components and the *flam* piRNA cluster, Tj exemplifies the efficiency of a single transcription factor in orchestrating complex gene networks. Tj's interaction partners which likely contribute to its finely tuned regulatory expression network, await further exploration.

Limitations of the Study

While our data suggest direct effects of Tj on the somatic piRNA pathway, we cannot entirely exclude the possibility of indirect effects. Tj regulates numerous somatic genes, including transcription factors, and its broad regulatory network may contribute indirectly to piRNA pathway activation. For example, the downregulation of *l(3)mbt*, a repressor of germline programs in somatic cells, observed upon *tj* knockdown could indirectly influence somatic piRNA pathway gene expression. Further investigation will be needed to disentangle these direct and indirect regulatory effects. Tj is known to regulate many genes critical for ovarian development, and therefore the ovarian atrophy observed upon Tj loss could be a consequence of both TE reactivation due to piRNA pathway disruption and/or misregulation of developmental genes critical for ovarian tissue homeostasis. In addition, our results with the *de novo* motif analysis with STREME did not yield the pseudo-palindromic motifs likely due to limitation in the detection of gapped motifs by the software, as well as due to frequent mutations and less frequent appearance of the

pseudo-palindromes at specific gapped lengths (1-4bp) within the peaks; however, we were able to map them using the motif enrichment (i.e., CentriMo in MEME-ChIP) and scanning tools (i.e., FIMO).

Author Contributions

AzA performed experiments and bioinformatic analyses, designed figures, and contributed to manuscript writing. AM conducted data analysis, Tj clonal experiments, functional assays with transcription factor knockdown, prepared figures, and contributed to manuscript writing. NM performed all the cloning, generated *Drosophila* transgenic lines, and conducted Tomato immunostaining and corresponding small RNA sequencing. JRS performed the experiments, analyzed, plotted RT-qPCR data and contributed to manuscript writing. BB designed CRISPR constructs, generated the corresponding *Drosophila* mutant lines, conducted TRAP experiments and bioinformatics analyses. NG was responsible for smFISH probe labeling. VM conducted TRAP experiments and bioinformatics analyses. AMP performed computational analyses. NCL designed the Tj peptide, funded antibody production, and supervised AJR in conducting affinity purification and validating the specificity of the anti-Tj antibody. AbA contributed to the Tj clonal experiments and performed ovary RT-qPCR experiment, analyzed and plotted the data. SMM contributed to the Tj clonal experiments, analysis, and manuscript writing. GJH supervised the project and contributed to funding acquisition. BCN conceptualized the project, designed the experiments, supervised the work, contributed to funding acquisition and wrote the manuscript. EB conceptualized the project, designed the experiments, supervised the work, contributed to funding acquisition and wrote the manuscript.

Acknowledgments

We thank Brasset and Hannon group members for fruitful discussions, and I. Vaillant for comments on the manuscript and S. Jensen for participating in the bioinformatics training of students. We thank the Bloomington *Drosophila* Stock Center for providing fly lines. We thank Dorothea Godt for providing Tj mutant *Drosophila* stocks. We thank Y. Renaud and P. Pouchin for helpful advices and development of the iGReD bioinfomatics platform. We also thank the CLIC facility (Clermont Imagerie Confocale). We thank the Scientific Computing core at the CRUK Cambridge Institute for HPC resources, Research Instrumentation and Cell Services for training, liquid nitrogen storage and mycoplasma testing, and the Genomics core for sequencing services. This work was supported by

grants from the Agence Nationale pour la Recherche (CHApiTRE ANR-20-CE12-0005, EB, BiopiC ANR-21-CE12-0022, EB). AM was supported by the Fondation pour la Recherche Médicale (ECO202006011583 and FDT202304016393, AM). This research was financed by the French government IDEX-ISITE initiative 16-IDEX-0001 (CAP 20-25). NCL is funded by NIH/NIGMS grant R01GM135215. GJH is a Royal Society Wolfson Research Professor (RSRP\R\200001). This research was funded in whole, or in part, by Cancer Research UK (G101107) and the Wellcome Trust (110161/Z/15/Z and 226627/Z/22/Z).

Declaration of interests

The authors declare no competing interests.

References

(1) Bourque, G.; Burns, K. H.; Gehring, M.; Gorbunova, V.; Seluanov, A.; Hammell, M.; Imbeault, M.; Izsvák, Z.; Levin, H. L.; Macfarlan, T. S.; Mager, D. L.; Feschotte, C. Ten Things You Should Know about Transposable Elements. *Genome Biol.* **2018**, *19* (1), 199. https://doi.org/10.1186/s13059-018-1577-z.

(2) Burns, K. H.; Boeke, J. D. Human Transposon Tectonics. *Cell* **2012**, *149* (4), 740–752. https://doi.org/10.1016/j.cell.2012.04.019.

(3) Kazazian, H. H. Mobile Elements: Drivers of Genome Evolution. *Science* **2004**, *303* (5664), 1626–1632. https://doi.org/10.1126/science.1089670.

(4) Chuong, E. B.; Elde, N. C.; Feschotte, C. Regulatory Activities of Transposable Elements: From Conflicts to Benefits. *Nat. Rev. Genet.* **2017**, *18* (2), 71–86. https://doi.org/10.1038/nrg.2016.139.

(5) Cosby, R. L.; Chang, N.-C.; Feschotte, C. Host-Transposon Interactions: Conflict, Cooperation, and Cooption. *Genes Dev.* **2019**, *33* (17–18), 1098–1116. https://doi.org/10.1101/gad.327312.119.

(6) Brennecke, J.; Aravin, A. A.; Stark, A.; Dus, M.; Kellis, M.; Sachidanandam, R.; Hannon, G. J. Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila. *Cell* **2007**, *128* (6), 1089–1103. https://doi.org/10.1016/j.cell.2007.01.043.

(7) Malone, C. D.; Brennecke, J.; Dus, M.; Stark, A.; McCombie, W. R.; Sachidanandam, R.; Hannon, G. J. Specialized piRNA Pathways Act in Germline and Somatic Tissues of the Drosophila Ovary. *Cell* **2009**, *137* (3), 522–535. https://doi.org/10.1016/j.cell.2009.03.040.

(8) Théron, E.; Dennis, C.; Brasset, E.; Vaury, C. Distinct Features of the piRNA Pathway in Somatic and Germ Cells: From piRNA Cluster Transcription to piRNA Processing and Amplification. *Mob. DNA* **2014**, *5* (1), 1–11. https://doi.org/10.1186/s13100-014-0028-y.

(9) Malone, C. D.; Hannon, G. J. Small RNAs as Guardians of the Genome. *Cell* **2009**, *136* (4), 656–668. https://doi.org/10.1016/j.cell.2009.01.045.

(10) Sienski, G.; Batki, J.; Senti, K.-A.; Dönertas, D.; Tirian, L.; Meixner, K.; Brennecke, J. Silencio/CG9754 Connects the Piwi–piRNA Complex to the Cellular Heterochromatin Machinery. *Genes Dev.* **2015**, *29* (21), 2258–2271. https://doi.org/10.1101/gad.271908.115.

(11) Yu, Y.; Gu, J.; Jin, Y.; Luo, Y.; Preall, J. B.; Ma, J.; Czech, B.; Hannon, G. J. Panoramix Enforces piRNA-Dependent Cotranscriptional Silencing. *Science* **2015**, *350* (6258), 339–342. https://doi.org/10.1126/science.aab0700.

(12) Fabry, M. H.; Ciabrelli, F.; Munafò, M.; Eastwood, E. L.; Kneuss, E.; Falciatori, I.; Falconio, F. A.; Hannon, G. J.; Czech, B. piRNA-Guided Co-Transcriptional Silencing Coopts Nuclear Export Factors. *eLife* **2019**, *8*, e47999. https://doi.org/10.7554/eLife.47999.

(13) Batki, J.; Schnabl, J.; Wang, J.; Handler, D.; Andreev, V. I.; Stieger, C. E.; Novatchkova, M.; Lampersberger, L.; Kauneckaite, K.; Xie, W.; Mechtler, K.; Patel, D. J.; Brennecke, J. The Nascent RNA Binding Complex SFiNX Licenses piRNA-Guided Heterochromatin Formation. *Nat. Struct. Mol. Biol.* **2019**, *26* (8), 720–731. https://doi.org/10.1038/s41594-019-0270-6.

(14) Murano, K.; Iwasaki, Y. W.; Ishizu, H.; Mashiko, A.; Shibuya, A.; Kondo, S.; Adachi, S.; Suzuki, S.; Saito, K.; Natsume, T.; Siomi, M. C.; Siomi, H. Nuclear RNA Export Factor Variant Initiates piRNA-Guided Co-Transcriptional Silencing. *EMBO J.* **2019**, *38* (17), e102870. https://doi.org/10.15252/embj.2019102870.

(15) Schnabl, J.; Wang, J.; Hohmann, U.; Gehre, M.; Batki, J.; Andreev, V. I.; Purkhauser, K.; Fasching, N.; Duchek, P.; Novatchkova, M.; Mechtler, K.; Plaschka, C.; Patel, D. J.; Brennecke, J. Molecular Principles of Piwi-Mediated Cotranscriptional Silencing through the Dimeric SFiNX Complex. *Genes Dev.* **2021**, *35* (5–6), 392–409. https://doi.org/10.1101/gad.347989.120.

(16) Eastwood, E. L.; Jara, K. A.; Bornelöv, S.; Munafò, M.; Frantzis, V.; Kneuss, E.; Barbar, E. J.; Czech, B.; Hannon, G. J. Dimerisation of the PICTS Complex via LC8/Cut-up Drives Co-Transcriptional Transposon Silencing in Drosophila. *eLife* **2021**, *10*, e65557. https://doi.org/10.7554/eLife.65557.
(17) Mohn, F.; Sienski, G.; Handler, D.; Brennecke, J. The Rhino-Deadlock-Cutoff Complex Licenses Noncanonical Transcription of Dual-Strand piRNA Clusters in *Drosophila*. *Cell* **2014**, *157* (6), 1364–1379. https://doi.org/10.1016/j.cell.2014.04.031.
(18) Gebert, D.; Neubert, L. K.; Lloyd, C.; Gui, J.; Lehmann, R.; Teixeira, F. K. Large Drosophila Germline piRNA Clusters Are Evolutionarily Labile and Dispensable for Transposon Regulation. *Mol. Cell* **2021**, *81* (19), 3965-3978.e5. https://doi.org/10.1016/j.molcel.2021.07.011.
(19) Akkouche, A.; Kneuss, E.; Bornelöv, S.; Renaud, Y.; Eastwood, E. L.; Lopik, J. van; Gueguen, N.; Jiang, M.; Creixell, P.; Maupetit-Mehouas, S.; Nicholson, B. C.; Brasset, E.; Hannon, G. J. A Dual Histone Code Specifies the Binding of Heterochromatin Protein Rhino to a Subset of piRNA Source Loci. bioRxiv January 11, 2024, p 2024.01.11.575256. https://doi.org/10.1101/2024.01.11.575256.
(20) Brasset, E.; Taddei, A. R.; Arnaud, F.; Faye, B.; Fausto, A. M.; Mazzini, M.; Giorgi, F.; Vaury, C. Viral Particles of the Endogenous Retrovirus ZAM from Drosophila Melanogaster Use a Pre-Existing Endosome/Exosome Pathway for Transfer to the Oocyte. *Retrovirology* **2006**, *3*, 1–9. https://doi.org/10.1186/1742-4690-3-25.
(21) Barckmann, B.; El-Barouk, M.; Pélisson, A.; Mugat, B.; Li, B.; Franckhauser, C.; Lavier, A. S. F.; Mirouze, M.; Fablet, M.; Chambeyron, S. The Somatic piRNA Pathway Controls Germline Transposition over Generations. *Nucleic Acids Res.* **2018**, *46* (18), 9524–9536. https://doi.org/10.1093/nar/gky761.
(22) Duc, C.; Yoth, M.; Mouniée, N.; Jensen, S.; Vaury, C.; Brasset, E. Trapping a Somatic Endogenous Retrovirus into a Germline piRNA Clus- Ter Immunizes the Germline against Further Invasion Keywords. **2019**.
(23) Hirakata, S.; Ishizu, H.; Fujita, A.; Tomoe, Y.; Siomi, M. C. Requirements for Multivalent Yb Body Assembly in Transposon Silencing in Drosophila. *EMBO Rep.* **2019**, *20* (7), e47708. https://doi.org/10.15252/embr.201947708.
(24) Ishizu, H.; Kinoshita, T.; Hirakata, S.; Komatsuzaki, C.; Siomi, M. C. Distinct and Collaborative Functions of Yb and Armitage in Transposon-Targeting piRNA Biogenesis. *Cell Rep.* **2019**, *27* (6), 1822-1835.e8. https://doi.org/10.1016/j.celrep.2019.04.029.
(25) Zhao, K.; Cheng, S.; Miao, N.; Xu, P.; Lu, X.; Zhang, Y.; Wang, M.; Ouyang, X.; Yuan, X.; Liu, W.; Lu, X.; Zhou, P.; Gu, J.; Zhang, Y.; Qiu, D.; Jin, Z.; Su, C.; Peng, C.; Wang, J.-H.; Dong, M.-Q.; Wan, Y.; Ma, J.; Cheng, H.; Huang, Y.; Yu, Y. A Pandas Complex Adapted for piRNA-Guided Transcriptional Silencing and Heterochromatin Formation. *Nat. Cell Biol.* **2019**, *21* (10), 1261–1272. https://doi.org/10.1038/s41556-019-0396-0.
(26) Iwasaki, Y. W.; Sriswasdi, S.; Kinugasa, Y.; Adachi, J.; Horikoshi, Y.; Shibuya, A.; Iwasaki, W.; Tashiro, S.; Tomonaga, T.; Siomi, H. Piwi-piRNA Complexes Induce Stepwise Changes in Nuclear Architecture at Target Loci. *EMBO J.* **2021**, *40* (18), e108345. https://doi.org/10.15252/embj.2021108345.
(27) Andreev, V. I.; Yu, C.; Wang, J.; Schnabl, J.; Tirian, L.; Gehre, M.; Handler, D.; Duchek, P.; Novatchkova, M.; Baumgartner, L.; Meixner, K.; Sienski, G.; Patel, D. J.; Brennecke, J. Panoramix SUMOylation on Chromatin Connects the piRNA Pathway to the Cellular Heterochromatin Machinery. *Nat. Struct. Mol. Biol.* **2022**, *29* (2), 130–142. https://doi.org/10.1038/s41594-022-00721-x.
(28) Duc, C.; Yoth, M.; Jensen, S.; Mouniée, N.; Bergman, C. M.; Vaury, C.; Brasset, E. Trapping a Somatic Endogenous Retrovirus into a Germline piRNA Cluster Immunizes the Germline against Further Invasion. *Genome Biol.* **2019**, *20*, 127. https://doi.org/10.1186/s13059-019-1736-x.

(29) van Lopik, J.; Alizada, A.; Trapotsi, M.-A.; Hannon, G. J.; Bornelöv, S.; Czech Nicholson, B. Unistrand piRNA Clusters Are an Evolutionarily Conserved Mechanism to Suppress Endogenous Retroviruses across the Drosophila Genus. *Nat. Commun.* **2023**, *14*, 7337. https://doi.org/10.1038/s41467-023-42787-1.
(30) Czech, B.; Munafò, M.; Ciabrelli, F.; Eastwood, E. L.; Fabry, M. H.; Kneuss, E.; Hannon, G. J. piRNA-Guided Genome Defense: From Biogenesis to Silencing. *Annu. Rev. Genet.* **2018**, *52*, 131–157. https://doi.org/10.1146/annurev-genet-120417-031441.
(31) Mével-Ninio, M.; Pelisson, A.; Kinder, J.; Campos, A. R.; Bucheton, A. The Flamenco Locus Controls the Gypsy and ZAM Retroviruses and Is Required for Drosophila Oogenesis. *Genetics* **2007**, *175* (4), 1615–1624. https://doi.org/10.1534/genetics.106.068106.
(32) Alizada, A.; Hannon, G. J.; Nicholson, B. C. Ovo Is a Master Regulator of the piRNA Pathway in Animal Ovarian Germ Cells. bioRxiv April 23, 2024. https://doi.org/10.1101/2024.04.23.590802.
(33) Li, M. A.; Alls, J. D.; Avancini, R. M.; Koo, K.; Godt, D. The Large Maf Factor Traffic Jam Controls Gonad Morphogenesis in Drosophila. *Nat. Cell Biol.* **2003**, *5* (11), 994–1000. https://doi.org/10.1038/ncb1058.
(34) Saito, K.; Inagaki, S.; Mituyama, T.; Kawamura, Y.; Ono, Y.; Sakota, E.; Kotani, H.; Asai, K.; Siomi, H.; Siomi, M. C. A Regulatory Circuit for Piwi by the Large Maf Gene Traffic Jam in Drosophila. *Nature* **2009**, *461* (7268), 1296–1299. https://doi.org/10.1038/nature08501.
(35) King, F. J.; Lin, H. Somatic Signaling Mediated by Fs(1)Yb Is Essential for Germline Stem Cell Maintenance during Drosophila Oogenesis. *Dev. Camb. Engl.* **1999**, *126* (9), 1833–1844. https://doi.org/10.1242/dev.126.9.1833.
(36) Handler, D.; Meixner, K.; Pizka, M.; Lauss, K.; Schmied, C.; Gruber, F. S.; Brennecke, J. The Genetic Makeup of the Drosophila piRNA Pathway. *Mol. Cell* **2013**, *50* (5), 762–777. https://doi.org/10.1016/j.molcel.2013.04.031.
(37) Niki, Y.; Yamaguchi, T.; Mahowald, A. P. Establishment of Stable Cell Lines of Drosophila Germ-Line Stem Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (44), 16325–16330. https://doi.org/10.1073/pnas.0607435103.
(38) Alizada, A.; Hannon, G. J.; Nicholson, B. C. Transcriptional Regulation of the piRNA Pathway by Ovo in Animal Ovarian Germ Cells. *Genes Dev.* **2024**. https://doi.org/10.1101/gad.352120.124.
(39) Kudron, M. M.; Victorsen, A.; Gevirtzman, L.; Hillier, L. W.; Fisher, W. W.; Vafeados, D.; Kirkey, M.; Hammonds, A. S.; Gersch, J.; Ammouri, H.; Wall, M. L.; Moran, J.; Steffen, D.; Szynkarek, M.; Seabrook-Sturgis, S.; Jameel, N.; Kadaba, M.; Patton, J.; Terrell, R.; Corson, M.; Durham, T. J.; Park, S.; Samanta, S.; Han, M.; Xu, J.; Yan, K.-K.; Celniker, S. E.; White, K. P.; Ma, L.; Gerstein, M.; Reinke, V.; Waterston, R. H. The ModERN Resource: Genome-Wide Binding Profiles for Hundreds of Drosophila and Caenorhabditis Elegans Transcription Factors. *Genetics* **2018**, *208* (3), 937–949. https://doi.org/10.1534/genetics.117.300657.
(40) Eferl, R.; Wagner, E. F. AP-1: A Double-Edged Sword in Tumorigenesis. *Nat. Rev. Cancer* **2003**, *3* (11), 859–868. https://doi.org/10.1038/nrc1209.
(41) Pogenberg, V.; Consani Textor, L.; Vanhille, L.; Holton, S. J.; Sieweke, M. H.; Wilmanns, M. Design of a bZip Transcription Factor with Homo/Heterodimer-Induced DNA-Binding Preference. *Structure* **2014**, *22* (3), 466–477. https://doi.org/10.1016/j.str.2013.12.017.
(42) Panchal, T.; Chen, X.; Alchits, E.; Oh, Y.; Poon, J.; Kouptsova, J.; Laski, F. A.; Godt, D. 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.* **2017**, *13* (5), e1006790. https://doi.org/10.1371/journal.pgen.1006790.
(43) Li, M.; Hu, X.; Zhang, S.; Ho, M. S.; Wu, G.; Zhang, L. Traffic Jam Regulates the Function of the Ovarian Germline Stem Cell Progeny Differentiation Niche during Pre-Adult Stage in Drosophila. *Sci. Rep.* **2019**, *9*, 10124. https://doi.org/10.1038/s41598-019-45317-6.

(44) Qi, H.; Watanabe, T.; Ku, H.-Y.; Liu, N.; Zhong, M.; Lin, H. The Yb Body, a Major Site for Piwi-Associated RNA Biogenesis and a Gateway for Piwi Expression and Transport to the Nucleus in Somatic Cells. *J. Biol. Chem.* **2011**, *286* (5), 3789–3797. https://doi.org/10.1074/jbc.M110.193888.
(45) Goriaux, C.; Renaud, Y.; Brasset, E.; Desset, S.; Vaury, C. Transcriptional Properties and Splicing of the Flamenco piRNA Cluster. *EMBO Rep.* **2014**, *15* (4), 411–418. https://doi.org/10.1002/embr.201337898.
(46) Munafò, M.; Lawless, V. R.; Passera, A.; MacMillan, S.; Bornelöv, S.; Haussmann, I. U.; Soller, M.; Hannon, G. J.; Czech, B. Channel Nuclear Pore Complex Subunits Are Required for Transposon Silencing in Drosophila. *eLife 10*, e66321. https://doi.org/10.7554/eLife.66321.
(47) Arnold, C. D.; Gerlach, D.; Stelzer, C.; Boryń, Ł. M.; Rath, M.; Stark, A. Genome-Wide Quantitative Enhancer Activity Maps Identified by STARR-Seq. *Science* **2013**, *339* (6123), 1074–1077. https://doi.org/10.1126/science.1232542.
(48) Arnold, C. D.; Gerlach, D.; Spies, D.; Matts, J. A.; Sytnikova, Y. A.; Pagani, M.; Lau, N. C.; Stark, A. Quantitative Genome-Wide Enhancer Activity Maps for Five Drosophila Species Show Functional Enhancer Conservation and Turnover during Cis-Regulatory Evolution. *Nat. Genet.* **2014**, *46* (7), 685–692. https://doi.org/10.1038/ng.3009.
(49) Zabidi, M. A.; Arnold, C. D.; Schernhuber, K.; Pagani, M.; Rath, M.; Frank, O.; Stark, A. Enhancer-Core-Promoter Specificity Separates Developmental and Housekeeping Gene Regulation. *Nature* **2015**, *518* (7540), 556–559. https://doi.org/10.1038/nature13994.
(50) Homolka, D.; Pandey, R. R.; Goriaux, C.; Brasset, E.; Vaury, C.; Sachidanandam, R.; Fauvarque, M.-O.; Pillai, R. S. PIWI Slicing and RNA Elements in Precursors Instruct Directional Primary piRNA Biogenesis. *Cell Rep.* **2015**, *12* (3), 418–428. https://doi.org/10.1016/j.celrep.2015.06.030.
(51) Özata, D. M.; Yu, T.; Mou, H.; Gainetdinov, I.; Colpan, C.; Cecchini, K.; Kaymaz, Y.; Wu, P.-H.; Fan, K.; Kucukural, A.; Weng, Z.; Zamore, P. D. Evolutionarily Conserved Pachytene piRNA Loci Are Highly Divergent among Modern Humans. *Nat. Ecol. Evol.* **2020**, *4* (1), 156–168. https://doi.org/10.1038/s41559-019-1065-1.
(52) Dennis, C.; Zanni, V.; Brasset, E.; Eymery, A.; Zhang, L.; Mteirek, R.; Jensen, S.; Rong, Y. S.; Vaury, C. “Dot COM”, a Nuclear Transit Center for the Primary piRNA Pathway in Drosophila. *PLoS ONE* **2013**, *8* (9), e72752. https://doi.org/10.1371/journal.pone.0072752.
(53) Dennis, C.; Brasset, E.; Sarkar, A.; Vaury, C. Export of piRNA Precursors by EJC Triggers Assembly of Cytoplasmic Yb-Body in Drosophila. *Nat. Commun.* **2016**, *7*, 13739. https://doi.org/10.1038/ncomms13739.
(54) Heiman, M.; Schaefer, A.; Gong, S.; Peterson, J. D.; Day, M.; Ramsey, K. E.; Suárez-Fariñas, M.; Schwarz, C.; Stephan, D. A.; Surmeier, D. J.; Greengard, P.; Heintz, N. A Translational Profiling Approach for the Molecular Characterization of CNS Cell Types. *Cell* **2008**, *135* (4), 738–748. https://doi.org/10.1016/j.cell.2008.10.028.
(55) Sokolova, O. A.; Ilyin, A. A.; Poltavets, A. S.; Nenasheva, V. V.; Mikhaleva, E. A.; Shevelyov, Y. Y.; Klenov, M. S. Yb Body Assembly on the Flamenco piRNA Precursor Transcripts Reduces Genic piRNA Production. *Mol. Biol. Cell* **2019**, *30* (12), 1544–1554. https://doi.org/10.1091/mbc.E17-10-0591.
(56) McGuire, S. E.; Mao, Z.; Davis, R. L. Spatiotemporal Gene Expression Targeting with the TARGET and Gene-Switch Systems in Drosophila. *Sci. STKE Signal Transduct. Knowl. Environ.* **2004**, *2004* (220), pl6. https://doi.org/10.1126/stke.2202004pl6.
(57) Rivera, A.; Lee, J.-H. R.; Gupta, S.; Yang, L.; Goel, R. K.; Zaia, J.; Lau, N. C. Traffic Jam Activates the Flamenco piRNA Cluster Locus and the Piwi Pathway to Ensure Transposon Silencing and Drosophila Fertility, 2024. https://doi.org/10.1101/2024.08.15.608167.
(58) Zhou, L.; Canagarajah, B.; Zhao, Y.; Baibakov, B.; Tokuhiro, K.; Maric, D.; Dean, J. BTBD18 Regulates a Subset of piRNA-Generating Loci through Transcription Elongation in Mice. *Dev. Cell* **2017**, *40* (5), 453-466.e5. https://doi.org/10.1016/j.devcel.2017.02.007.

(59) Yu, T.; Biasini, A.; Cecchini, K.; Säflund, M.; Mou, H.; Arif, A.; Eghbali, A.; de Rooij, D. G.; Weng, Z.; Zamore, P. D.; Özata, D. M. A-MYB/TCFL5 Regulatory Architecture Ensures the Production of Pachytene piRNAs in Placental Mammals. *RNA* **2023**, *29* (1), 30–43. https://doi.org/10.1261/rna.079472.122.
(60) Yoth, M.; Maupetit-Méhouas, S.; Akkouche, A.; Gueguen, N.; Bertin, B.; Jensen, S.; Brasset, E. Reactivation of a Somatic Errantivirus and Germline Invasion in Drosophila Ovaries. *Nat. Commun.* **2023**, *14* (1), 6096. https://doi.org/10.1038/s41467-023-41733-5.
(61) Senti, K.-A.; Handler, D.; Rafanel, B.; Kosiol, C.; Schlötterer, C.; Brennecke, J. Functional Adaptations of Endogenous Retroviruses to the Drosophila Host Underlie Their Evolutionary Diversification, 2023. https://doi.org/10.1101/2023.08.03.551782.
(62) Tcheressiz, S.; Calco, V.; Arnaud, F.; Arthaud, L.; Dastugue, B.; Vaury, C. Expression of the Idefix Retrotransposon in Early Follicle Cells in the Germarium of Drosophila Melanogaster Is Determined by Its LTR Sequences and a Specific Genomic Context. *Mol. Genet. Genomics* **2002**, *267* (2), 133–141. https://doi.org/10.1007/s00438-002-0641-1.
(63) Meignin, C.; Dastugue, B.; Vaury, C. Intercellular Communication between Germ Line and Somatic Line Is Utilized to Control the Transcription of ZAM, an Endogenous Retrovirus from Drosophila Melanogaster. *Nucleic Acids Res.* **2004**, *32* (13), 3799–3806. https://doi.org/10.1093/nar/gkh708.
(64) Yang, Q.; Li, R.; Lyu, Q.; Hou, L.; Liu, Z.; Sun, Q.; Liu, M.; Shi, H.; Xu, B.; Yin, M.; Yan, Z.; Huang, Y.; Liu, M.; Li, Y.; Wu, L. Single-Cell CAS-Seq Reveals a Class of Short PIWI-Interacting RNAs in Human Oocytes. *Nat. Commun.* **2019**, *10* (1), 3389. https://doi.org/10.1038/s41467-019-11312-8.
(65) Kim, D.; Tran, A.; Kim, H. J.; Lin, Y.; Yang, J. Y. H.; Yang, P. Gene Regulatory Network Reconstruction: Harnessing the Power of Single-Cell Multi-Omic Data. *NPJ Syst. Biol. Appl.* **2023**, *9* (1), 51. https://doi.org/10.1038/s41540-023-00312-6.
(66) Andersen, P. R.; Tirian, L.; Vunjak, M.; Brennecke, J. A Heterochromatin-Dependent Transcription Machinery Drives piRNA Expression. *Nature* **2017**, *549* (7670), 54–59. https://doi.org/10.1038/nature23482.
(67) Dennis, C.; Pouchin, P.; Richard, G.; Mirouse, V. Basement Membrane Diversification Relies on Two Competitive Secretory Routes Defined by Rab10 and Rab8 and Modulated by Dystrophin and the Exocyst Complex. *PLOS Genet.* **2024**, *20* (3), e1011169. https://doi.org/10.1371/journal.pgen.1011169.
(68) Vialat, M.; Baabdaty, E.; Trousson, A.; Kocer, A.; Lobaccaro, J.-M. A.; Baron, S.; Morel, L.; de Joussineau, C. Cholesterol Dietary Intake and Tumor Cell Homeostasis Drive Early Epithelial Tumorigenesis: A Potential Modelization of Early Prostate Tumorigenesis. *Cancers* **2024**, *16* (11), 2153. https://doi.org/10.3390/cancers16112153.
(69) Mugat, B.; Akkouche, A.; Serrano, V.; Armenise, C.; Li, B.; Brun, C.; Fulga, T. A.; Van Vactor, D.; Pélisson, A.; Chambeyron, S. MicroRNA-Dependent Transcriptional Silencing of Transposable Elements in Drosophila Follicle Cells. *PLoS Genet.* **2015**, *11* (5), e1005194. https://doi.org/10.1371/journal.pgen.1005194.
(70) Vachias, C.; Tourlonias, C.; Grelée, L.; Gueguen, N.; Renaud, Y.; Venugopal, P.; Richard, G.; Pouchin, P.; Brasset, É.; Mirouse, V. Gap Junction and Amino Acid Import in Somatic Cells Promote Germ Cell Growth. bioRxiv January 26, 2024, p 2023.05.15.540837. https://doi.org/10.1101/2023.05.15.540837.
(71) Quinlan, A. R.; Hall, I. M. BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* **2010**, *26* (6), 841–842. https://doi.org/10.1093/bioinformatics/btq033.
(72) Gu, Z.; Eils, R.; Schlesner, M. Complex Heatmaps Reveal Patterns and Correlations in Multidimensional Genomic Data. *Bioinforma. Oxf. Engl.* **2016**, *32* (18), 2847–2849. https://doi.org/10.1093/bioinformatics/btw313.

Figure legends

Figure 1: Traffic jam regulates the expression of somatic piRNA pathway genes in *Drosophila* ovarian somatic cells

(A) Heatmap showing germline-specific and soma-expressed piRNA pathway genes ranked by gene expression fold-enrichments in the FACS-sorted germline cells (*vas*-GFP+) compared to somatic cells (*vas*-GFP-) dissociated from fly ovaries (*<0.01, DEseq2, RNA-seq n=3 replicates from distinct samples). Normalized RNA-seq counts (DEseq2) from OSCs, vas-GFP+ germline cells, and vas-GFP- somatic cells are also shown. The presence of Tj motifs within ovary ATAC-seq peaks and Tj ChIP-seq peaks (ENCODE, female whole-fly) within ±500 bp of gene TSS is depicted with dots.

(B) Venn diagram showing Tj as the only motif shared between OSC ATAC-seq peaks of the promoter regions (±500 bp of TSS) of somatic-enriched piRNA pathway genes. Full TF motif list is shown in **Table S1**.

(C) Tj ChIP-seq from whole female flies showing Tj binding events to the promoters of somatic piRNA pathway genes (dm6). Grey bars indicate Tj ChIP-seq peaks with the central red lines showing peak summits.

(D) RT-qPCR plot showing 48 h siRNA knockdown of *tj* in OSCs resulting in downregulation of somatic piRNA pathway genes (normalised to *rpL32*; n=3 replicates from distinct samples; error bars indicate standard deviation).

(E) Volcano plot showing differential RNA-seq analysis (DEseq2) between *tj* and *renilla* siRNA knockdowns (48 h; n=3 replicates from distinct samples). Blue dots are showing soma-enriched piRNA pathway genes (*fs(1)Yb*, *nxf2*, *panx*, *soYb* and *armi*); green dots showing general piRNA factors (e.g., *piwi*); red dots are showing *tj*.

(F) Genomic binding profiles of Tj at the promoters of germline- and soma-enriched genes using normalized ChIP-seq (ENCODE; whole-fly; n=3) signals with input subtraction (top) and $\log_2$ of chip rpm/input rpm (bottom).

(G) The top-scoring *de novo* motif within Tj ChIP-seq peaks identified using STREME (MEME-ChIP) (top-left). The numbers and motif logo (MEME) of genomic sites matching the 1-bp gapped pseudo-palindromic motif (MA0659.1; JASPAR 2024) within Tj ChIP-seq peaks identified using FIMO tool ($p<1x10^{-04}$) (top-right). Enrichment of the known

vertebrate motifs within Tj ChIP-seq peak centers using CentriMo (MEME-ChIP) (bottom).
(H) Confocal images of *tj*^{eo2} clonal egg chamber showing Armi (black) and Tj (green) protein expression via immunostaining. Clone is circled in white. Nuclei are labelled with DAPI (white). Scale bars: 10 µm.

Figure 2: Tj binds to conserved *cis*-regulatory elements in *flamenco*

(A) Conservation of ovary ATAC-seq peaks in the *flam* promoter region of *D. melanogaster* (*dm6*) and *D. yakuba* (*droYak2*) species (orthologous regions derived with UCSC LiftOver and chain files). OSC ATAC-seq peaks, OSC STARR-seq peaks, OSC RNA-seq, RNAPol2 ChIP-seq (ovary)[66] and Tj ChIP-seq (ENCODE, whole-fly) are shown.
(B) Venn diagram showing motifs for 21 TFs shared between ATAC-seq peak #7 in *D. melanogaster* and its orthologous peak region in *D. yakuba*. Dot plot showing expression of 21 candidate TFs in the ovaries of *D. melanogaster* and *D. yakuba* (read depth and gene length-normalized DEseq2 counts, RNA-seq n=4).
(C) Soma and germline-enrichment of 21 candidate TFs based on differential gene expression between somatic (*vas*-GFP-) and germline (*vas*-GFP+) cells from the FACS-sorted fly ovaries (DEseq2 was performed between *vas*-GFP- (n=3) and *vas*-GFP+ (n=3) RNA-seq libraries using counts from all genes).

Figure 3: An enhancer of *flam* is located downstream of the TSS

(A) Schematic representation of the transgenic Tomato reporter constructs. The positions of the Tj ChIP-seq peak (line in the middle showing the peak summit), Tj motifs (red) and ATAC-seq peaks (#5-8) are highlighted above.
(B) Confocal images of ovarioles of the indicated reporters showing on the left panels Tomato immunostaining (grey) and on the right panels the merge of Tomato immunostaining (red) and nuclei stained with DAPI (white). Scale bars: 10 µm.
(C) Relative transgene expression levels in ovaries measured by RT-qPCR. *tomato* expression was normalized to *rpl32*. Mean expression is shwon (n=6) with error bars representing standard deviation. Statistical significance was determined using the Mann-Whitney test (** indicating a P-value < 0.005 and * a P-value < 0.05).

(D) Amount of piRNAs produced from the *tomato* reporter sequence, expressed in reads per million (RPM). Nucleotide bias of the first nucleotide of piRNAs are shown as indicated.

Figure 4: The minimal *flam* promoter is crucial for its expression and piRNA biogenesis

(A) Schematic representation of the reporter constructs and *flam* promoter deletion (*flam*△*prom* locus).

(B) Confocal images of ovarioles of the indicated reporters showing Tomato (red) expression by immunostaining. Nuclei are labelled with DAPI (white). Scale bars: 10 µm.

(C) Relative transgene expression levels in ovaries measured by RT-qPCR. Tomato expression was normalized to *rpL32*. Mean expression is shwon (n=6), with error bars representing standard deviation. Statistical analysis was performed using the Mann-Whitney test (** indicates P-value of 0.005).

(D) Boxplot showing changes in piRNA levels for the indicated clusters (split into 1kb tiles) in *flam* mutants (*flam*△*prom*) relative to controls (*flam* WT). The boxplot displays the median (line) and the first and third quartiles (box).

(E) Genome uniquely mapping piRNAs (normalized per million of genome-mapped reads with 0 mismatches) are shown from heterozygous (top) and homozygous *flam* mutant (bottom) ovaries.

(F) Morphology of *Drosophila* ovaries from *flam* heterozygous (top) and homozygous mutant (bottom) flies.

Figure 5: Tj controls *flam* expression

(A) Confocal images of tj^{eo2} clonal egg chamber showing *flam* RNA (magenta) by smFISH RNA and Tj (green) by immunostaining. The clone are circled in white. Nuclei are labelled with DAPI (white). Scale bars: 10 µm.

(B) Uniquely mapped piRNAs are plotted over the *42AB*, *20A* and *flam* piRNA clusters. The piRNA profiles of *tj*-Gal80ts-Gal4; *w-sh* (control) and *tj*-Gal80ts-Gal4; *tj-sh* adult ovaries are shown. Normalized per million of unique genome-mapping piRNAs (0 mismatches).

(C) Confocal images of egg chambers with indicated genotype showing *flam* RNA (magenta) and Tj (green) by Immuno-FISH. A zoom on one egg chamber (white square) is shown to the right. Scale bars: 10 µm.

(D) Relative transgene expression in the indicate ovaries as measured by RT-qPCR. *tomato* expression levels were normalized to the expression of *rpL32.* Mean expression is shwon (n=6) with error bars representing standard deviation.

Figure 6: A subset of transposons appears directly regulated by Tj

(A) Volcano plot showing upregulation of *gypsy*-family TEs by 96 h of *tj* siRNA knockdowns in OSCs using differential mRNA-seq analysis (DEseq2) between *tj* and *renilla* siRNA knockdowns (mRNA-seq, n=3 replicates from distinct samples). Grey circles are showing the genes; blue triangles are showing the transposons.

(B) Shown are beta-gal stainings of egg chambers as readout for gypsy silencing in *tj*-SKD and *w* SKD flies ovary

(C) Volcano plot showing upregulation of *gypsy*-family TEs in *flam∆prom* flies ovary using differential mRNA-seq analysis between *flam∆prom/flam∆prom* and *flam∆prom/FM7* (DEseq2; n=3 replicates from distinct samples). Grey circles are showing the genes; blue triangles are showing the transposons.

(D) Confocal images of *tj*-SKD clonal egg chambers showing *gypsy, ZAM RNA* (grey), Tj (green) and clonal marker (magenta) by Immuno-FISH. *tj*-SKD clones are circled in white. Nuclei are labelled with DAPI (white). Scale bars: 10 µm.

(E) Heatmap showing $\log_2$ fold-change of Tj ChIP-seq signal over input for the indicated TEs that were selected based on expression in *flam∆prom* versus control fly ovaries or *tj* vs control knowckdown in OSCs (padj <0.01 and baseMean normalized by TE length in the top 70% of values for each dataset).

(F) Position plots showing the ChIP-seq and input signals and positions of single and palindromic Tj motifs along the TE consensus sequence for the *gypsy* (top) and *blood* (bottom).

Supplementary Figure legends

Figure S1: Knockdown of *tj* in OSCs downregulates somatic piRNA pathway genes

(A) Volcano plot showing differential RNA-seq analysis (DEseq2) between *tj* and *renilla* siRNA knockdowns (96 h; n=3 replicates from distinct samples) in OSCs. Blue dots are showing the soma-enriched piRNA pathway genes (*fs(1)Yb*, *nxf2*, *panx*, *soYb* and *armi*); green dots showing general piRNA factors (e.g., *piwi*); red dots showing *tj*.

(B) Pseudo-palindromic Tj motifs are found within Tj ChIP-seq peaks at promoters of somatic piRNA pathway components and somatic non-piRNA pathway target genes (indicated by asterisk). Motif logos shown above were generated from the shown Tj motif half-sites. Motif logo for MAFA shown below was taken from JASPAR database (MA1521.1).

Figure S2: The +770; +2086 region of *flam* is insufficient for its transcription

(A) Schematic representation of the transgenic constructs.

(B) Confocal images of egg chambers with the indicated genotypes showing Tomato expression (red) by immunostaining. Nuclei are stained with DAPI (white). Scale bars: 10 µm.

Figure S3: Tj depletion disrupts *flam* expression

Confocal images of tj^{eo2} clonal egg chambers showing *flam* RNA (magenta) and Tj protein (green) levels measured by Immuno-FISH. tj^{eo2} clones are circled in white. Nuclei are labelled with DAPI (white). Scale bars: 10 µm.

Figure S4: TRAP experiments identify additional candidate TFs regulating *flam* expression

(A) Schematic representation of the Translating ribosome affinity purification (TRAP) method performed on somatic or germ cells of ovaries expressing GFP-tagged ribosomal protein RpL10a, driven by *tj*-Gal4 or *nos*-Gal4 respectively. Comparison between UAS-GFP::Rpl10AGFP; *tj*-Gal4 to UAS-GFP::Rpl10AGFP; *nos*-Gal4 conditions.

(B)-(C). Gene ontology analyses of genes enriched in follicles cells. List of the significantly enriched GO Biological Process **(B)** and Molecular Function **(C)** terms.

Figure S5: Effects of candidates from TRAP experiment on *flam* expression

Confocal images of egg chambers with indicated genotype stained for *flam* RNA using smRNA-FISH (magenta). The driver used is *tj*-Gal80ts-Gal4. Nuclei are labelled with DAPI (white). Scale bars: 10 µm.

Figure S6 : **A subset of transposons appears directly regulated by Tj**

(A) Fold-changes in *gypsy* TE expression upon *tj* knockdown in OSCs after 48 h are shown by RT-qPCR (*rpL32* housekeeping control, n=3, error bars indicate standard deviation).

(B) As in **(A)** but comparing fold-changes between 48 h and 96 h of *tj* knockdown

(C) Volcano plot showing upregulation of transposons by 48 h of *tj* siRNA in OSCs using differential mRNA-seq analysis (DEseq2) compared to *renilla* siRNA knockdowns (mRNA-seq, n=3 replicates from distinct samples). Grey circles are showing the genes; blue triangles are showing the transposons.

(D) Fold-changes in TE expression upon *tj*-SKD in ovaries are shown by RT-qPCR (*rpL32* housekeeping control, n=3, error bars indicate standard deviation)

(E) Confocal images of *tj*-SKD clonal egg chambers showing *mdg1, 412 RNA* (grey), Tj (green) and clonal marker (magenta) by Immuno-FISH. *tj*-SKD clones are circled in white. Nuclei are labelled with DAPI (white). Scale bars: 10 µm.

Figure S7: Control of germline and somatic piRNA pathway components relies on Ovo and Tj

(A) Ovo and Tj ChIP-seq showing antagonistic binding at the promoters of germline and somatic genes.

(B) Model depicting the germline and somatic transcriptional regulators of the piRNA pathway components in *Drosophila* ovaries and their putative co-option by transposons for mobilization and infection from soma into the germline.

STAR★Methods

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Goat Polyclonal anti-Armitage (dD-17) antibody	Santa Cruz Biotechnology	Cat# SC-34564, RRID: AB_1134112
Chicken polyclonal anti-mCherry antibody	NOVUS	Cat# NBP2-25158, RRID: AB_2636881
Guinea pig polyclonal anti-Traffic jam antibody	Gift form Dorothea GODT Lab	N/A
Rabbit polyclonal anti-Traffic jam	Gift from Nelson Lau lab	N/A
Cy™3 AffiniPure Donkey Anti-Goat IgG (H+L)	Jackson ImmunoResearch Labs	Cat# 705-165-147, RRID: AB_2307351
Cy3-AffiniPure Donkey Anti-Chicken IgY (IgG) (H+L)	Jackson ImmunoResearch Labs	Cat# 703-165-155, RRID: AB_2340363
Alexa Fluor 488-AffiniPure Donkey Anti-Guinea Pig IgG (H+L)	Jackson ImmunoResearch Labs	Cat# 706-545-148, RRID: AB_2340472
Cy3-AffiniPure Donkey Anti-Rat IgG (H+L)	Jackson ImmunoResearch Labs	Cat# 712-165-150, RRID:AB_2340666
Bacterial and virus strains		
Biological samples		
Fly extract	*Drosophila* Genomics Resource Center (DGRC)	Cat#1645670
Fetal bovine serum	Sigma-Aldrich	Cat#F9665-500ML
Chemicals, peptides, and recombinant proteins		
Formaldehyde solution 16%	Polysciences	Cat#18814_10
Triton X100	euromedex	Cat#2000_B
Formamide	Sigma-Aldrich	Cat#F7503
Dextran sulfate	Sigma-aldrich	Cat#9011-18-1
Atto-488	Lumiprobe	Cat#41698
Atto 550	Sigma-aldrich	Cat#92835
SSC20X	euromedex	Cat#EU0300-C
VectaShield	Vector Laboratory	Cat#H1000
BSA	euromedex	Cat#04-100-812-E
TRIzol Reagent	invitrogen	Cat#15596026
Shields and Sang M3 Insect Medium	Sigma-Aldrich	Cat#S3652-6X1L
Glutathione	Sigma-Aldrich	Cat#G6013-25G
Human insulin	Sigma-Aldrich	Cat#I9278-5ML
Penicillin-Streptomycin	Gibco	Cat#15070063
Critical commercial assays		

NEB Gibson Assembly Cloning Kit	New England Biolabs	Cat#E5510
QuickExtractTM DNA Extraction Solution	Cambio	Cat#QE0905T
SuperScript IV reverse transcriptase	Invitrogen	Cat#18090010
RQ1 RNase-Free DNase	Promega	Cat#M6101
LightCycler 480 Green Master	Roche Diagnostic	Cat#4887352001
Nucleofection solution V	Lonza	Cat#VVCA-1003
Deposited data		
Small RNA-seq and mRNA-seq	This paper	GSE282040
TRAP experiment and RNA sequencing analysis	Vachias and colleagues (85)	GSE230452
Experimental models: Cell lines		
Ovarian somatic cells (OSCs)	*Drosophila* Genomics Resource Center (DGRC)	Stock # 288
Experimental models: Organisms/strains		
D. melanogaster: line expressing dsRNA for RNAi of ***w*** under UAS control in the VALIUM20 vector. y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00045}attP2	Bloomington Drosophila Stock Center	Flybase : FBgn0003996 RRID:BDSC_33644
D. melanogaster: line expressing dsRNA for RNAi of ***tj*** under UAS control in the VALIUM10 vector. y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02009}attP2	Bloomington Drosophila Stock Center	Flybase:FBgn00009 64 RRID:BDSC_25987
D. melanogaster: line expressing dsRNA for RNAi of ***tj*** under UAS control in the VALIUM10 vector. y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01069}attP2	Bloomington Drosophila Stock Center	Flybase: FBgn0000964 RRID:BDSC_34595
D. melanogaster: line expressing dsRNA for RNAi of ***Stat92E*** under UAS control in the VALIUM22 vector. y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.GL00437}attP40/CyO	Bloomington Drosophila Stock Center	Flybase: FBgn0016917 RRID :BDSC_35600
D. melanogaster: line expressing dsRNA for RNAi of ***tgo*** under UAS control in the VALIUM20 vector. y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMC03580}attP40	Bloomington Drosophila Stock Center	Flybase: FBgn0264075 RRID :BDSC_53351
D. melanogaster: line expressing dsRNA for RNAi of ***Chd64*** under UAS control in the VALIUM20 vector. y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMJ23773}attP40/CyO	Bloomington Drosophila Stock Center	Flybase: FBgn0035499 RRID :BDSC_62377
D. melanogaster: line expressing dsRNA for RNAi of ***mamo*** under UAS control in the VALIUM20 vector. y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMC05105}attP2	Bloomington Drosophila Stock Center	Flybase : FBgn0267033 RRID :BDSC_60111
D. melanogaster: line expressing dsRNA for RNAi of ***shn*** under UAS control in the VALIUM20 vector. y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMS06091}attP40	Bloomington Drosophila Stock Center	Flybase : FBgn0003396 RRID :BDSC_82982
D. melanogaster: line expressing dsRNA for RNAi of ***piwi*** under UAS control in the VALIUM20 vector. y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00606}attP2	Bloomington Drosophila Stock Center	Flybase : FBgn0004872 RRID :BDSC_33724

D. melanogaster: line *tj*-Gal4 driver Expresses GAL4 under the control of the traffic jam promoter. y[*] w[*]; P{w[+mW.hs]=GawB}NP1624 / CyO, P{w[-]=UAS-lacZ.UW14}UW14	Kyoto Stock Center (DGRC)	Flybase: FBti0034540 RRID:DGGR_104055
D. melanogaster: line *Tub*-Gal80ts-tj-Gal4 driver Expresses temperature-sensitive GAL80 under the control of the alphaTub84B promoter. Restrictive temp is 30 degrees C. y[*] w[*]; *P{tubP-GAL80[ts]}10,* P{w[+mW.hs]=GawB}NP1624 / CyO, , P{w[-]=UAS-lacZ.UW14}UW14	Construction realized by Mirouse Lab[67]. Recombination between DGGR#10455 and BDSC#67059	FlyBase : FBti0027796, RRID:BDSC_67059 Flybase: FBti0034540 RRID:DGGR_104055
D. melanogaster: line hsFlp; FRT40A-nls-RFP y,w, hs:FLP : P{ry[+t7.2]=hsFLP}22, w[*] ; P{w[+mC]=Ubi-GFP(S65T)nls}2L P{ry[+t7.2]=neoFRT}40A/CyO	Construction realized by Mirouse Lab[67] between BDSC#8862 and BDSC#5629. Bloomington Drosophila Stock Center	FlyBase ID: FBst0008862; BDSC: 8862 FlyBase ID: FBti0015576; BDSC: 5629
D. melanogaster: line b tjeo2 FRT40A/ CyO, bw	from Dorothea Godt lab, University of Toronto	
D. melanogaster: hs-hid(Y); tj Gal4, gyspsy_lacZ/CyO; ;	Vienna BioCenter Core Facilities (VDRC)	VDRC : 313222
D. melanogaster: y y,w,HS:flp122/+;Act:FRTstopFRTGal4, UAS:GFP/CyO;	[68]	
D. melanogaster: line expressing *TdTomato* under the control of *flam[-515; +30]* *Chr 2, 53B2, 2R:16298254 flam[-515; +30]TdTomato-K10 polyA*	This paper	N/A
D. melanogaster: line expressing *TdTomato* under the control of *flam[-515; +432]* *Chr 2, 53B2, 2R:16298254 flam[-515; +432]TdTomato-K10 polyA*	This paper	N/A
D. melanogaster: line expressing *TdTomato* under the control of *flam[-515; +718]* *Chr 2, 53B2, 2R:16298254 flam[-515; +718]TdTomato-K10 polyA*	This paper	N/A
D. melanogaster: line expressing *TdTomato* under the control of *flam[-515; +1200]* *Chr 2, 53B2, 2R:16298254 flam[-515; +1200]TdTomato-K10 polyA*	This paper	N/A
D. melanogaster: line expressing *TdTomato* under the control of *flam[-515; +2086]* *Chr 2, 53B2, 2R:16298254 flam[-515; +2086]TdTomato-K10 polyA*	This paper	N/A
D. melanogaster: line expressing *TdTomato under the control of flam[-515; +6334]* *Chr 2, 53B2, 2R:16298254 flam[-515; +6334]TdTomato-K10 polyA*	This paper	N/A

D. melanogaster: line expressing TdTomato under the control of *flam[-1626; +718]* *Chr 2, 53B2, 2R:16298254 flam[-1626; +718]TdTomato-K10 polyA*	This paper	N/A
D. melanogaster: line expressing *TdTomato* under the control of *flam[-1626; +2086]* *Chr 2, 53B2, 2R:16298254 flam[-1626; +2086]TdTomato-K10 polyA*	This paper	N/A
D. melanogaster: line expressing *TdTomato* under the control of *flam[-770; +2086]* *Chr 2, 53B2, 2R:16298254 flam[+770; +2086]TdTomato-K10 polyA*	This paper	N/A
D. melanogaster: line expressing *TdTomato* under the control of *14XUAS flam[-515; +432]* *Chr 2, 53B2, 2R:16298254 14xUAS flam[-515; +432]TdTomato-K10 polyA.*	This paper	N/A
D. melanogaster: line with a deletion of ChrX:21631436-21632105 in the promoter of *flam*	This paper	N/A
Oligonucleotides		
Probes for *in situ* hybridization, see Table S3	This paper	N/A
Primers for plasmids construction, see Table S3	This paper	N/A
sgRNA see, Table S3	This paper	N/A
Primers used to check *flam* mutant lines, see Table S3	This paper	N/A
RT-qPCR primers, see Table S3	This paper	N/A
Recombinant DNA		
Plasmid constructions listed see on Table S3	This paper	N/A
pCFD5	Addgene	Cat#73914
W+AttB	Addgene	Cat#30326
Software and algorithms		
FastQC (fastqcr_0.1.3)	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/	
Bowtie2 (version 2.4.2)	86	
HTSeq count	87	
Features count (v2.0.1)		
DESeq2 (DESeq2_1.30.1)	73	
Bioconductor	88	
R version 4.0.4	https://www.r-project.org	
sRNAPipe	84	
bedtools	https://bedtools.readthedocs.io/en/latest/	
Omero	https://www.openmicroscopy.org/omero/	
UCSC Genome Browser	https://genome.ucsc.edu/	
UCSC LiftOver tool	https://genome.ucsc.edu/cgi-bin/hgLiftOver	

FIMO tool in MEME Suite (v5.4.1)	https://meme-suite.org/meme/tools/fimo	
MEME- ChIP tool in MEME Suite (v5.4.1)	https://meme-suite.org/meme/tools/meme-chip	
CentriMo in MEME Suite (v5.4.1)	https://meme-suite.org/meme/tools/centrimo	
UCSC Table Browser	https://genome.ucsc.edu/cgi-bin/hgTables	
Burrows-Wheeler Aligner (BWA) (v0.7.17)		
SAMtools (v1.9)		
MACS2 (v2.1.1.20160309)		
deepTools (v3.5.1)	https://deeptools.readthedocs.io/en/3.5.1/	
STAR (v2.7.3a) and v(2.7.8a)		
featureCounts tool (Subread package v1.5.3)	https://subread.sourceforge.net/	
MACS3 (v3.0.0a6)		
irreproducible discovery rate (IDR) tool (v2.0.2)		
Cutadapt tool (v1.18)		
Picard tool (v2.9.0)	https://broadinstitute.github.io/picard/	
bedtools	https://bedtools.readthedocs.io/en/latest/	
ataqv (v1.0.0)	https://github.com/ParkerLab/ataqv	
DSIR tool	http://biodev.cea.fr/DSIR/DSIR.html	
Other		
Public ChIP-seq, ATAC-seq, RNA-seq, see Table S4		
Fly TF motifs	FlyFactorSurvey (http://pgfe.umassmed.edu/TFDBS/)	
Vertebrate TF motifs	JASPAR2022_CORE_vertebrates_non-redundant, EUKARYOTE/jolma2013, and MOUSE/uniprobe_mouse motif databases	
D. melanogaster gene annotations	Ensembl genome database	Drosophila_melanogaster.BDGP6.28.100.gtf
Orthologous *D. yakuba* genes	Genes from chained tBLASTn in UCSC Table Browser	blastDm2FB
Consensus *Drosophila* transposon sequences	https://github.com/bergmanlab/transposons	

Experimental model and study participant details

D. melanogaster flies
Genotypes: listed in Key resources table.
Age: Adult stages, 3 to 6 days old
Sex: female (to study ovaries)
Housing: 20°C or 25°C, on standard medium.
For heat shock experiments flies are placed 1h at 37°C on standard Medium.

Wild-type OSCs
Species: D. melanogaster
Sex: female
Tissue: ovary
Growth media: cultured at 26 °C in Shields and Sang M3 Insect Medium (Sigma-Aldrich, S3652-6X1L) supplemented with 10% fetal bovine serum (Sigma-Aldrich, F9665-500ML), 10% fly extract (DGRC, Stock # 1645670), 0.6 mg/mL glutathione (Sigma-Aldrich, G6013-25G), and 10 mU/mL human insulin (Sigma-Aldrich, I9278-5ML) + 1% Penicillin-Streptomycin (Gibco, 15070063).

Method details

Cell culture and growth media

Wild-type OSCs[34] were purchased from the *Drosophila* Genomics Resource Center (DGRC, cat # 288) and cultured at 26°C in Shields and Sang M3 Insect Medium (Sigma-Aldrich, cat # S3652-6X1L) supplemented with 10% fetal bovine serum (Sigma-Aldrich, cat # F9665-500ML), 10% fly extract (DGRC, cat # 1645670), 0.6 mg/mL glutathione (Sigma-Aldrich, cat # G6013-25G), 10 mU/mL human insulin (Sigma-Aldrich, cat # I9278-5ML), and 1% Penicillin-Streptomycin (Gibco, cat # 15070063).

siRNA knockdowns in cell culture

Sense and antisense 21 nt siRNA sequences for target genes were designed using the DSIR tool (http://biodev.cea.fr/DSIR/DSIR.html) (**Table S3**) and ordered from IDT. The RNA oligonucleotides were resuspended to 400 μM final concentration in RNase-free water. Equal volumes of the resuspended sense and antisense siRNA were mixed and then added to an equal amount of 2x annealing buffer (60 mM potassium acetate; 200 mM HEPES, pH 7.5). The mix was boiled for 5 min at 75°C, then ramped down to 25°C

(- 0.1°C/s) to anneal the siRNA sequences into the siRNA duplex (100 μM final concentration). 2 μL of the final 100 μM siRNA duplex was mixed with 100 μL of Nucleofection solution V (Lonza, VVCA-1003) and transfected into 4 million cells (48 h knockdown) or 5 μL of 100 μM siRNA duplex was mixed with 100 μL of Nucleofection solution V and transfected into 10 million cells (96 h knockdown) using a NucleofectorTM 2b Device and program T-029 (Lonza). Cells were plated into 6-well plates (5 million cells) or 10 cm plates (10 million cells). RNA was either harvested at 48 h or nucleofection was repeated after 48 h and RNA was harvested at 96 h. RNA was either harvested at 48 h or nucleofection was repeated after 48 h and RNA was harvested at 96 h.

RNA isolation and RT-qPCR

OSCs were homogenized by addition of Buffer RLT directly on the cell culture vessel surface. RNA was isolated using RNeasy Mini Kit (QIAGEN, 74106) with RNase-Free DNase Set (QIAGEN, 79254) for DNase digestion during RNA purification as per manufacturer's instruction. Reverse transcription was performed with 100 ng to 1 μg RNA using SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific, 18090010). RT-qPCR was performed on 1:10 diluted cDNA using Fast SYBR Green Master Mix (Thermo Fisher Scientific, 4385610) with Bio-Rad C1000 Thermal Cycler. Primers were designed against exon-exon junctions for genes with introns and *rpL32* was used as an internal standard (**Table S3**). Relative expression was analysed using the delta-delta Ct method[62].

When total RNAs were extracted from fresh ovaries, we used TRIzol (Invitrogen) following the manufacturer's instructions. After DNase treatment (RQ1 RNase-free DNase Promega), cDNA was synthesized using random priming of 1 μg total RNA and SuperScript IV Reverse Transcriptase (Invitrogen). Quantitative PCR was performed with the LightCycler480 SYBR Green I Master mix system (Roche), and the data were analyzed using the LightCycler480 system (Roche). Each experiment was performed 6 times with technical duplicates. RNA levels were calculated with the relative quantification method with *rpL32* housekeeping gene as the reference (for primer sequences, see **Table S3**). Statistical differences were evaluated using Kruskal-Wallis non-parametric test, with corrections for multiple comparisons (Holm–Bonferroni adjustment method).

For mRNA-seq experiments, total RNA was extracted from 30 ovaries using TRIzol (Invitrogen) following the manufacturer's instructions.

RNA-seq

For mRNA-seq from OSCs, selection was performed using Poly(A) mRNA Magnetic Isolation Module (NEB, E7490) with 1 ug of total RNA as input. Libraries were prepared using The NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB, E7420L). Library quality was assessed using Agilent 2100 bioanalyzer (High Sensitivity DNA Kit). Libraries were quantified with KAPA Library Quantification Kit for Illumina (Kapa Biosystems, cat# KK4873) submitted to sequencing on a NovaSeq X System with 2 × 50 bp reads.

For mRNA-seq from ovaries 1 µg of total RNA was send to BGI for library construction using the DNBSEQ Eukaryotic Strand-specific mRNA library workflow. Libraries were sequenced on platform DNBSEQ to a read length of 150 bp paired-end. Raw data were filtered by BGI in order to remove adapter sequences or low-quality reads.

Fly stocks and crosses

All *D. melanogaster* stocks were grown on standard medium at 20°C. RNAi lines expressing dsRNAs[80] or shRNAs[81] against *w* (#33644), *tj* (#25987, #34595), *Stat92E* (#35600), *tango* (#53351), *chd64* (#62377), *mamo* (#60111)*, schnurri (#82982), piwi* (#33724) were obtained from the Bloomington *Drosophila* stock center (BDSC). The *tj*-Gal4 driver was originated from Kyoto Stock Center DGRC, #104055. The *gypsy*-lacZ strain (#313222) was obtained from Vienna *Drosophila* Resource Center (VDRC). The lines *tub*-Gal80ts-tj-Gal4 driver and hsFlp; FRT40A-nls-RFP were a kind gift of V. Mirouse at iGReD[67] and the tj^{eo2} FRT40A/CyO, bw line from Dorothea Godt (University of Toronto). The line *y,w,HS:flp122/+;Act:FRTstopFRTGal4, UAS:GFP/CyO;* is a kind gift of C. De Joussineau. The *tj*-Gal4 driver strain was crossed with the transgenic 14xUAS *flam*[-515; +482] *tomato* flies and with all RNAi strains. *tj*-Gal80ts-Gal4 line was crossed with RNAi lines except RNAi lines against *piwi*; the Gal80ts and Gal4 are driving by *tj*; Gal80ts is inactivated at 30°C. *tj*-Gal4 driver was crossed with shRNA *tj* or shRNA *w* as a control at 25°C. F1 flies obtained were used for mRNA-seq analysis.

Clones and Constructs

All plasmids were constructed using NEB Gibson Assembly Cloning Kit (cat # E5510) in accordance with the manufacturer's instructions and the recommendations of the NEBuilder assembly tool (https://nebuilder.neb.com/). Inserts (listed in the **Table S3**) were amplified by PCR using Phusion taq polymerase (Thermofisher) using primers summarized in **Table S3**, and cloned in the W+attB plasmid (addgene #30326). Ten vectors have been then elaborated in W+attB backbone: nine of them possessed different *flam* fragments sizes upstream of tdTomato reporter sequence directly followed by polyA signal sequence from *K10* gene:

flam[-515; +30]TdTomato-K10 polyA, *flam[-515; +432]TdTomato-K10 polyA*, *flam[-515; +718]TdTomato-K10 polyA*, *flam[-515; +1200]TdTomato-K10 polyA*, *flam[-515; +2086]TdTomato-K10 polyA*, *flam[-515; +6334]TdTomato-K10 polyA*, *flam[-1626; +718]TdTomato-K10 polyA*, *flam[-1626; +2086]TdTomato-K10 polyA*, *flam[+770; +2086]TdTomato-K10 polyA*. The last construct contains the fourteen UAS sequence upstream of the *flam[-515; +432]TdTomato-K10 polyA fragment and is named 14xUAS flam[-515; +432]TdTomato-K10 polyA*.

Transgenic animal production

Plasmid constructions were purified using Plasmid Plus Maxi Kit (QIAGEN). Purified DNA was used for PhiC31 integrase-mediated transgenesis, which was carried out by BestGene (http://www.thebestgene.com/). Transgenes were integrated into the genomic cytological position 53B2.

Generation of *flam* knockout line by CRISPR/Cas9

To remove the entire promoter region of *flamenco* we decided to delete the specific locus ChrX:21631436-21632105 (-455; +215 around TSS). To do so, we used two flanking sgRNAs. The two gRNAs were designed using CRISPR optimal target finder[82]. The individual guide sequences are listed in the **Table S3**. Bolded sequences correspond to each sgRNA. Those primers were used on a pCFD5 template. The resulting PCR product was assembled with the linearised pCFD5 backbone in a single Gibson Assembly

reaction. Then the plasmid was injected in *vas*-Cas9 embryos (BestGene, BL#51324). To establish KO lines, molecular characterisation of target loci was performed as described[83]. Briefly, genomic DNA was extracted from individual larvae using QuickExtract DNA Extraction Solution (Cambio) according to the supplier's instructions. Mutant flies were selected by PCR reactions. Exact deletion of *flam* was precisely identified by PCR and sanger sequencing once homozygote mutant fly lines were obtained (primers listed in **Table S3**).

Mitotic clones of *tj* mutant or knockdown analysis

For mitotic clones of tj mutant, the *Drosophila* strain *hsFlp; FRT40A - nls-RFP*; was crossed with the *tjeo2 FRT40A/CyO, bw* strain. For mitotic clones of *tj* knockdown, *y,w,HS:flp122/+;Act:FRTstopFRTGal4, UAS:GFP/CyO;* flies were crossed with lines carrying shRNAs against *tj* or *w* . All these flies were maintained at 25°C on a rich medium. Three heat shocks at 37°C for 1 h were applied to the F1 progeny, one day before eclosion (pupal stage), on the day of eclosion and one day after. The following day, ovaries were dissected and fixed for either FISH or IF experiments.

β-Gal staining on ovarioles

Ovaries from 5-day-old flies were dissected in PBS, kept on ice, fixed in 0.2% glutaraldehyde/2% formaldehyde/PBS at room temperature for 5 min and rinsed three times with PBS. They were then incubated in staining solution (1x PBS pH7.5, 1mM MgCl2, 4mM potassium ferricyanide, 4mM potassium ferrocyanide, 1% Triton, 2.7mg/ml X-Gal) at 37°C for 1h[69].

FISH and immunostaining

Five to ten ovary pairs were dissected into ice-cold PBS and fixed in formaldehyde solution (4% formaldehyde, 0.3% TritonX-100 in PBS) for 20 min at RT with agitation. The fixed ovaries were then washed three times for 10 min in 0.3% Triton X-100/PBS and permeabilized overnight at 4°C in 70% ethanol. For probe hybridization, permeabilized ovaries were first rehydrated for 5 min in RNA FISH wash buffer (10% (v/w) formamide in 2x SSC). The ovaries were then resuspended in 50 µl hybridization buffer (10% (v/w)

dextran sulphate and 10% (v/w) formamide in 2x SSC). Subsequently, 1.5 µL of a homemade *flam* RNA probe labelled with Atto-488 (Sigma, ref 41698) or Atto-550 (Lumiprobe, ref 92835) was added with or without primary antibody. The samples were incubated at 37°C with agitation. The ovaries were then rinsed twice for 15 min at RT in RNA FISH wash buffer. For FISH-IF, incubation with a secondary antibody conjugated to either Alexa-488, Cy3, or Cy5, was performed in 2x SSC for 1 h 30 at room temperature. After two washes in 2x SSC, DNA was stained with DAPI (1/500). The tissues were then mounted between a slide and coverslip using VectaShield mounting medium (Vector laboratories).

Immunofluorescence

Ovaries from 3- to 5-day-old flies were dissected in Schneider's *Drosophila* Medium, fixed in 4% formaldehyde/PBS for 20 min, rinsed three times with PBT (× 1 PBS, 0.1% Triton, 1% BSA), and incubated in PBT for at least 1 h 30. They were then incubated with primary antibodies overnight. After 3 washes in PBT, ovaries were incubated with the corresponding secondary antibodies (1:1,000), conjugated to Alexa-488, Cy3, or Cy5, respectively, for 90 min. After two washes, DNA was stained with DAPI. The tissues were then mounted between a slide and coverslip using VectaShield mounting medium (Vector laboratories). All antibodies used for immunostaining (IF) along with probe sequences for in situ hybridization (FISH) are listed in **Table S3**. FISH and IF experiments were imaged using Zeiss LSM 800 or LSM 980 Airyscan confocal with a 20x and 40x objectives and analyzed using OMERO figure software.

TRAP experiment and RNA sequencing analysis

The TRAP experiment was described by Vachias and colleagues [70] and data was publicly available (see **Table S4**). Quality control of the sequencing data was evaluated using FastQC software. High quality reads were mapped to the *Drosophila melanogaster* dm6 reference genome using bowtie2 with default parameters[86]. Reads per gene were counted using HTSeqcount[87]. Normalization and differential gene expression analysis were performed using DESeq2[73]. Only genes with an adjusted p-value <0.05 were considered as differentially expressed between the two conditions. Gene expression data

from Tj positive cells were compared to Nanos dataset by applying a fold change ≥ 2, $p < 0.05$ in order to generate list of enriched somatic genes. We computed and compared GO biological processes and molecular functions from this list using an R package cluster profiler from BioConductor[88].

Small RNA sequencing

Total RNA was extracted from 80–100 pairs of ovaries from 3- to 5-day-old flies (for the analysis of piRNA production in somatic follicle cells) using TRIzol reagent (Invitrogen). After 2S rRNA depletion, deep sequencing of 18–30-nt small RNAs was performed by Fasteris S.A. (Geneva/CH) (**Table S4**).

Small RNA-seq analysis

Illumina small RNA-Seq reads were analysed with the small RNA pipeline sRNAPipe[84] to map them to various genomic sequence categories of the *D. melanogaster* genome (release 6.03). All libraries were normalized per million of either genome-mapped reads or unique genome-mapping piRNAs (0 mismatches). All the analyses were performed using genome-unique piRNAs mapped to piRNA clusters, as defined by Brennecke and colleagues [6] (no mismatch allowed). The size profile for *flam* piRNA cluster was obtained by extracting the reads of 23–29 nt sequences uniquely mapped to this locus. Subsequent quantification of reads mapping to 1-kb tiles was done using bedtools, while relative quantification and plotting were done in R. Briefly, small RNA-seq tile signal was normalized to the estimated mappability scores for each 1-kb window. A pseudo-count of 1 was then added to each tile value and log2(fold-change) values relative to control samples was performed.

ATAC-seq data analysis

The read quality was assessed with FastQC (v0.11.8). The paired-end reads were trimmed of adapter sequences TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG and GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG using Cutadapt tool (v1.18; default parameters)[64]. Burrows-Wheeler Aligner (BWA) (v0.7.17, bwa mem -M -t 4)[65] was used to align the trimmed paired reads to 2014 (BDGP Release 6 + ISO1 MT/dm6)

assembly of the *D. melanogaster* genome. Duplicates were marked using Picard tool (v2.9.0) (MarkDuplicates, validation stringency=lenient). SAMtools (v1.9) was used for indexing and filtering[65]. The quality metrics for the aligned ATAC-seq reads were assessed using ataqv (v1.0.0) (https://github.com/ParkerLab/ataqv)[66]. The ATAC-seq peaks were called with MACS2 (v2.1.1.20160309) using --nomodel --shift -37 --extsize 73 parameters and FDR cut-off $q \leq 0.05$[67]. RPKM normalized bigWig files were generated using deepTools (v3.5.1, bamCoverage)[68]. Data were visualized using the UCSC Genome Browser[69].

RNA-seq data analysis

The reads were trimmed of adapter sequences using Cutadapt tool (v1.18; -m 1 specified to not have reads trimmed to zero)[64]. The trimmed reads were aligned to the genome assemblies using the RNA-seq aligner STAR (v2.7.3a)[70]. Gene counts were calculated with the featureCounts tool (Subread package v1.5.3; -s 2 for stranded libraries prepared by dUTP method)[71]. *D. melanogaster* gene annotations were taken from Drosophila_melanogaster.BDGP6.28.100.gtf in the Ensembl genome database[72] and annotations of orthologous *D. yakuba* genes were derived from *D. melanogaster* proteins mapped onto *D. yakuba* genome by chained tBLASTn (blastDm2FB in UCSC Table Browser). The read depth-normalized ovary RNA-seq counts were then normalized to species gene lengths. Consensus *Drosophila* transposon sequences were taken from (https://github.com/bergmanlab/transposons). Samtools (v1.9) was used to merge bam files from replicates[65]. Differential gene expression analysis was performed using DESeq2 (v1.30.1)[73]. The RPKM normalized bigWig files were generated for each strand using bamCoverage tool (--filterRNAstrand specified for dUTP stranded libraries) in deepTools (v3.5.1)[68]. Raw data for publicly available RNA-seq reads were downloaded from sources listed in **(Table S4)** and processed as above. Data were visualized using the UCSC Genome Browser[69].

After quality control of reads using FASTQC (fastqcr_0.1.3), high quality reads were mapped to the *Drosophila melanogaster* dm6 reference genome (release 6_48 from flybase database), and consensus *Drosophila* transposon sequences (https://github.com/bergmanlab/transposons) using STAR (v2.7.8a) [70]. Gene counts were

calculated with the featureCounts tool (featureCounts v2.0.1). Normalization and differential gene expression analysis were performed using DESeq2[73] (DESeq2_1.30.1). Volcano plots were generated using ggplot2 (v 3.3.3) in R (v 4.0.4).

ChIP-seq data analysis

The reads were trimmed of adapter sequences using Cutadapt tool (v1.18)[64]. The trimmed reads were aligned to the genome assemblies and TE consensus sequence assemblies (https://github.com/bergmanlab/drosophila-transposons/) using the Burrows-Wheeler Aligner (BWA) (v0.7.17, bwa aln)[74]. SAMtools (v1.9) was used for sorting, merging, and indexing[65]. ChIP-seq peaks were called using MACS3 (v3.0.0a6) callpeak command (-q 0.01)[67] with data from ChIP input (the whole cell lysate) used as the control (-c). The reproducibility of ChIP-seq peaks between replicates was evaluated using the irreproducible discovery rate (IDR) tool (v2.0.2)[75]. The RPKM normalized bigWig files were generated using bamCoverage tool in deepTools (v3.5.1)[68] with the parameter (--extendReads 120) specified. The input-normalized bigWig files were generated using the bamCompare tool in the deepTools (v3.5.1)[68]. ChIP-seq heatmaps and profiles were generated using the plotHeatmap and plotProfile tools in the deepTools (v3.5.1)[68]. Raw data for publicly available ChIP-seq datasets were downloaded from sources listed in **(Table S4)** and processed as describe above. Data were visualized using the UCSC Genome Browser[69].

Coverage plots over the TE consensus sequences were generated using the genomecov function in bedtools (v 2.26.0)[71] with 1 nt bins and plotted with ggplot2 (v 3.5.1) in R (v 4.4.0). Normalised Tj ChIP coverage across the TE consensus sequences were created using the bamCompare function in deeptools (v 3.5.1) with RPKM normalisation and the log2FC operation. All TE consensus sequences were scaled to 5,000 bits using linear interpolation and the resulting coverage was plotted using ComplexHeatmap (v 2.20.0)[72] in R (v 4.4.0).

Motif discovery and ATAC-seq peak conservation

De novo motif analysis of ChIP-seq peaks and known motif enrichments within peak centers were performed using the MEME-ChIP tool with the positional distribution of

motifs calculated by CentriMo tool in MEME Suite (v5.4.1)[77]. Motif scanning was performed using FIMO tool (match p<0.001) in MEME Suite (v5.4.1) (https://meme-suite.org/meme/index.html)[76]. The list of *Drosophila* motifs was downloaded from the FlyFactorSurvey (http://pgfe.umassmed.edu/TFDBS/) database [78] and lists of known vertebrate motifs for peak central CentriMo enrichments were taken from JASPAR2022_CORE_vertebrates_non-redundant, EUKARYOTE/jolma2013, and MOUSE/uniprobe_mouse motif databases. fastaFromBed (v2.26.0) was used for conversion of bed file coordinates into fasta format using *D. melanogaster* BDGP6.28.dna.toplevel genome. Genomic loci were visualized on the UCSC genome browser[69]. Orthologous ATAC-seq peak regions between the species genomes were derived using the UCSC LiftOver tool (https://genome.ucsc.edu/cgi-bin/hgLiftOver)[79]. *D. melanogaster* ATAC-seq peaks were classified as conserved if the peak orthologous LiftOver regions in the *D. yakuba* genome overlapped with the *D. yakuba* ATAC-seq peaks.

Data and code availability

RNA-seq and small-RNA-seq data generated in this study have been deposited to GEO database under series **GSE282040**.

Quantification and statistical analysis

Each experiment of qPCR data relative to *tomato* expression on fresh ovaries was performed 6 times with technical duplicates. Statistical differences were evaluated using Kruskal-Wallis non-parametric test, with corrections for multiple comparisons (Holm–Bonferroni adjustment method). Each experiment of qPCR relative to TE expression on ovaries was performed 3 times with technical duplicates. Fold changes of TE form *tj*-SKD compared to *w*-SKD were calculated. Statistical differences between *tj*-SKD and control (*w*-SKD) were evaluated using Wilcoxon non-parametric test using R software.

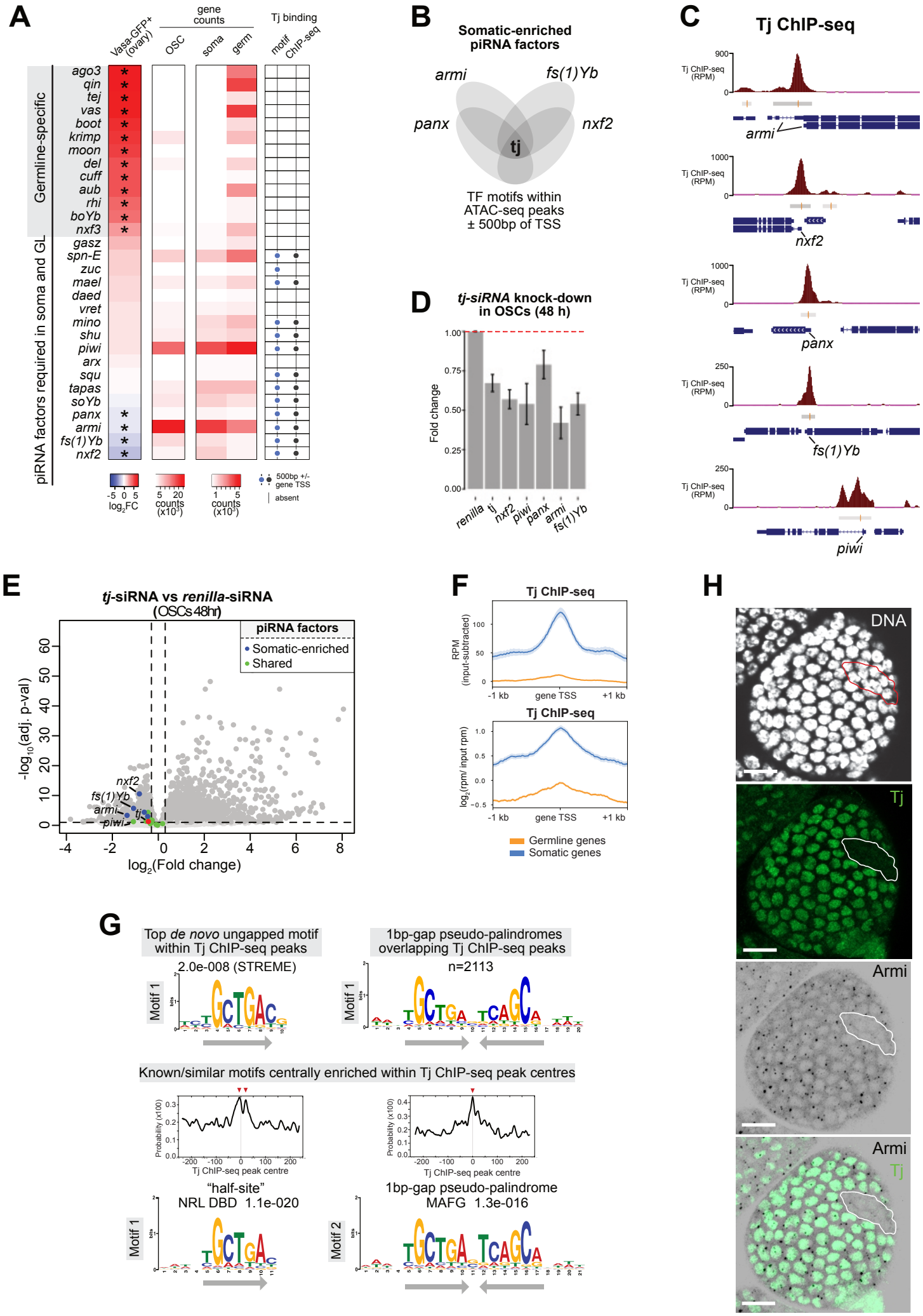
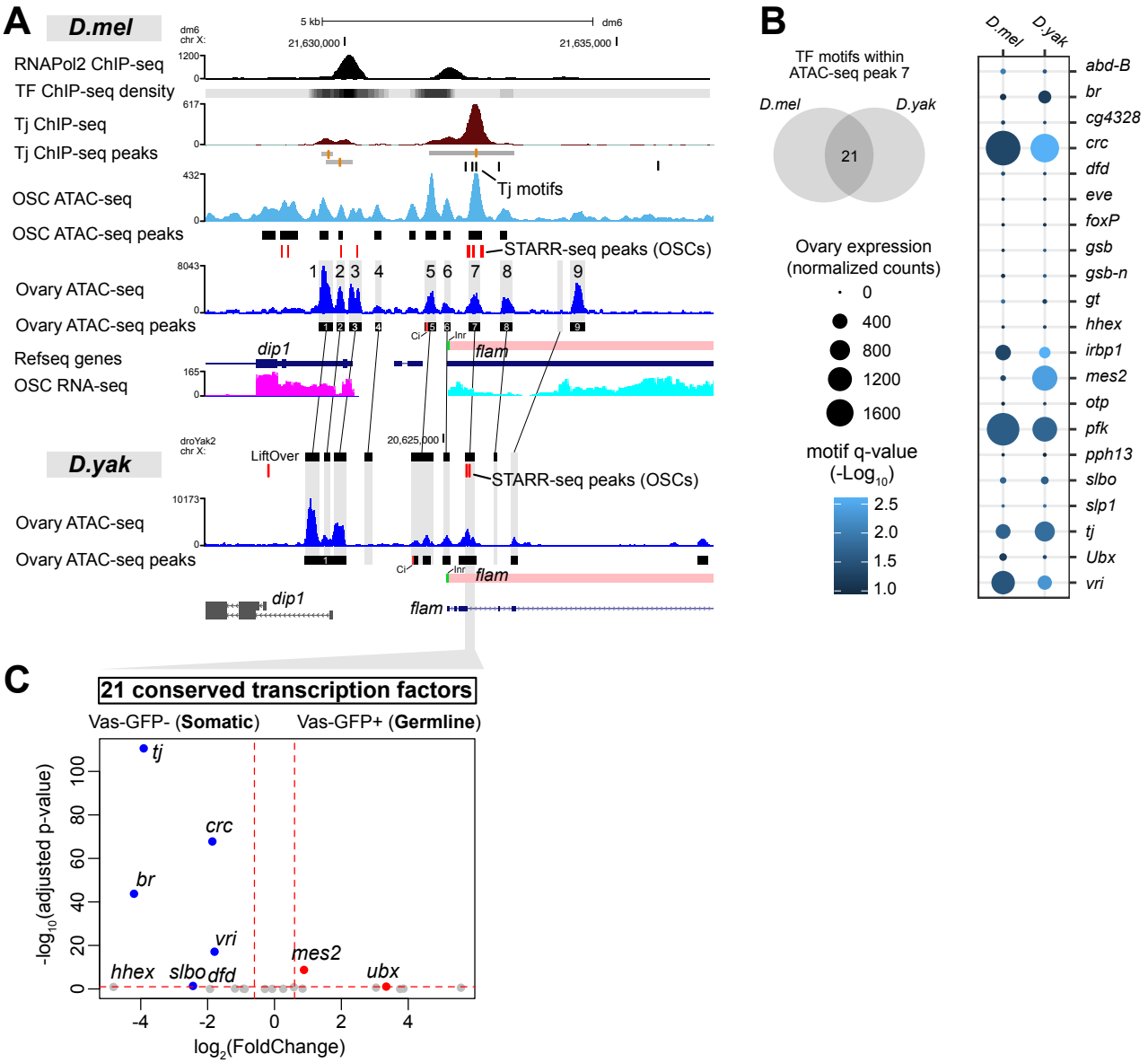
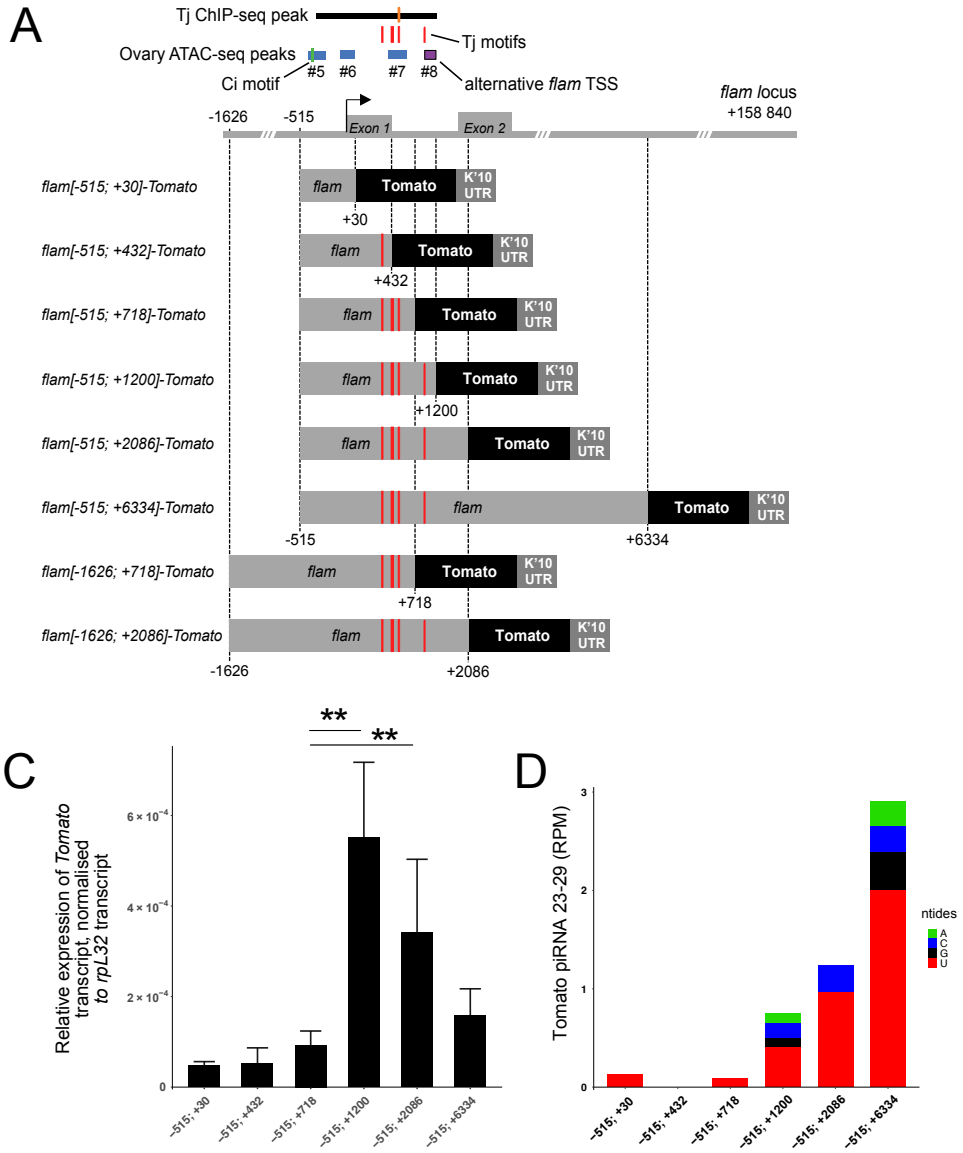
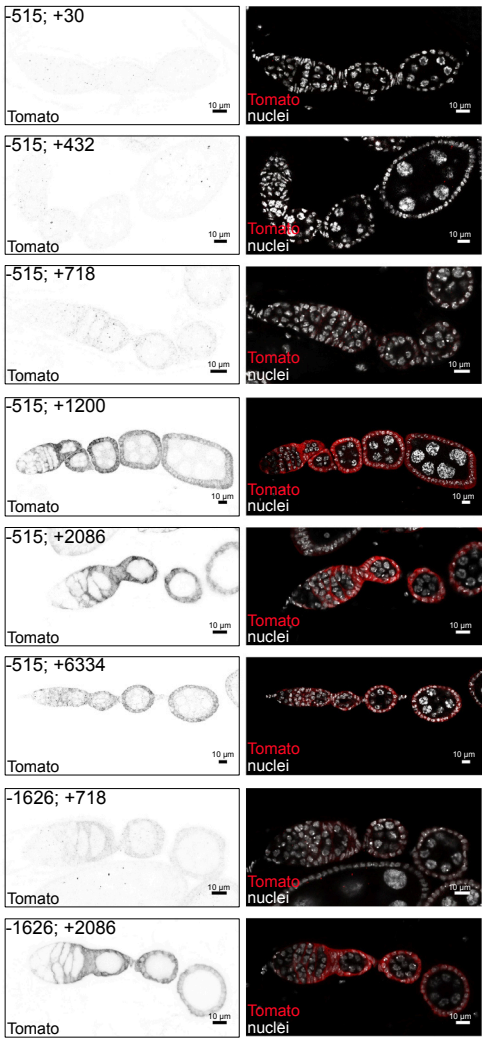
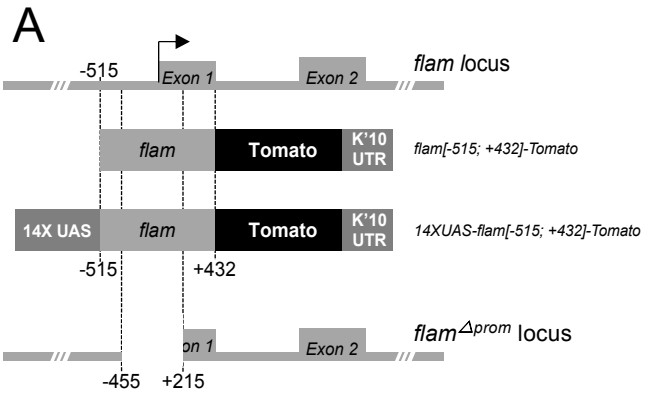

A
Vasa-GFP+ (ovary)
gene counts
OSC
soma
germ
Tj binding
motif
ChIP-seq
Germline-specific
piRNA factors required in soma and GL
ago3
qin
tej
vas
boot
krimp
moon
del
cuff
aub
rhi
boYb
nxf3
gasz
spn-E
zuc
mael
daed
vret
mino
shu
piwi
arx
squ
tapas
soYb
panx
armi
fs(1)Yb
nxf2
log2FC
counts (x10³)
counts (x10⁵)
500bp +/- gene TSS
absent
B
Somatic-enriched piRNA factors
armi
fs(1)Yb
panx
nxf2
tj
TF motifs within ATAC-seq peaks ± 500bp of TSS
C
Tj ChIP-seq
Tj ChIP-seq (RPM)
armi
nxf2
panx
fs(1)Yb
piwi
D
tj-siRNA knock-down in OSCs (48 h)
Fold change
renilla
tj
nxf2
piwi
panx
armi
fs(1)Yb
E
tj-siRNA vs renilla-siRNA (OSCs 48hr)
piRNA factors
Somatic-enriched
Shared
-log10(adj. p-val)
log2(Fold change)
nxf2
fs(1)Yb
armi
tj
piwi
F
Tj ChIP-seq
RPM (input-subtracted)
gene TSS
-1 kb
+1 kb
log2(rpm/ input rpm)
Germline genes
Somatic genes
H
DNA
Tj
Armi
Armi
Tj
G
Top de novo ungapped motif within Tj ChIP-seq peaks
2.0e-008 (STREME)
Motif 1
1bp-gap pseudo-palindromes overlapping Tj ChIP-seq peaks
n=2113
Motif 1
Known/similar motifs centrally enriched within Tj ChIP-seq peak centres
Probability (x100)
Tj ChIP-seq peak centre
"half-site"
NRL DBD 1.1e-020
Motif 1
1bp-gap pseudo-palindrome
MAFG 1.3e-016
Motif 2

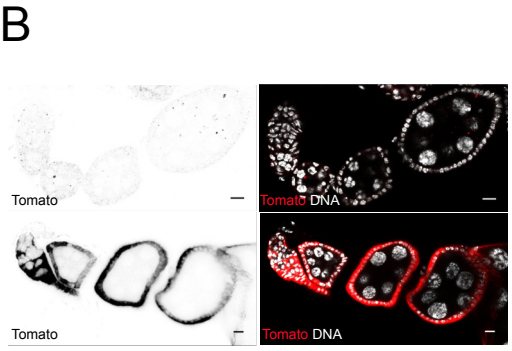
Figure 2

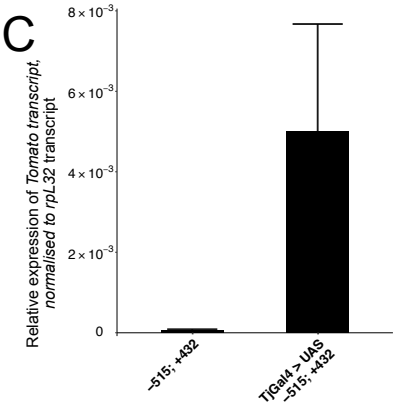

A
D.mel
RNAPol2 ChIP-seq
TF ChIP-seq density
Tj ChIP-seq
Tj ChIP-seq peaks
OSC ATAC-seq
OSC ATAC-seq peaks
Tj motifs
STARR-seq peaks (OSCs)
Ovary ATAC-seq
Ovary ATAC-seq peaks
Refseq genes
OSC RNA-seq
dip1
flam
D.yak
LiftOver
STARR-seq peaks (OSCs)
Ovary ATAC-seq
Ovary ATAC-seq peaks
B
TF motifs within ATAC-seq peak 7
D.mel
D.yak
21
Ovary expression (normalized counts)
0
400
800
1200
1600
motif q-value (-Log10)
2.5
2.0
1.5
1.0
abd-B
br
cg4328
crc
dfd
eve
foxP
gsb
gsb-n
gt
hhex
irbp1
mes2
otp
pfk
pph13
slbo
slp1
tj
Ubx
vri
C
21 conserved transcription factors
Vas-GFP- (Somatic)
Vas-GFP+ (Germline)
-log10(adjusted p-value)
log2(FoldChange)
tj
crc
br
vri
hhex
slbo
dfd
mes2
ubx

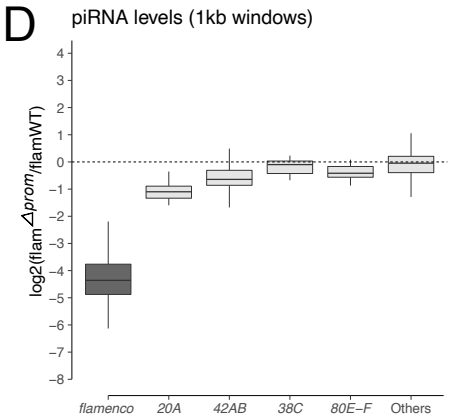

A
Tj ChIP-seq peak
Tj motifs
Ovary ATAC-seq peaks
#5
#6
#7
#8
Ci motif
alternative flam TSS
flam locus
+158 840
-1626
-515
Exon 1
Exon 2
flam[-515; +30]-Tomato
flam[-515; +432]-Tomato
flam[-515; +718]-Tomato
flam[-515; +1200]-Tomato
flam[-515; +2086]-Tomato
flam[-515; +6334]-Tomato
flam[-1626; +718]-Tomato
flam[-1626; +2086]-Tomato
flam
Tomato
K'10 UTR
+30
+432
+1200
+718
+2086
+6334
C
Relative expression of Tomato transcript, normalised to rpL32 transcript
D
Tomato piRNA 23-29 (RPM)
ntides
A
C
G
U

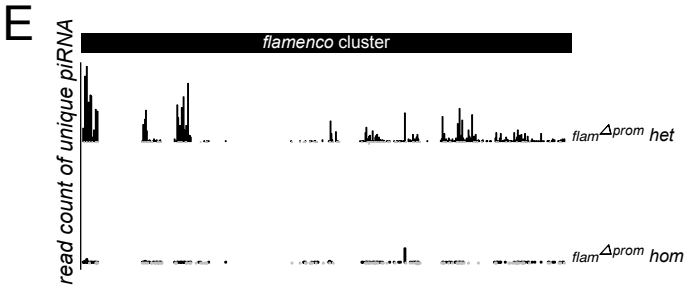

B
-515; +30
-515; +432
-515; +718
-515; +1200
-515; +2086
-515; +6334
-1626; +718
-1626; +2086
Tomato
nuclei


A
-515
Exon 1
Exon 2
flam locus
flam
Tomato
K'10 UTR
flam[-515; +432]-Tomato
14X UAS
flam
Tomato
K'10 UTR
14XUAS-flam[-515; +432]-Tomato
-515
+432
flam^Δprom locus
Exon 2
-455
+215


B
Tomato
Tomato DNA
Tomato
Tomato DNA


C
Relative expression of Tomato transcript, normalised to rpL32 transcript
8 × 10⁻³
6 × 10⁻³
4 × 10⁻³
2 × 10⁻³
0
-515; +432
TjGal4 > UAS -515; +432


D
piRNA levels (1kb windows)
log2(flam^Δprom/flamWT)
4
3
2
1
0
−1
−2
−3
−4
−5
−6
−7
−8
flamenco
20A
42AB
38C
80E−F
Others


E
flamenco cluster
read count of unique piRNA
flam^Δprom het
flam^Δprom hom

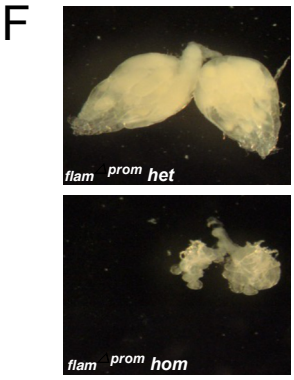
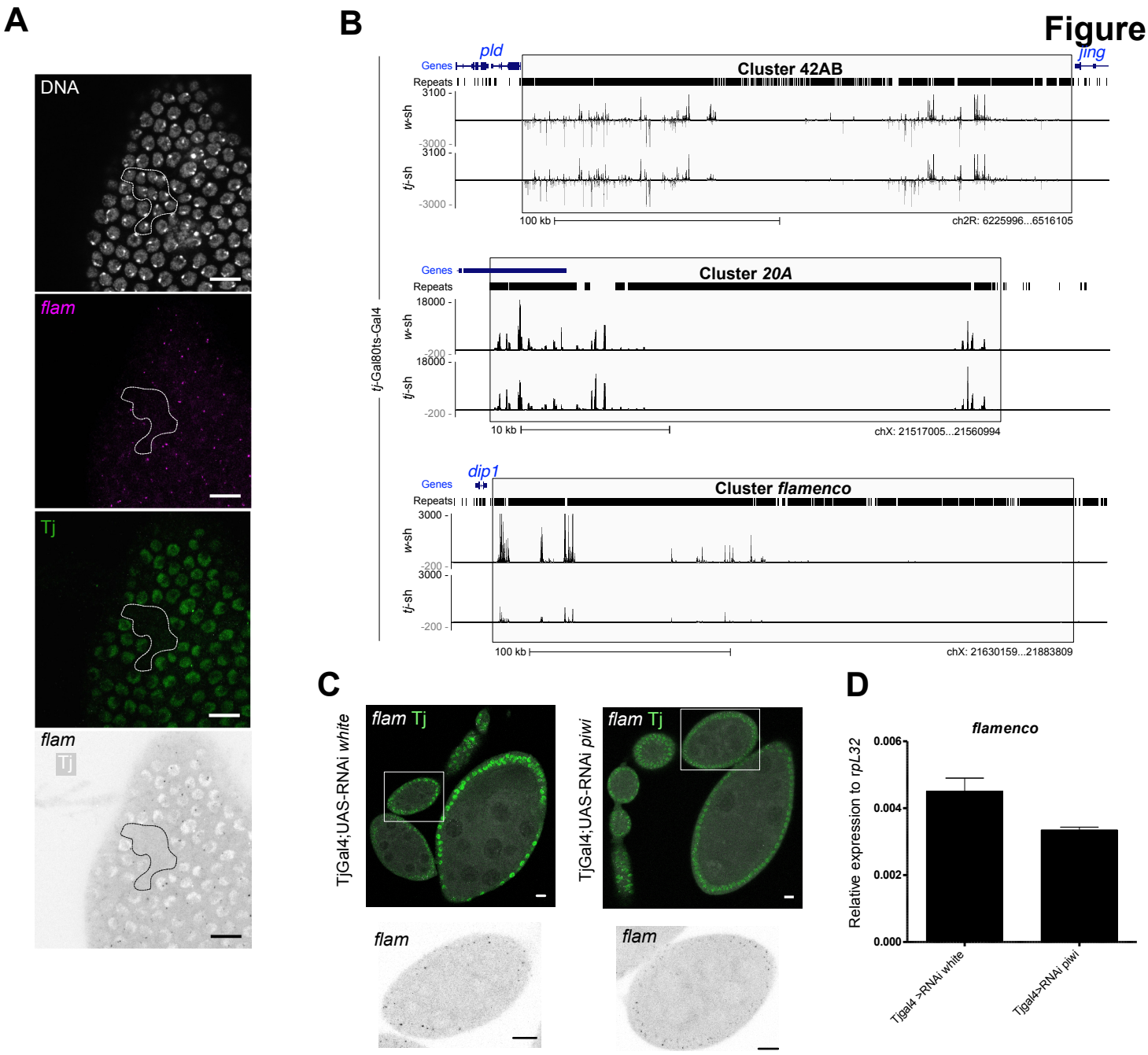
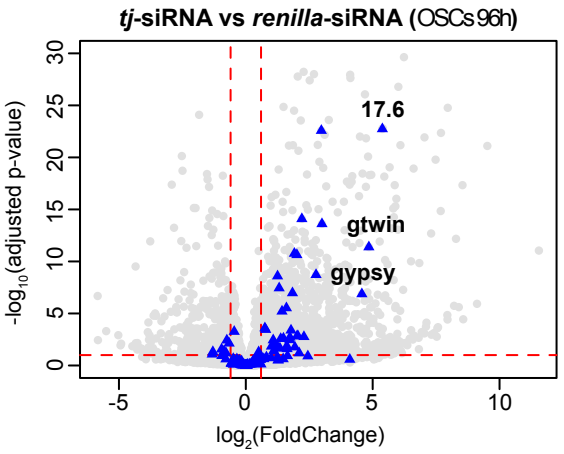
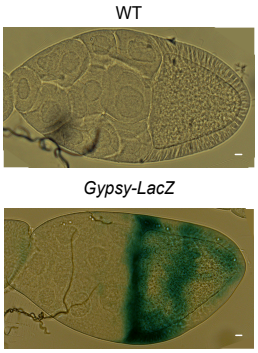
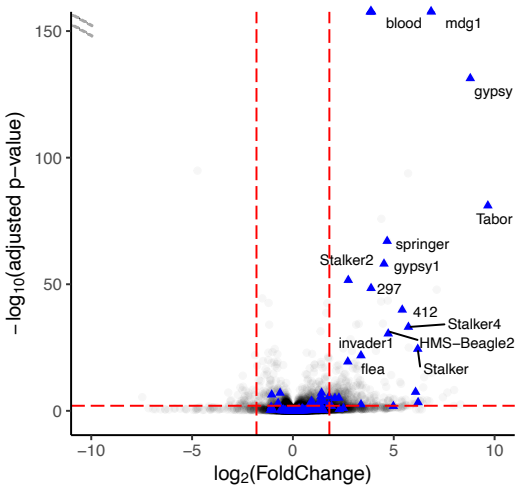

F
flam prom het
flam prom hom

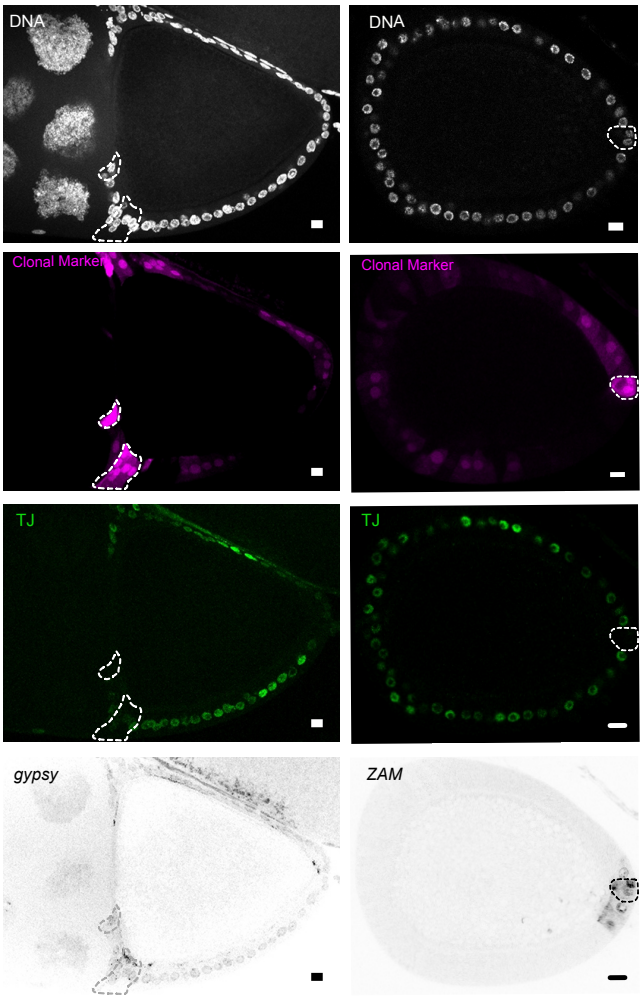
Figure 5

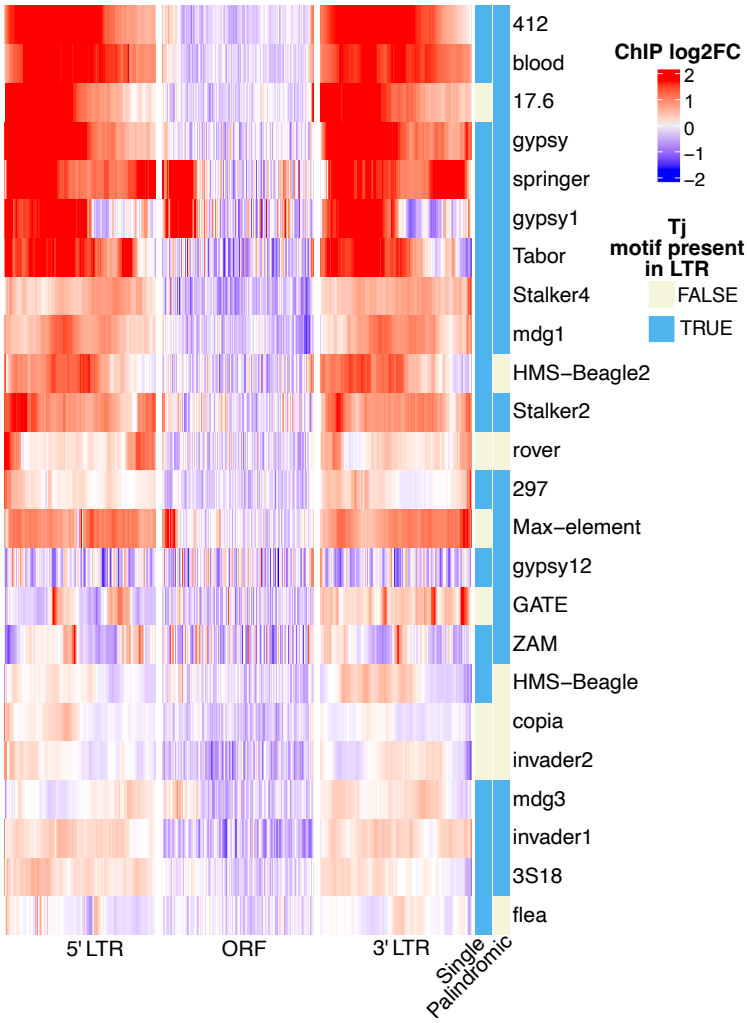

A
DNA
flam
Tj
flam
Tj
B
pld
jing
Genes
Repeats
Cluster 42AB
w-sh
tj-sh
3100
-3000
100 kb
ch2R: 6225996...6516105
Cluster 20A
18000
-200
10 kb
chX: 21517005...21560994
dip1
Cluster flamenco
3000
100 kb
chX: 21630159...21883809
tj-Gal80ts-Gal4
C
TjGal4;UAS-RNAi white
TjGal4;UAS-RNAi piwi
flam Tj
flam
D
flamenco
Relative expression to rpL32
0.006
0.004
0.002
0.000
Tjgal4 >RNAi white
Tjgal4>RNAi piwi

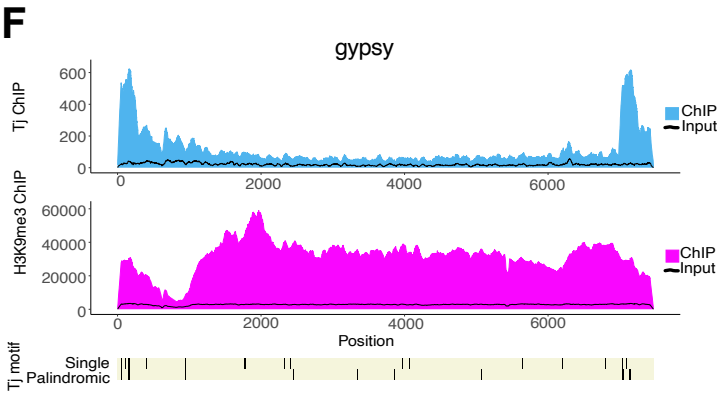

A
tj-siRNA vs renilla-siRNA (OSCs 96h)
17.6
gtwin
gypsy
-log10(adjusted p-value)
log2(FoldChange)


B
WT
Gypsy-LacZ


C
flamΔprom vs control (flamΔprom/WT)
blood
mdg1
gypsy
Tabor
springer
Stalker2
gypsy1
297
412
Stalker4
invader1
HMS-Beagle2
flea
Stalker
−log10(adjusted p−value)
log2(FoldChange)


D
DNA
DNA
Clonal Marker
Clonal Marker
TJ
TJ
gypsy
ZAM


E
412
blood
17.6
gypsy
springer
gypsy1
Tabor
Stalker4
mdg1
HMS−Beagle2
Stalker2
rover
297
Max−element
gypsy12
GATE
ZAM
HMS−Beagle
copia
invader2
mdg3
invader1
3S18
flea
5' LTR
ORF
3' LTR
Single
Palindromic
ChIP log2FC
Tj motif present in LTR
FALSE
TRUE


F
gypsy
Tj ChIP
H3K9me3 ChIP
ChIP
Input
Position
Tj motif
Single
Palindromic

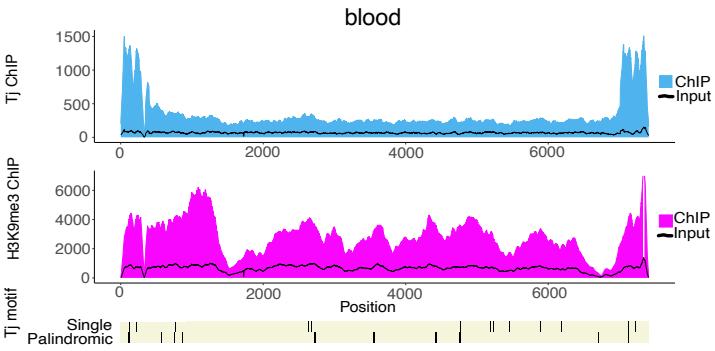

blood
Tj ChIP
H3K9me3 ChIP
ChIP
Input
Position
Tj motif
Single
Palindromic

Figure S1

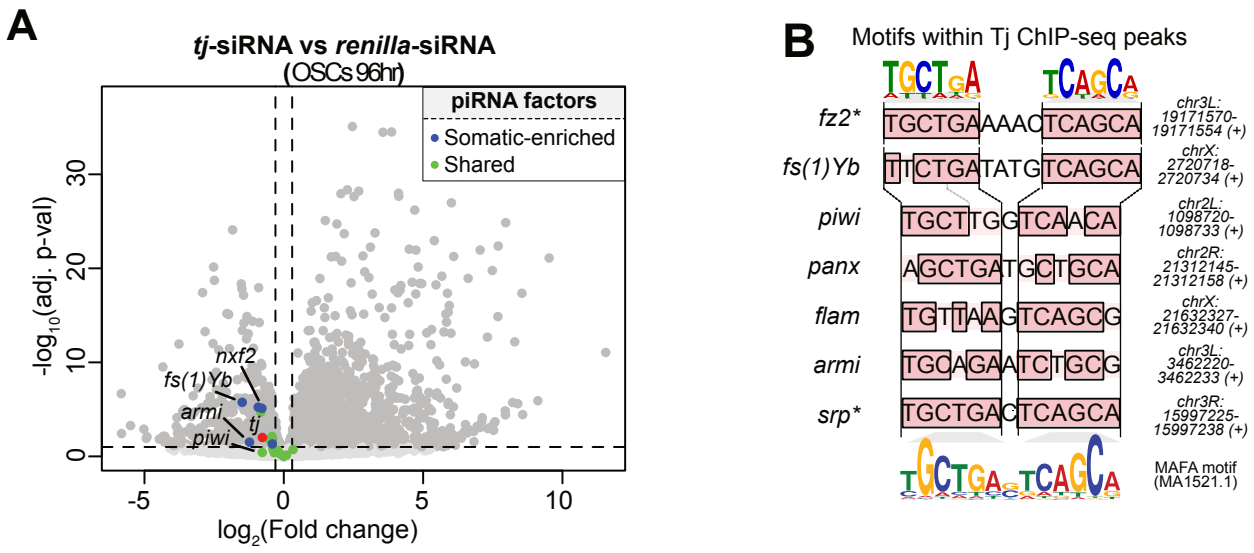

A
tj-siRNA vs renilla-siRNA
(OSCs 96hr)
piRNA factors
Somatic-enriched
Shared
nxf2
fs(1)Yb
armi
tj
piwi
-log10(adj. p-val)
log2(Fold change)
B
Motifs within Tj ChIP-seq peaks
fz2*
fs(1)Yb
piwi
panx
flam
armi
srp*
chr3L: 19171570-19171554 (+)
chrX: 2720718-2720734 (+)
chr2L: 1098720-1098733 (+)
chr2R: 21312145-21312158 (+)
chrX: 21632327-21632340 (+)
chr3L: 3462220-3462233 (+)
chr3R: 15997225-15997238 (+)
MAFA motif (MA1521.1)

Figure S2

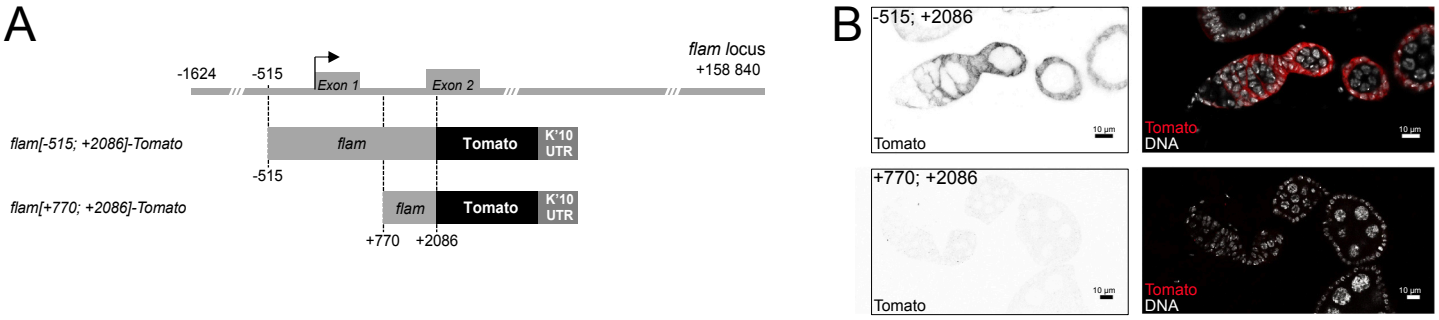

A
flam locus
+158 840
-1624
-515
Exon 1
Exon 2
flam[-515; +2086]-Tomato
flam
Tomato
K'10
UTR
-515
flam[+770; +2086]-Tomato
flam
Tomato
K'10
UTR
+770
+2086
B
-515; +2086
Tomato
10 µm
Tomato
DNA
10 µm
+770; +2086
Tomato
10 µm
Tomato
DNA
10 µm

Figure S3

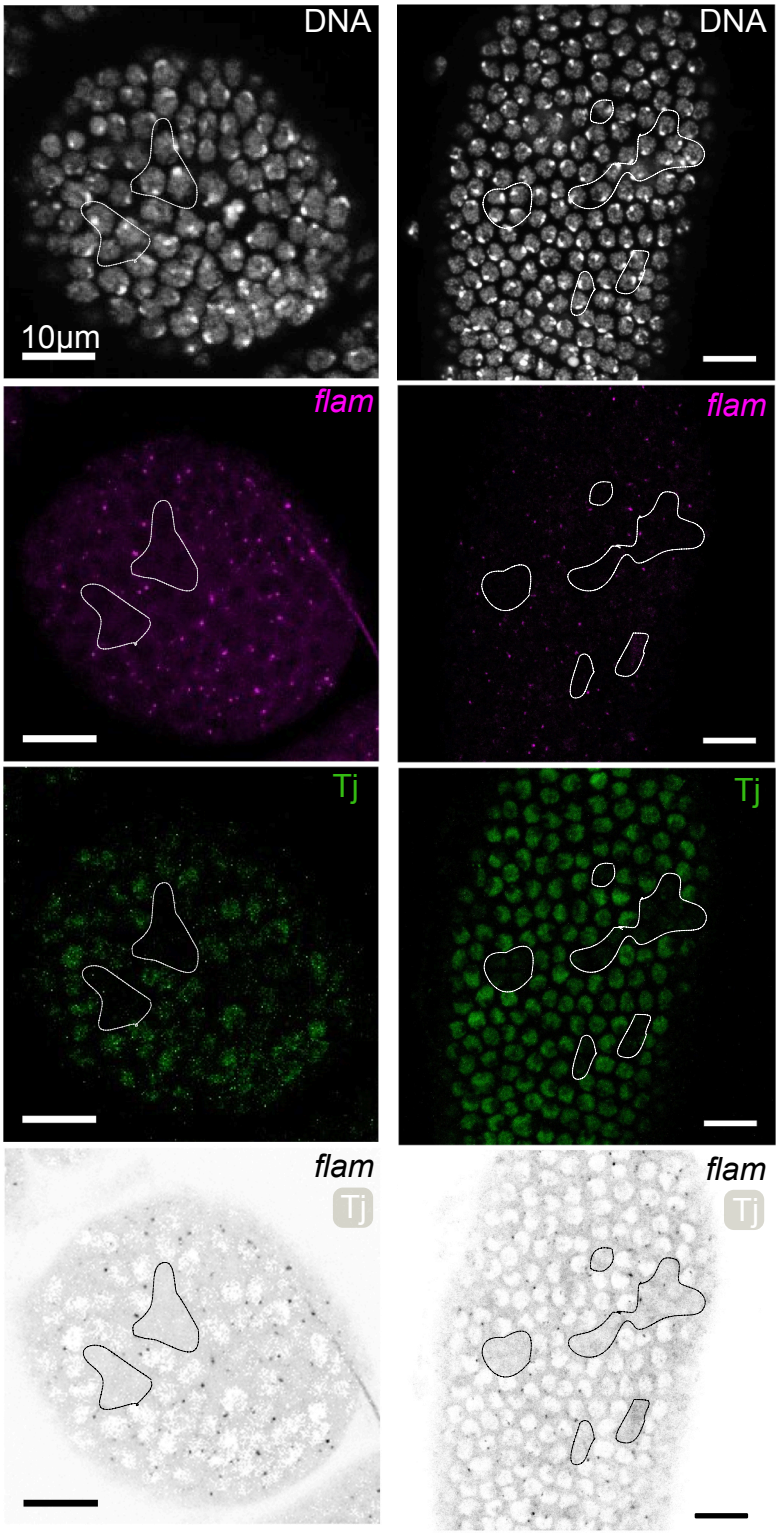


Figure S4

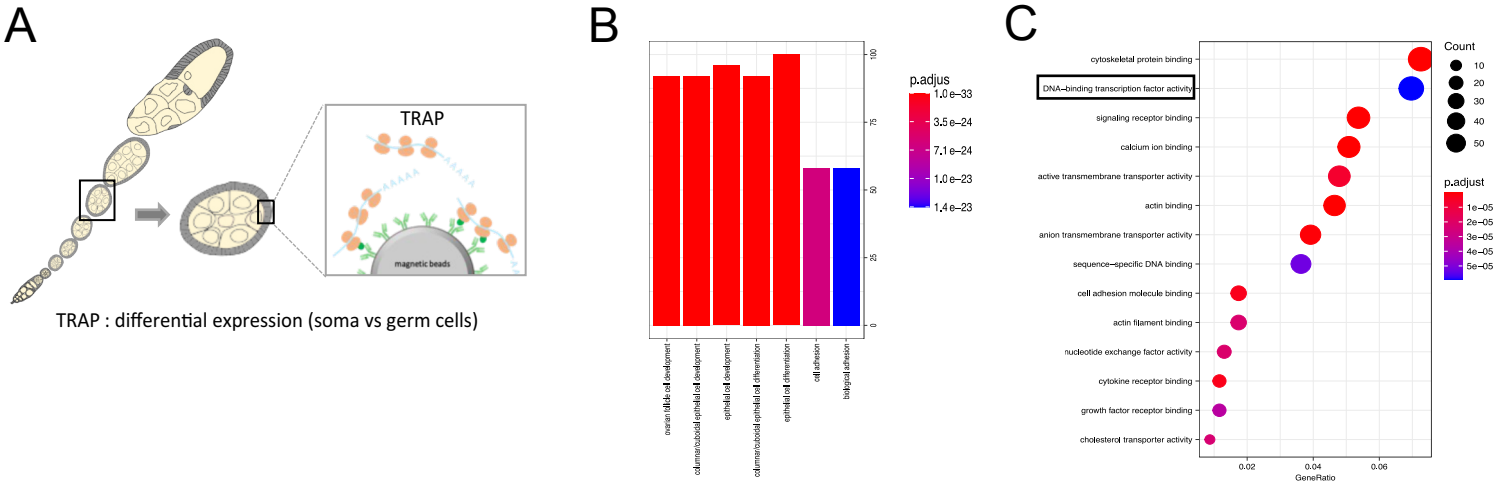

A
TRAP
magnetic beads
TRAP : differential expression (soma vs germ cells)
B
p.adjus
1.0 e−33
3.5 e−24
7.1 e−24
1.0 e−23
1.4 e−23
0
25
50
75
100
ovarian follicle cell development
columnar/cuboidal epithelial cell development
epithelial cell development
columnar/cuboidal epithelial cell differentiation
epithelial cell differentiation
cell adhesion
biological adhesion
C
cytoskeletal protein binding
DNA−binding transcription factor activity
signaling receptor binding
calcium ion binding
active transmembrane transporter activity
actin binding
anion transmembrane transporter activity
sequence−specific DNA binding
cell adhesion molecule binding
actin filament binding
nucleotide exchange factor activity
cytokine receptor binding
growth factor receptor binding
cholesterol transporter activity
0.02
0.04
0.06
GeneRatio
Count
10
20
30
40
50
p.adjust
1e−05
2e−05
3e−05
4e−05
5e−05

Figure S5

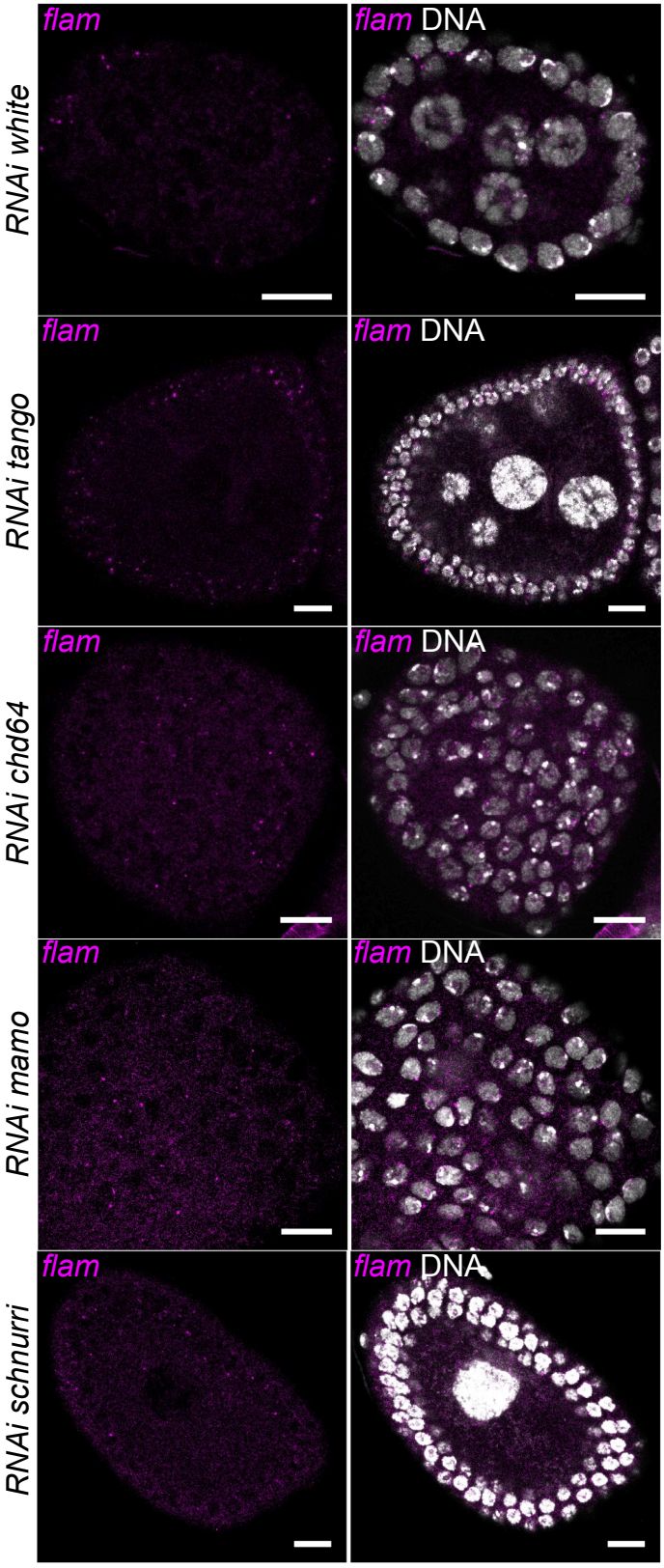
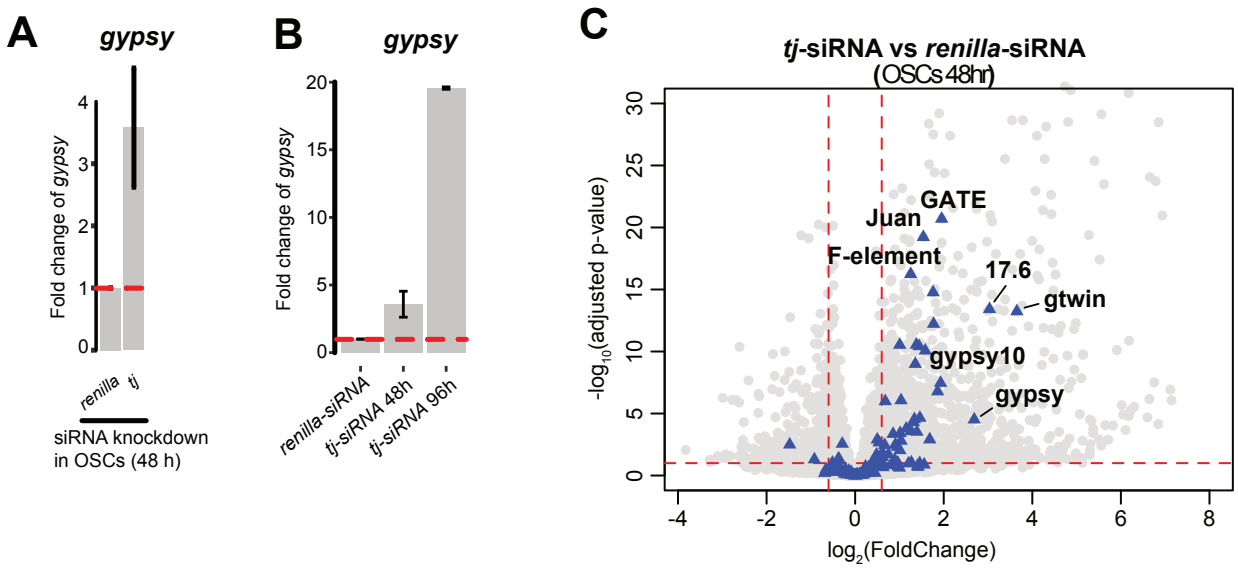
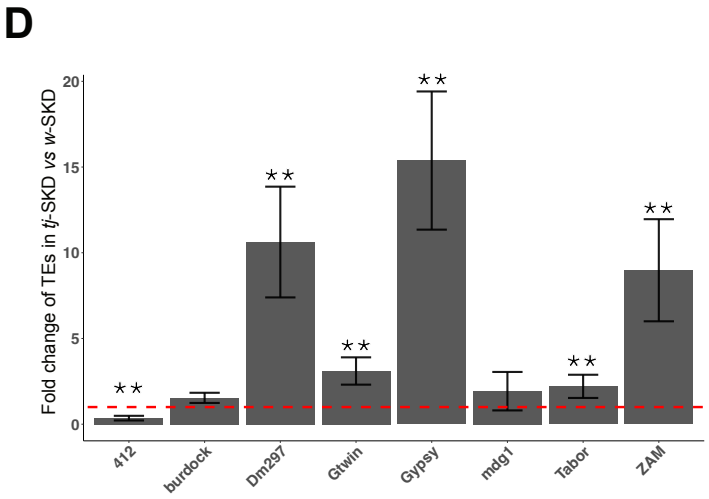

RNAi white
flam
flam DNA
RNAi tango
flam
flam DNA
RNAi chd64
flam
flam DNA
RNAi mamo
flam
flam DNA
RNAi schnurri
flam
flam DNA

Figure S6


A
gypsy
Fold change of gypsy
renilla
tj
siRNA knockdown in OSCs (48 h)
B
gypsy
Fold change of gypsy
renilla-siRNA
tj-siRNA 48h
tj-siRNA 96h
C
tj-siRNA vs renilla-siRNA (OSCs 48hr)
-log10(adjusted p-value)
log2(FoldChange)
GATE
Juan
F-element
17.6
gtwin
gypsy10
gypsy


D
Fold change of TEs in tj-SKD vs w-SKD
412
burdock
Dm297
Gtwin
Gypsy
mdg1
Tabor
ZAM

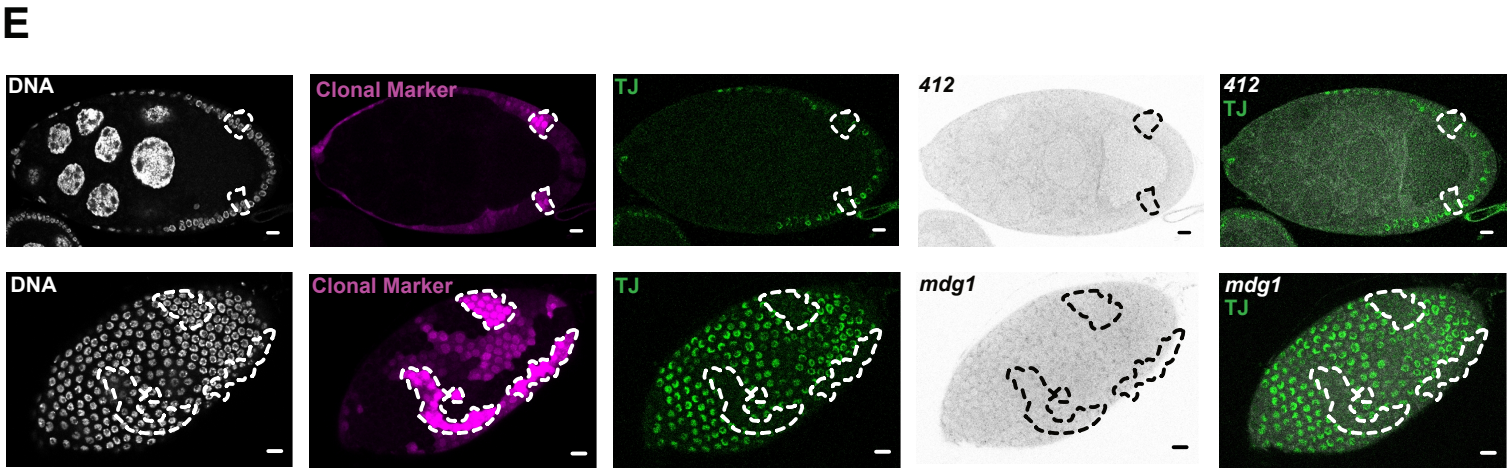
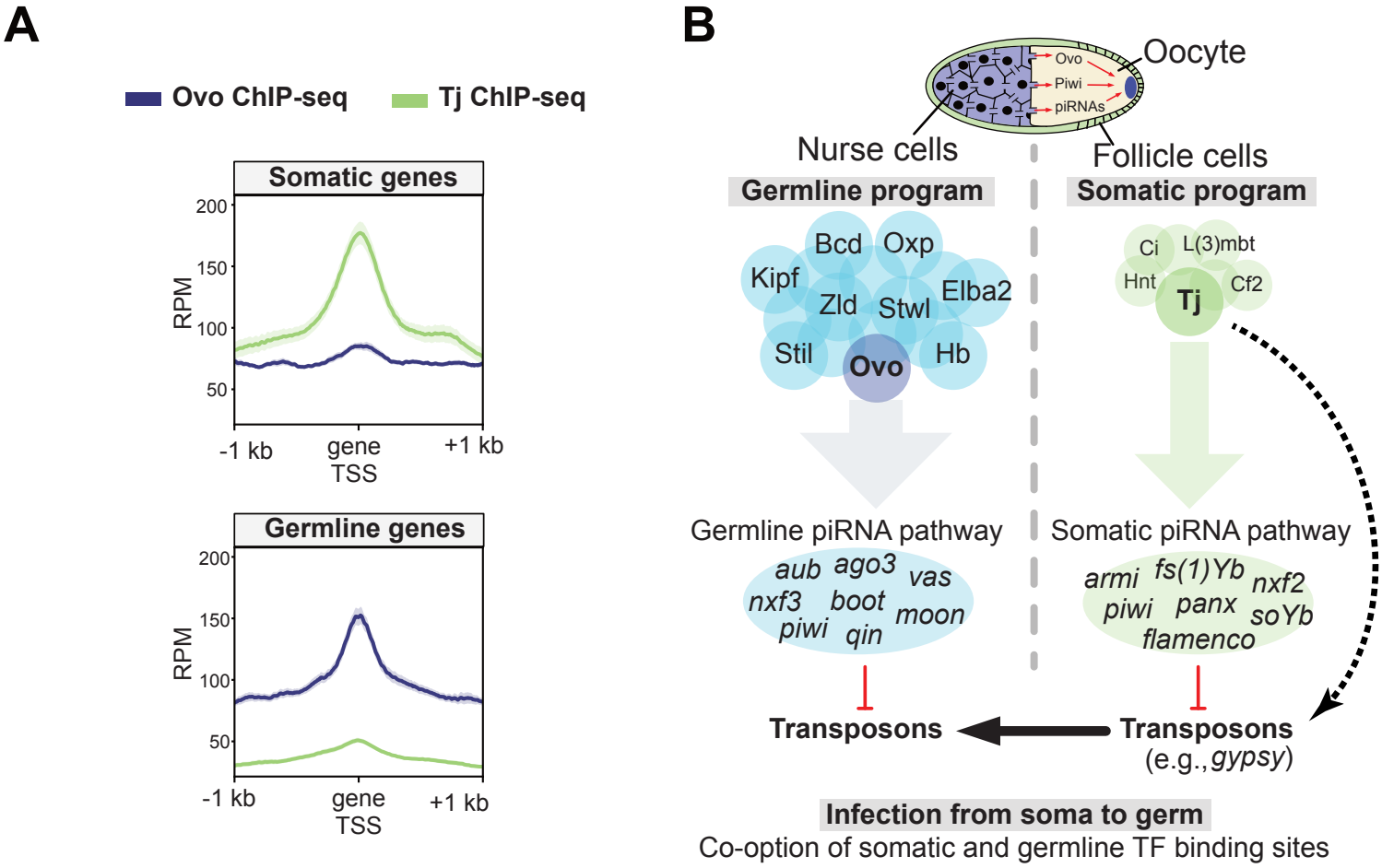

E
DNA
Clonal Marker
TJ
412
412
TJ
DNA
Clonal Marker
TJ
mdg1
mdg1
TJ

Figure S7


A
Ovo ChIP-seq
Tj ChIP-seq
Somatic genes
RPM
200
150
100
50
-1 kb
gene TSS
+1 kb
Germline genes
RPM
200
150
100
50
-1 kb
gene TSS
+1 kb
B
Oocyte
Ovo
Piwi
piRNAs
Nurse cells
Follicle cells
Germline program
Somatic program
Bcd
Oxp
Kipf
Zld
Stwl
Elba2
Stil
Ovo
Hb
Ci
L(3)mbt
Hnt
Tj
Cf2
Germline piRNA pathway
aub
ago3
vas
nxf3
boot
moon
piwi
qin
Somatic piRNA pathway
armi
fs(1)Yb
nxf2
piwi
panx
soYb
flamenco
Transposons
Transposons
(e.g., gypsy)
Infection from soma to germ
Co-option of somatic and germline TF binding sites