

Transcriptional regulation of the piRNA pathway by Ovo in animal ovarian germ cells

Azad Alizada, Gregory J. Hannon, and Benjamin Czech Nicholson

Cancer Research UK Cambridge Institute, Li Ka Shing Centre, University of Cambridge, Cambridge CB2 0RE, United Kingdom

The gene-regulatory mechanisms controlling the expression of the germline PIWI-interacting RNA (piRNA) pathway components within the gonads of metazoan species remain largely unexplored. In contrast to the male germline piRNA pathway, which in mice is known to be activated by the testis-specific transcription factor A-MYB, the nature of the ovary-specific gene-regulatory network driving the female germline piRNA pathway remains a mystery. Here, using *Drosophila* as a model, we combined multiple genomics approaches to reveal the transcription factor Ovo as regulator of the germline piRNA pathway in ovarian germ cells. Ectopic expression of Ovo in ovarian somatic cells activates germline piRNA pathway components, including the ping-pong factors Aubergine, Argonaute-3, and Vasa, leading to assembly of perinuclear cellular structures resembling nuage bodies of germ cells. We found that in ovarian somatic cells, transcription of *ovo* is repressed by *l(3)mbt*, thus preventing expression of germline piRNA pathway genes in the soma. Cross-species ChIP-seq and motif analyses demonstrate that Ovo is binding to genomic CCGTTA motifs within the promoters of germline piRNA pathway genes, suggesting a regulation by Ovo in ovaries analogous to that of A-MYB in testes. Our results also show consistent engagement of the Ovo transcription factor family at ovarian piRNA clusters across metazoan species, reflecting a deep evolutionary conservation of this regulatory paradigm from insects to humans.

[Keywords: piRNAs; oogenesis; germ cells; transposon silencing; transcriptional regulation]

Supplemental material is available for this article.

Received July 20, 2024; revised version accepted November 11, 2024.

The PIWI-interacting RNA (piRNA) pathway is an evolutionarily conserved defense system in metazoans that silences transposons in gonads, thereby serving crucial functions in genomic stability, gametogenesis, and fertility (Czech et al. 2018; Ozata et al. 2019). Mutants affecting the piRNA pathway negatively impact the integrity of the germline genome, disrupt gamete development, and typically result in sterility (Carmell et al. 2007; Chen et al. 2007; Klattenhoff et al. 2007; Pane et al. 2007). This host-parasite conflict has driven the piRNA pathway to evolve several specialized adaptations to silence transposons in gonadal cells (Malone et al. 2009). The *Drosophila* ovary is one of the key models that has been used to decipher the workings of the piRNA pathway (Czech et al. 2013; Handler et al. 2013; Muerdter et al. 2013). Within the *Drosophila* ovary, unlike somatic follicular cells, the germline nurse cells use a germline-specific version of the pathway encompassing nuage bodies, the ping-pong

cycle, and noncanonical transcription of dual-strand piRNA clusters (Czech et al. 2018; Ozata et al. 2019).

In the nuclei of *Drosophila* germ cells (Fig. 1A), transcription of piRNA precursors from dual-strand piRNA clusters is facilitated by Rhino (Rhi), a paralog of heterochromatin protein 1a that acts in complex with Deadlock (Del) and Cutoff (Cuff) (Klattenhoff et al. 2009; Mohn et al. 2014). Del interacts with Moonshiner (Moon), which recruits TBP-related factor 2 (Trf2) to enable dual-strand piRNA cluster transcription (Andersen et al. 2017). The resulting piRNA precursors are then exported out of the nucleus via a dedicated, noncanonical export machinery using nuclear export factor 3 (Nxf3), Bootlegger (Boot), UAP56, and Nxt1 (Zhang et al. 2012; Hur et al. 2016; ElMaghraby et al. 2019; Kneuss et al. 2019) and directed to germline-specific, perinuclear piRNA processing structures called nuage, where they are processed into functional piRNAs via the ping-pong cycle (Lim and Kai 2007).

In nuage bodies of *Drosophila* germ cells (Fig. 1A), ping-pong amplification operates via the PIWI proteins Argonaute-3 (Ago3) and Aubergine (Aub), which cleave piRNA

Corresponding authors: ben.nicholson@cruk.cam.ac.uk, greg.hannon@cruk.cam.ac.uk

Article published online ahead of print. Article and publication date are online at <http://www.genesdev.org/cgi/doi/10.1101/gad.352120.124>. Freely available online through the *Genes & Development* Open Access option.

© 2025 Alizada et al. This article, published in *Genes & Development*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

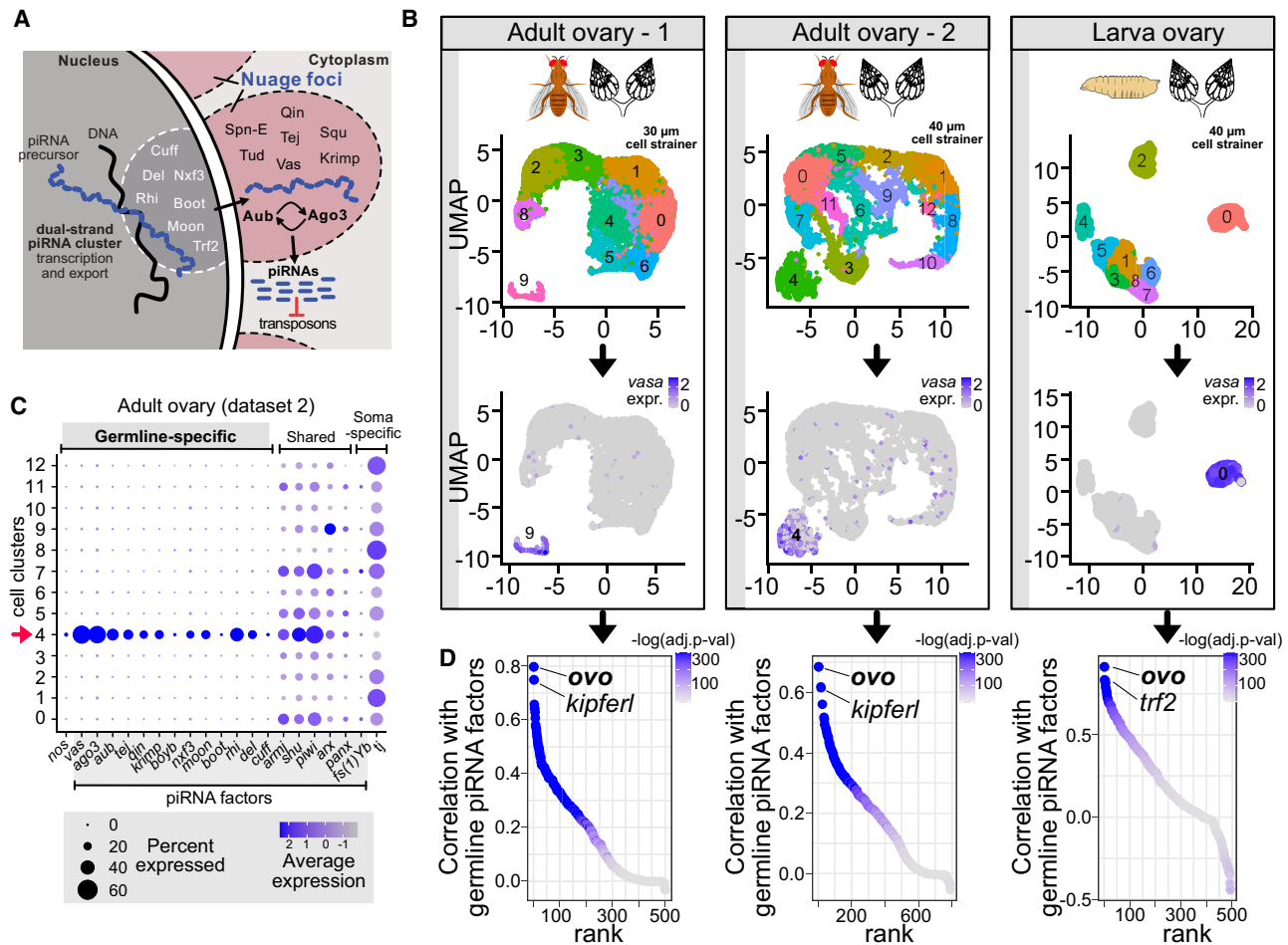


Figure 1. Ovo is the top transcription factor coexpressed with the germline piRNA pathway genes in the *Drosophila* ovary. (A) Model depicting the germline piRNA pathway in *Drosophila* germ cells. (B) UMAP clustering of *Drosophila* ovary single-cell RNA-seq data sets (adult ovary 1 from Rust et al. 2020; adult ovary 2 from Jevitt et al. 2020; LL3 larva ovary from Slaidina et al. 2020). Expression of the germline marker *vasa* is shown to mark germ cell clusters. (C) Dot plot showing expressions of the germline-specific, shared, and soma-specific piRNA pathway genes across the clusters identified in the adult ovary 2 data set (40 μ m cell strainer; clusters from A; cluster 4 is the germline cluster). (D) Ranking of DNA-binding proteins and transcription factors by their average expression correlation values (Pearson's r) with the expression of the germline piRNA pathway factors *aub*, *vas*, and *ago3*. The color scale indicates correlation P -values adjusted with Bonferroni correction for multiple testing. Also see Supplemental Data Set S1.

precursor and transposon transcripts in an alternating loop, hence shaping and amplifying the piRNA pool against active transposons (Brennecke et al. 2007; Gunawardane et al. 2007). Ago3 in complex with a sense piRNA recognizes and cleaves cluster transcripts through sequence complementarity, generating antisense pre-piRNAs that are loaded into Aub. Antisense piRNA-loaded Aub in turn detects and cleaves target transposon mRNAs, thereby forming a new Ago3 complex loaded with a sense piRNA and completing the cycle (Brennecke et al. 2007; Gunawardane et al. 2007). This process is assisted by several nuage-localized protein components, including the DEAD-box RNA helicase Vasa (Vas), the putative nuclease Squash (Squ), and Tudor domain proteins Tejas (Tej), Tapas, Qin, Krimper (Krimp), and Spindle-E (Spn-E) (Lim and Kai 2007; Pane et al. 2007; Patil and Kai 2010; Zhang et al. 2011; Patil et al. 2014; Sato et al. 2015; Webster et al. 2015; Ryazansky et al. 2016).

This is in sharp contrast to piRNA biogenesis in somatic cells, which lack nuage bodies and ping-pong amplification, and where piRNA precursors are transcribed from unistrand clusters (e.g., *flamenco*), exported to the cytoplasm by a canonical machinery, and processed into piRNAs via Zucchini-mediated biogenesis (Czech et al. 2018; Ozata et al. 2019).

The gene-regulatory network underlying the control of the germline-specific piRNA pathway in the ovaries of metazoan species have been largely unexplored. To date, clues from previous studies have pointed toward the existence of both positive and negative regulatory mechanisms. A study in mouse testes identified A-MYB as a transcriptional coordinator of piRNA factors such as *Piwi11* (*Miwi*) and germline pachytene piRNA clusters in male germ cells (Bolcun-Filas et al. 2011; Li et al. 2013), yet such regulatory pathways controlling the expression of the female germline piRNA pathway in the ovaries of

metazoan species, including *Drosophila*, are not known. Nonetheless, earlier work in *Drosophila* ovaries has suggested negative regulation of germline piRNA factors in somatic cells (Janic et al. 2010; Sumiyoshi et al. 2016; Coux et al. 2018). Here, germline-specific components of the piRNA pathway such as *aub*, *ago3*, and *vas* were up-regulated upon deletion of the tumor suppressor gene *lethal (3) malignant brain tumor [l(3)mbt]*; however, the mechanism underlying this regulation remains unclear. Several intriguing questions remain unanswered: What are the sex-specific differences in gene-regulatory networks underlying female and male germ cells? What are the mechanisms through which the coordinated transcription of germline piRNA factors and piRNA clusters are achieved? How does this vary between ovaries and testes, and what is their evolutionary conservation across the animal kingdom.

In this study, using *Drosophila* as a model, we uncovered key elements of the gene-regulatory network controlling the female germline piRNA pathway. In a systematic analysis integrating multiple approaches, including single-cell RNA-seq, ATAC-seq, and ChIP-seq, as well as depletion and overexpression screens, we identified the transcription factor Ovo as the key positive transcriptional regulator of the germline piRNA pathway in ovaries. Ectopic expression of Ovo in somatic cells activates expression of the germline piRNA pathway genes, including the ping-pong cycle components *Aub*, *Ago3*, and *Vas*, leading to formation of perinuclear cellular structures mimicking the nuage bodies of germ cells. This is orchestrated through binding of Ovo to conserved CCGTTA motifs within the promoters of these genes. We also reveal the mechanistic link between *l(3)mbt*, *ovo*, and the germline piRNA pathway genes. In addition, we show that Ovo-binding motifs are highly enriched within germline piRNA clusters in ovaries of metazoan species ranging from insects to humans. ChIP-seq experiments show that fly Ovo and its mouse and human orthologs, *OVOL2*, are recruited to the motifs within promoters of the piRNA pathway genes and ovarian piRNA clusters. Interestingly, the same CCGTTA motifs are also bound by the male-specific transcription factor A-MYB in testes of vertebrates to control the male germline piRNA pathway (Li et al. 2013). Our results indicate that Ovo is an ovary-specific counterpart of A-MYB, performing a role in ovaries analogous to that of A-MYB in vertebrate testes. Overall, our results reveal gene-regulatory interactions between the ovary-specific DNA-binding TFs and the CCGTTA *cis*-regulatory elements underlying the control of the germline piRNA pathway in ovaries across the metazoan species.

Results

Ovo is coexpressed with germline piRNA pathway genes in Drosophila ovary

Previous studies in mice identified the testis-specific transcription factor A-MYB as a regulator of male germline piRNA pathway genes (Bolcun-Filas et al. 2011; Li et al. 2013). To identify a potential master regulator of the fe-

male germline piRNA pathway in the *Drosophila* ovary, we performed a coexpression analysis using three separate single-cell RNA-seq data sets from *Drosophila* ovaries to determine the ovarian TFs coexpressed with the germline-specific piRNA pathway genes *aub*, *vas*, and *ago3* (Fig. 1B). Two of the scRNA-seq data sets were from adult ovaries (Jevitt et al. 2020; Rust et al. 2020), and one was from larval ovaries (Slaidina et al. 2020). Clustering of each of the data sets revealed distinct germline clusters showing specific expression of germline markers (e.g., *vas* and *nos*) and germline-specific piRNA pathway genes (i.e., *aub*, *vas*, *ago3*, *tej*, *qin*, *krimp*, *boyb*, *nxf3*, *moon*, *boot*, *rhi*, *del*, and *cuff*) (Fig. 1B,C; Supplemental Fig. S1A–D).

Correlation analysis (Pearson's r) identified *ovo* as the TF most significantly coexpressed with germline-specific piRNA pathway genes in all three data sets (Pearson's $r > 0.7$ in adult flies and Pearson's $r > 0.9$ in larva, adjusted $P < 1.0 \times 10^{-300}$) (Fig. 1D). The top coexpressed genes mostly represented germline-expressed and piRNA pathway-related genes (Supplemental Data Set S1). Interestingly, the second most significantly coexpressed TF in adult ovaries was *kipferl* (Fig. 1D), a recently identified zinc finger protein that interacts with Rhino and is required for piRNA production from most dual-strand piRNA clusters (Baumgartner et al. 2022). Of note, the second top-ranking TF in larval ovaries was *trf2* (Fig. 1D), which has been shown to assist Moon with transcription of dual-strand piRNA clusters (Andersen et al. 2017).

Drosophila ovaries are made up of ovarioles that comprise distinct stages of oogenesis, starting from the germline stem cells (GSCs) at the tip of germarium that produce cystoblasts, which give rise to nurse cells and the oocyte (Bastock and St Johnston 2008). Therefore, the germ cell cluster identified by scRNA-seq represents an aggregate of cells from distinct stages of germline development. To delineate the TFs coexpressed with germline piRNA pathway genes within germ cells across oogenesis stages, we isolated and reclustered the germ cell cluster and performed a correlation analysis specifically with the germ cells (Fig. 2; Supplemental Fig. S2).

Using early (*bgn* and *bam*) and late (*osk*) markers of oogenesis (Ohlstein et al. 2000; Chen and McKearin 2003; Snee et al. 2007), we computed a pseudotime of germline development and tracked the expression pattern of the germline piRNA pathway genes from GSCs to nurse cells (Fig. 2A,B; Supplemental Fig. S2A,B). Germline piRNA pathway genes showed expression in both the early and middle stages of oogenesis captured by scRNA-seq data. We then searched for TFs showing the closest match in expression pattern. This germ cell cluster-specific correlation analysis revealed *ovo* as the top-ranking TF coexpressed with the germline piRNA pathway genes over the course of germline development (adjusted $P < 3.4 \times 10^{-5}$) (Fig. 2C; Supplemental Fig. S2C).

To functionally validate the top coexpressed gene-regulatory candidates, we performed an overexpression screen in ovarian somatic cells (OSCs) (Fig. 2D). Ectopic expression of *ovo* (isoform *ovo-B*, NM_080338 transcript) resulted in upregulation of the germline piRNA pathway genes

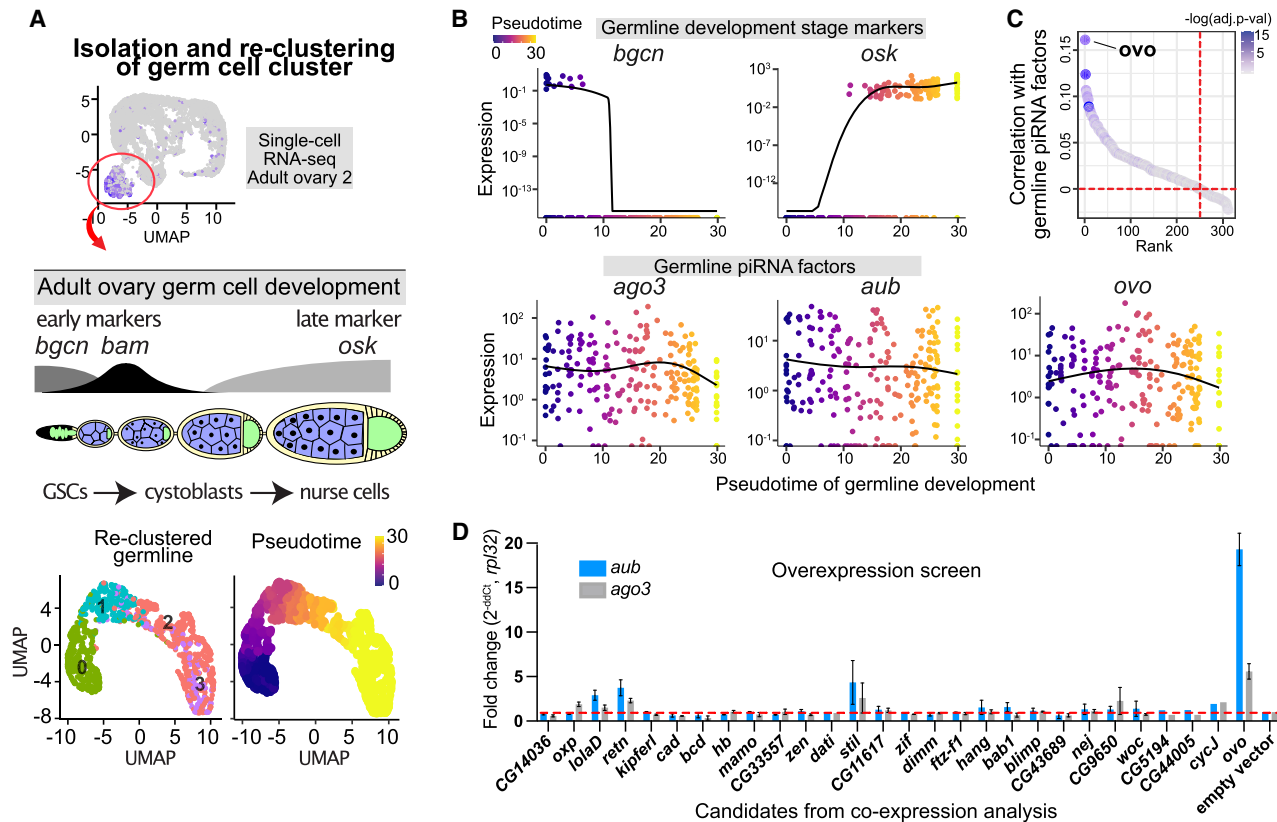


Figure 2. Overexpression screen of germline candidates coexpressed with piRNA pathway genes reveals Ovo as the positive regulator. (A) Diagram showing isolation and reclustering of the germ cells (cluster 4) from adult ovary single-cell RNA-seq (Supplemental Data Set S2) and computation of the pseudotime trajectory by rooting the *bgcn*-expressing germline stem cells (GSCs) as the starting point. (B) Expression pattern of the early (*bgcn*) and late (*osk*) stage markers of germline differentiation along the pseudotime trajectory of the germline development shown together with the germline piRNA pathway genes (*ago3* and *aub*) and the top coexpressed transcription factor *ovo*. (C) Ranking of the DNA-binding proteins and transcription factors by the average expression correlation (Pearson's *r*) with the germline piRNA pathway genes *aub*, *vas*, *qin*, and *ago3* within the reclustered germ cell cluster (cluster 4 in adult ovary scRNA-seq in Supplemental Data Set S2). Using the average of all germline piRNA pathway genes gave similar results. The color scale shows correlation *P*-values adjusted with Bonferroni correction for multiple testing. (D) Overexpression screen in ovarian somatic cells (OSCs) using the top coexpressed candidates (TFs and chromatin-binding proteins) from the coexpression analysis (RT-qPCR, 48–72 h after nucleofection in OSCs, *n* = 3 replicates from distinct samples; error bars indicate standard error of the mean). Overexpression of the Ovo-B isoform (matching NM_080338) is indicated for *ovo*. Also see Supplemental Data Set S1.

in OSCs (*aub* ~20-fold, *ago3* approximately fivefold, and *vas* approximately fivefold; $P < 0.01$, RT-qPCR), whereas other candidates had weaker or no effects on their expression (Fig. 2D).

Ovo is the top germline-enriched transcription factor in the *Drosophila* ovary

Differential expression analyses of germline and somatic cell clusters in the three separate ovarian single-cell RNA-seq data sets showed *ovo* as the most enriched germline TF (Supplemental Fig. S3A). To validate this finding, we sought to experimentally identify all germline-enriched TFs within *Drosophila* ovaries that could potentially be responsible for promoting germline fate along with the germline piRNA program.

To delineate germline-specific TFs, we first performed RNA-seq on FACS-sorted germline (*vas*-GFP⁺) and

somatic (*vas*-GFP⁻) cells from *vas*-GFP *Drosophila* ovaries (Fig. 3A). A differential RNA-seq analysis between the sorted germline and somatic cells (DESeq2) revealed *ovo* as the top germline-enriched TF (~87-fold, $P < 1.0 \times 10^{-300}$; *ovo*-B isoform matching NM_080338 transcript) (Fig. 3A). Cross-tissue RNA-seq analysis showed that ovaries are the most expressed site for *ovo* (Supplemental Fig. S3B). This result corroborates the hypothesis that *ovo* is the principal ovarian germline TF, potentially controlling a germline-specific expression program encompassing the piRNA pathway in ovarian germ cells. To confirm our findings, we performed RNA-seq on the ovarian germline/somatic coculture line (fGS/OSS) and compared it with purely somatic OSCs, which also revealed *ovo* as the most enriched germline-specific TF (~155-fold, $P < 1.0 \times 10^{-300}$) (Fig. 3B).

The germline–somatic differential RNA-seq analysis of sorted cells from the *vas*-GFP *Drosophila* ovaries showed

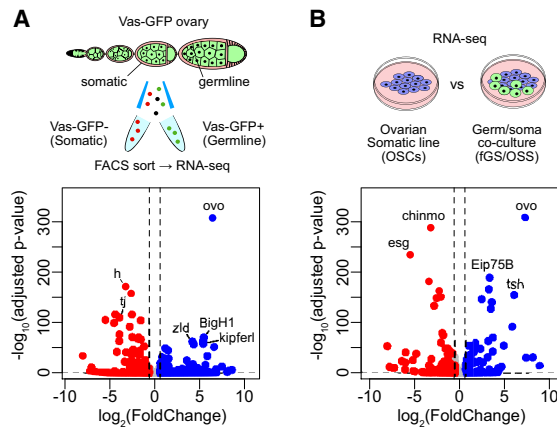


Figure 3. Ovo is the top enriched transcription factor in the *Drosophila* ovarian germ cells. (A) Differential gene expression of the DNA-binding TFs and chromatin proteins between the FACS-sorted *vas-GFP*⁺ (germline) and *vas-GFP*⁻ (somatic) cells from the transgenic *vas-GFP* fly ovaries (DESeq2; RNA-seq; *n* = 3 replicates from distinct samples). (B) Differential gene expression of the DNA-binding TFs and chromatin proteins between the ovarian somatic cells (OSCs) and the ovarian germ/soma coculture line (fGS/OSS) (DESeq2; RNA-seq; *n* = 4 replicates from distinct samples). Also see Supplemental Data Set S2.

a total of 1019 germline-enriched and 1316 soma-enriched genes ($\log_2FC > 1$ and $\log_2FC < -1$, respectively; *Padj* < 0.01; DESeq2). Among the piRNA pathway genes, only one, *fs(1)Yb*, had a soma-specific expression, whereas 13 components of the germline piRNA pathway (namely, *aub*, *tej*, *ago3*, *vas*, *qin*, *boYb*, *krimp*, *boot*, *moon*, *nxf3*, *cuff*, *rhi*, and *del*) showed strong germline-specific expression (Supplemental Data Set S2) and thus could potentially be driven by Ovo.

Ectopic ovo expression in OSCs activates germline piRNA pathway components

To determine the gene-regulatory effects of Ovo genome-wide, we performed RNA-seq in OSCs following ectopic *ovo* expression (overexpression vector carrying FLAG-tagged *ovo-B* isoform driven by the *act5C* promoter) and compared it with OSCs transfected with an empty vector. Western blotting confirmed the presence of the Ovo-B-FLAG protein at the expected size (~114 kDa) (Fig. 4A). Immunofluorescence showed that overexpressed Ovo-B-FLAG localized to the nucleus (Fig. 4B; Supplemental Fig. S4A). Differential RNA-seq analysis (DESeq2) revealed that ectopic Ovo expression significantly upregulated ~70% of the germline-specific piRNA pathway genes in OSCs; namely, *ago3*, *aub*, *tej*, *qin*, *vas*, *moon*, *boot*, and *nxf3* (>1.5-fold, adjusted *P* < 0.05) (Fig. 4C; Supplemental Fig. S4B,C; Supplemental Data Set S2). The strongest effects were observed for *aub* and *ago3* (16-fold and sixfold, adjusted *P* < 1.0×10^{-41}) (Fig. 4C; Supplemental Fig. S4B,C). In contrast to components of the germline piRNA pathway, the soma-specific *fs(1)Yb* was downregulated upon Ovo overexpression, with no change in expression in all other somatic factors involved in biogenesis and

transcriptional gene silencing (Fig. 4C; Supplemental Fig. S4C).

Using immunofluorescence, we observed the appearance of *Aub*, *Ago3*, and *Vas* proteins following ectopic Ovo expression in OSCs that assembled into perinuclear “nuage”-like structures (Fig. 4D–F; Supplemental Fig. S4D). Colocalization analysis showed that *Aub*, *Ago3*, and *Vas* proteins assembled as foci around nuclei in a ring-shaped manner (Fig. 4E,F), resembling the nuage structures of the *Drosophila* germ cells where piRNAs are processed by the ping-pong cycle. Of note, we confirmed that these “nuage”-like structures were distinct from the somatic Yb bodies of OSCs (Supplemental Fig. S4E).

A total of 693 genes was upregulated by ectopic Ovo expression in OSCs ($\log_2FC > 0.6$, *Padj* < 0.1; DESeq2), with 182 (~26%) showing strong germline-enriched expression in ovaries (*vas-GFP* $\log_2FC > 1$, *Padj* < 0.01) (Supplemental Data Set S2). Ovo could account for the regulation of at least 18% of all strong germline-enriched genes in ovaries (182 out of 1019 genes). Ovo target genes in OSCs represented ~70% (nine out of 13) of the germline-specific piRNA pathway genes and 16.8% of the germline-enriched genes. Notably, among the top germline Ovo targets was *nanos* (*nos*) (Supplemental Fig. S4F), a gene essential for germline formation and maintenance (Kobayashi et al. 1996; Forbes and Lehmann 1998). Therefore, our results suggest that Ovo controls germline piRNA pathway factors in addition to a more general role in ovarian germline development. Next, we sought to decipher mechanisms of Ovo regulation to understand how it activates piRNA factor expression exclusively in germ cells.

L(3)mbt suppresses the germline piRNA pathway by repressing *ovo* in ovarian somatic cells

Early clues regarding the regulation of germline piRNA pathway components in *Drosophila* came from studies that deleted the *l(3)mbt* gene in somatic tissue, including the brain, ovary, and ovary-derived OSCs, which uniformly resulted in the upregulation of several germline-specific genes such as the ping-pong components *Aub*, *Vas*, and *Ago3* (Janic et al. 2010; Sumiyoshi et al. 2016; Coux et al. 2018). The molecular mechanism underlying this activation has remained unknown. *L(3)mbt* has been shown to interact with chromatin and cause histone compaction, leading to suppression and insulation of gene expression (Trojer et al. 2007; Richter et al. 2011; Blanchard et al. 2014). Although a direct regulation by *L(3)mbt* could explain prior observations, we hypothesized that *L(3)mbt* deletion could also upregulate one or several specific TFs that in turn control the expression of germline piRNA pathway genes.

To determine the precise mechanistic link between *L(3)mbt* and germline piRNA pathway components, we performed RNA-seq (*n* = 3 replicates) and ATAC-seq (*n* = 2 replicates) on wild-type OSCs and OSCs carrying a deletion of *l(3)mbt*, referred to as $\Delta l(3)mbt$ (Fig. 5A; Sumiyoshi et al. 2016). We aimed to uncover regulatory elements, transcription factors, and genes that are differentially

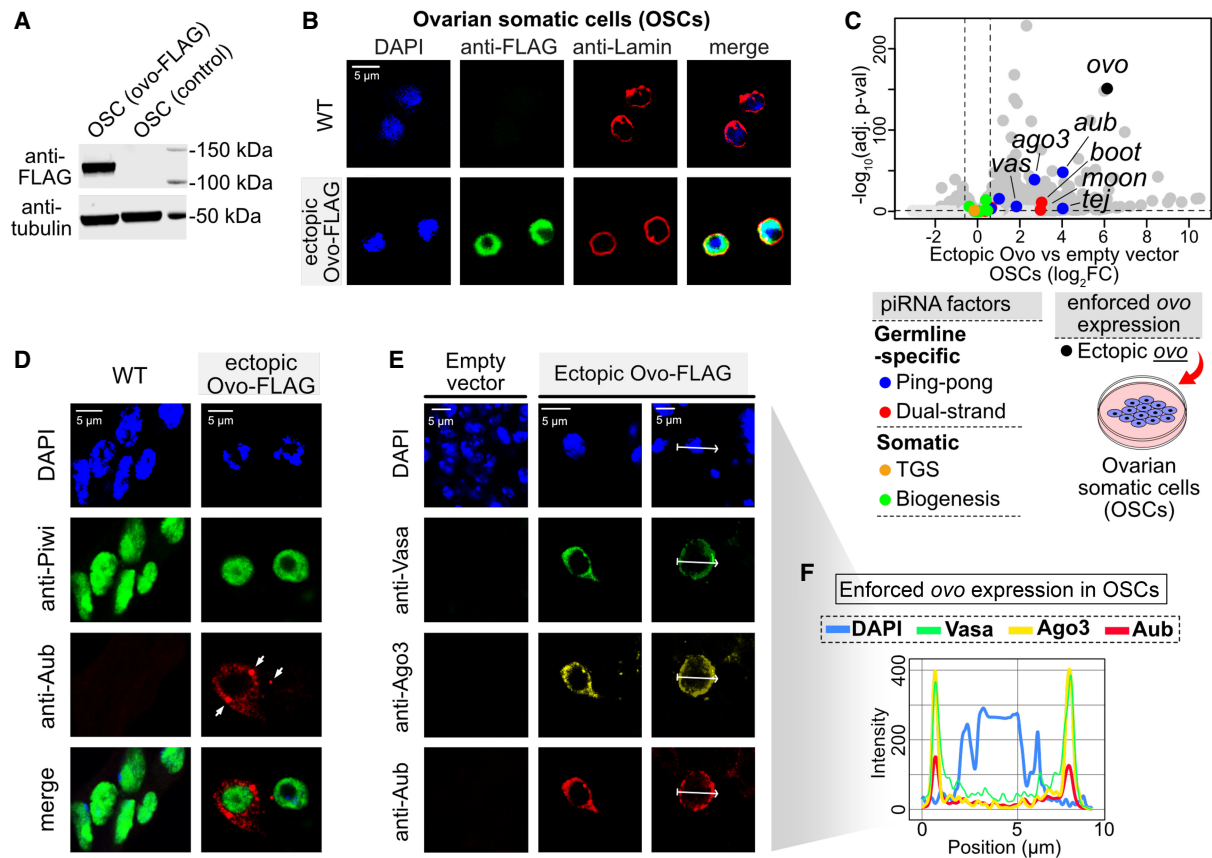


Figure 4. Ectopic expression of *ovo* in ovarian somatic cells (OSCs) activates the germline piRNA pathway components, leading to formation of cellular structures resembling nuage bodies of germ cells. (A) Western blot showing the presence of the Ovo-FLAG protein following ectopic expression in OSCs. (B) Immunofluorescence images showing the nuclear localization of the Ovo-FLAG protein in OSCs after nucleofection with the *ovo-FLAG* construct (48 h; Ovo-B isoform, NM_080338 transcript). (Blue) DAPI, (red) Lamin. (C) Differential gene expression between *ovo-FLAG* nucleofected OSCs relative to empty vector (DESeq2; RNA-seq $n = 3$ replicates from distinct samples; the piRNA pathway genes are labeled according to the color key). (TGS) Transcriptional gene silencing. (D) Immunofluorescence images showing the appearance of Aub, Ago3, and Vas as perinuclear nuage-like structures and foci (arrowheads) in the *ovo-FLAG* nucleofected OSCs. (Blue) DAPI, (green) Piwi, (red) Aub. (E) As in D but showing Vasa (green), Ago3 (yellow), and Aub (red). (F) Colocalization of Aub, Ago3, and Vas proteins within nuage-like bodies formed around the nuclei (DAPI) of the *ovo-FLAG* nucleofected OSCs. The fluorescence intensity along the white arrows in E is normalized to the highest value.

regulated and could be responsible for the activation of the germline piRNA pathway components upon *L(3)mbt* deletion.

Differential expression analysis of DNA-binding TFs (DESeq2 on RNA-seq data) between $\Delta l(3)mbt$ and wild-type OSCs revealed *ovo* as the top upregulated TF upon *l(3)mbt* deletion (~ 97 -fold, adjusted P value $< 1.0 \times 10^{-300}$) (Fig. 5A). The degree of *ovo* upregulation was markedly higher compared with other TFs (Fig. 5A). To assess whether this expression difference was due to increased accessibility at the *ovo* promoter upon *l(3)mbt* loss, we analyzed differential chromatin accessibility using ATAC-seq data (DiffBind) from $\Delta l(3)mbt$ and wild-type OSCs (Fig. 5B–D). This analysis uncovered 207 genomic regions with significantly increased accessibility upon *l(3)mbt* deletion (> 1.5 -fold, FDR $q < 0.05$), of which 111 were at gene promoters (± 1 kb of TSS) (Fig. 5B,C). Out of these, 30 were germline-specific and included only one germline TF, *ovo* (approximately twofold, FDR $q < 0.01$) (Fig. 5D).

The only other TF that showed increased promoter accessibility in $\Delta l(3)mbt$ -OSCs was *bigmax*, which was not germline-enriched in ovaries. Among piRNA pathway genes, only *vas* showed increased promoter accessibility in $\Delta l(3)mbt$ -OSCs and thus could potentially be a piRNA factor that is directly regulated by *L(3)mbt*. This is also supported by the *L(3)mbt* ChIP-seq binding signal at the *vas* promoter (± 0.5 kb) (Fig. 5H) and disproportionately stronger upregulation of *vas* in *l(3)mbt* mutants compared with other piRNA factors or compared with ectopic Ovo alone in wild-type OSCs (Supplemental Data Set S2). Therefore, upregulation of *vas* in *l(3)mbt* knockouts and knockdowns is unlikely to be attributed to activated Ovo alone and is likely due to the lack of *l(3)mbt* repression over the *vas* promoter followed by activation by Ovo.

We speculated that direct binding of *L(3)mbt* to the *ovo* promoter could be responsible for the decreased chromatin accessibility in somatic cells, leading to its somatic-specific transcriptional repression. Indeed, ChIP-seq data

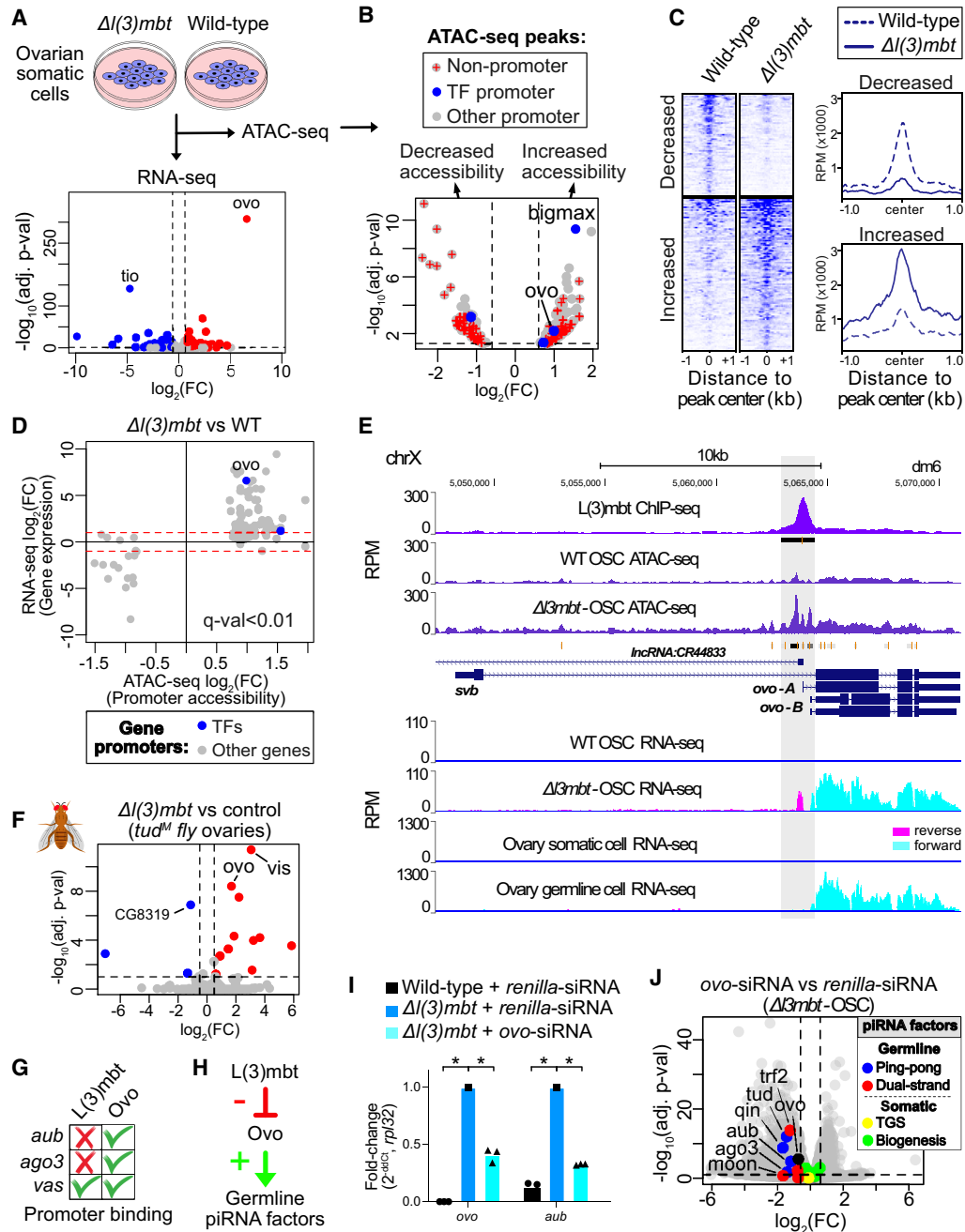


Figure 5. L(3)mbt negatively regulates the germline piRNA pathway via repression of *ovo* expression in somatic cells. (A) Differential gene expression of the DNA-binding TFs between the wild-type ovarian somatic cells (OSCs) and $\Delta l(3)mbt$ OSCs (DESeq2; RNA-seq; $n = 4$ replicates from distinct samples). (B) Differential chromatin accessibility between the wild-type OSCs and $\Delta l(3)mbt$ OSCs (DiffBind; ATAC-seq; $n = 2$ replicates from distinct samples). Gene promoters are defined as ATAC-seq peaks overlapping ± 1 kb of the transcription start site (TSS). (C) Heat maps and genomic profiles of ATAC-seq reads corresponding to regions of increased and decreased chromatin accessibility (FDR $q < 0.05$; DiffBind). (D) All genes showing significant differential gene expression (RNA-seq; DESeq2; adjusted P -value < 0.01) and promoter chromatin accessibility (ATAC-seq; ± 0.5 kb of gene TSS; DiffBind, FDR $q < 0.01$) between wild-type and $\Delta l(3)mbt$ OSCs. (E) Increased ATAC-seq peak accessibility (identified using DiffBind) at the *ovo* promoter region in $\Delta l(3)mbt$ OSCs ($n = 2$ replicates from distinct samples; merged) (our data) and L(3)mbt ChIP-seq from OSCs showing L(3)mbt binding at the region ($n = 2$ replicates from distinct samples; merged) (data from Yamamoto-Matsuda et al. 2022). RNA-seq tracks below show the expression of the germline *ovo* isoform in $\Delta l(3)mbt$ OSCs and *Drosophila* ovary germline cells ($n = 4$ replicates from distinct samples; merged) (our data). (F) Differential gene expression (DESeq2) between control and $\Delta l(3)mbt$ fly ovaries, both lacking germline due to *tud* maternal mutations (*tud^M*; $n = 3$ replicates from distinct samples) (RNA-seq data from Coux et al. 2018). (G) Table summarizing L(3)mbt and Ovo ChIP-seq peaks within the promoters of the ping-pong genes *aub*, *ago3*, and *vas* (± 1 kb of TSS). (H) Model for the L(3)mbt and Ovo regulation of the germline piRNA pathway genes. (I) Ovo siRNA knockdown experiments in $\Delta l(3)mbt$ OSCs (RT-qPCR; $n = 3$ replicates from distinct samples). (*) P -value: < 0.01 , one-tailed two-sample t -test. (J) Volcano plot showing downregulation of the germline-specific piRNA pathway genes on day 2 of *ovo* siRNA knockdowns in $\Delta l(3)mbt$ OSCs using differential RNA-seq analysis (DESeq2) between *ovo* and *renilla* siRNA knockdowns ($n = 3$ replicates from distinct samples).

from OSCs showed strong L(3)mbt binding specifically at the promoter region of *ovo* (Fig. 5E), which increased in accessibility upon *l(3)mbt* deletion. This suggests that direct L(3)mbt binding to the *ovo* promoter in somatic cells represses its expression. Elimination of *l(3)mbt* consequently leads to increased accessibility at the *ovo* promoter and results in drastically increased expression of *ovo*, mimicking its germline expression in the *Drosophila* ovary (Fig. 5E). We further confirmed our results by reanalyzing recently published RNA-seq data from *l(3)mbt* knockdown in OSCs (Yamamoto-Matsuda et al. 2022), which also showed *ovo* as the most significantly upregulated gene (Supplemental Fig. S5A).

To extend our findings to an in vivo setting, we analyzed previously published RNA-seq data from Lehmann and colleagues (Coux et al. 2018) from *l(3)mbt* mutant and control ovaries of transgenic flies that lack germ cells due to maternal *tud* mutations (*tud^M*). This allowed us to perform differential expression analysis between $\Delta l(3)mbt$ and control ovaries that were devoid of germline cells (Fig. 5F). This analysis identified *ovo* as the second most significantly upregulated TF upon *l(3)mbt* deletion in vivo within somatic cells of fly ovaries (Fig. 5F). It was previously shown that L(3)mbt forms a complex with Lint-O at chromatin to silence gene expression (Yamamoto-Matsuda et al. 2022). Differential expression analysis of RNA-seq data from *lint-O* and control knockdowns in OSCs revealed *ovo* as the top upregulated TF upon Lint-O depletion (Supplemental Fig. S5B), suggesting that Lint-O, together with L(3)mbt, forms a repressive complex at the *ovo* promoter in somatic cells, which we confirmed using Lint-O ChIP-seq data in OSCs (Supplemental Fig. S5C).

We then analyzed whether L(3)mbt could be exerting its effects on germline piRNA pathway genes, such as *aub*, indirectly via regulation of Ovo. Most importantly, using Ovo ChIP-seq, we observed strong Ovo binding at the promoters of germline-specific piRNA pathway genes where L(3)mbt binding was lacking (Fig. 5G; Supplemental Fig. S5F), suggesting that L(3)mbt was not directly controlling their expression. We therefore hypothesized that L(3)mbt indirectly represses the expression of germline-specific piRNA pathway genes in somatic cells via its regulation of *ovo* expression (Fig. 5H).

To establish a causal link, we performed siRNA knockdowns of *ovo* in $\Delta l(3)mbt$ OSCs, which express *ovo* due to the absence of L(3)mbt repression (Fig. 5I). Knockdown of *ovo* in $\Delta l(3)mbt$ OSCs resulted in the downregulation of germline piRNA pathway genes, including *aub* and *ago3* ($P < 0.01$, RT-qPCR), further supporting our hypothesis (Fig. 5I; Supplemental Fig. S5D). To validate our finding, we additionally performed RNA-seq on *ovo* siRNA knockdowns in $\Delta l(3)mbt$ OSCs (Fig. 5J; Supplemental Fig. S5E), which confirmed decreased expression of germline piRNA pathway genes, including *aub* (~2.3-fold decrease, $Padj = 2.8 \times 10^{-6}$; DESeq2), *qin* (approximately threefold decrease, $Padj = 2.6 \times 10^{-8}$), and *ago3* (approximately threefold decrease, $Padj = 0.01$) (Fig. 5J; Supplemental Fig. S5E).

Overall, in our analysis, we observed that ~25% ($n = 275$) of all upregulated genes in $\Delta l(3)mbt$ OSCs ($n = 1085$) were germline-enriched, representing ~16% of all germline-

enriched genes ($n = 1700$; $\log_2FC > 0.6$, $Padj < 0.05$). These genes could be either under a direct L(3)mbt repression [i.e., L(3)mbt promoter binding] or under an indirect repression via germline TFs such as Ovo that are repressed by L(3)mbt (Supplemental Data Set S2). Of the 315 L(3)mbt-repressed germline genes, 103 were upregulated by ectopic *ovo* expression alone in OSCs; therefore, ~33% of the upregulated germline genes in $\Delta l(3)mbt$ OSCs were likely to be indirectly regulated through activation of *ovo* expression in the absence of *l(3)mbt* (Supplemental Data Set S2). RNA-seq of *ovo* siRNA knockdowns in $\Delta l(3)mbt$ OSCs revealed that ~35% ($n = 419$ genes) of all L(3)mbt-repressed genes in OSCs ($n = 1195$) were downregulated by day 2 of *ovo* siRNA knockdown and thus could be accounted for by indirect regulation via Ovo (with ~27%, 112 out of the 419 genes, being germline-enriched) (Supplemental Data Set S2). Out of a total of 2395 downregulated genes by day 2 of *ovo* siRNA treatment, 488 were germline-enriched, including germline piRNA pathway genes. Thus, we could conclude that at least 20% of all upregulated germline genes in $\Delta l(3)mbt$ OSCs were upregulated indirectly due to activation of Ovo. This fraction captured ~62% of all germline-specific piRNA pathway components (e.g., *aub*, *ago3*, *tej*, *boot*, and *qin*). This proportion of upregulated germline genes remained at ~20% (485 out of 2394) by day 4 of *ovo* siRNA treatment (second nucleofection on day 2) (Supplemental Data Set S2; Supplemental Fig. S5E).

Ovo regulates the germline piRNA pathway genes via binding to conserved CNGTTA motifs

To capture Ovo binding events that lead to activation of the germline piRNA pathway genes, we performed Ovo ChIP-seq ($n = 2$) following ectopic *ovo* expression in OSCs ($n = 4769$ peaks, $FDR < 0.05$, more than fivefold enrichment). Using Ovo ChIP-seq data from adult female flies (ENCODE) (The modENCODE Consortium et al. 2010) we were also able to map Ovo binding events ($n = 4477$ peaks) across the *Drosophila* genome (Fig. 6A). We performed ATAC-seq in *Drosophila* ovaries to locate Ovo binding regions corresponding to open chromatin ($n = 4007$) and closed chromatin ($n = 470$) within ovaries (Fig. 6A). Ectopic Ovo expression in OSCs recapitulated ~65% of the in vivo binding events observed in flies, of which the majority (~93%) were within open chromatin in ovaries (Fig. 6A). Interestingly, Ovo binding sites at ~72% of the germline piRNA pathway gene promoters in OSCs were restricted in chromatin accessibility before addition of Ovo (e.g., promoter regions of *aub*, *vas*, *qin*, *ago3*, *tej*, *nxf3*, and *boot*), whereas others, such as the promoter regions of *squ* and *piwi*, were already accessible. This could point to a pioneering factor-like ability of Ovo to bind motifs in closed chromatin configurations and potentially remodel and increase their accessibility, though this hypothesis requires further testing.

The top-scoring de novo motif within in vivo Ovo ChIP-seq peaks in flies was CCGTTA (MEME-ChIP, e -value = 1.3×10^{-6}), while the secondary motif was CNGTTA (MEME-ChIP, e -value = 8.3×10^{-68}) (Fig. 6B). As expected, the

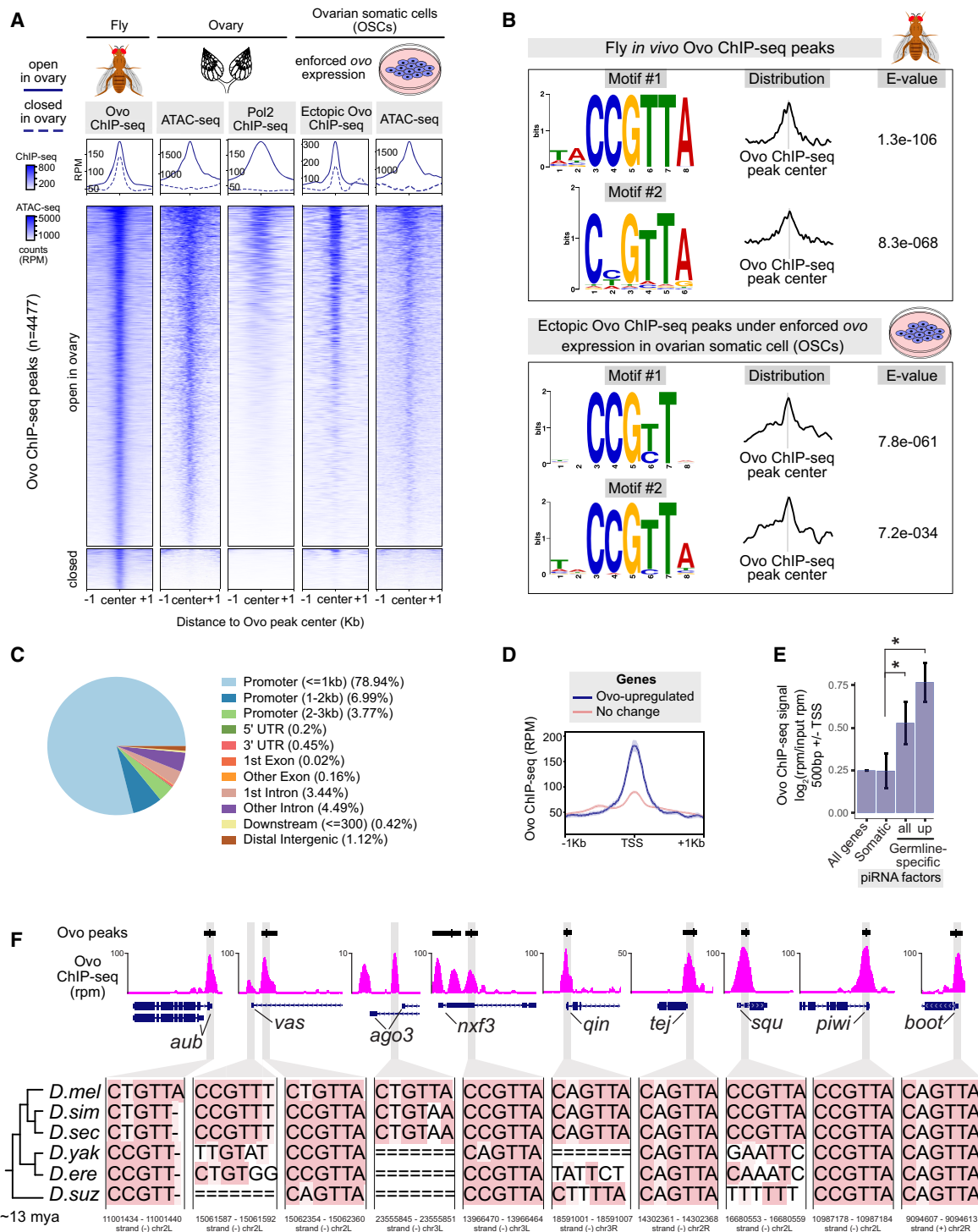


Figure 6. Ovo regulates germline piRNA pathway genes via binding to conserved CNGTTA motifs. (A) Heat map of fly Ovo ChIP-seq peaks ($n=2$ replicates from distinct samples merged; ENCODE) corresponding to open and closed chromatin states in fly ovaries based on ovary ATAC-seq peaks ($n=2$ replicates from distinct samples merged). The binding pattern of the ectopic Ovo ChIP-seq following ectopic *ovo* expression in OSCs ($n=2$ replicates from distinct samples) and ATAC-seq accessibility of OSCs ($n=2$ replicates from distinct samples) centered on the fly Ovo ChIP-seq peaks are shown by the heat maps at the right. The engagement of RNA polymerase II (Pol2) in ovaries is also depicted using the ovary Pol2 ChIP-seq signals ($n=1$) (data from Andersen et al. 2017). (B) The top-scoring de novo motifs discovered within Ovo ChIP-seq peaks using MEME-ChIP. (C) Genomic annotations (UCSC; Ensembl genes; dm6) of the fly Ovo ChIP-seq peaks using ChIPseeker. (D) A genomic profile of fly Ovo ChIP-seq signals (RPM) within ± 1 kb of the TSS of the upregulated, downregulated, and unresponsive genes following ectopic *ovo* expression in OSCs is shown. (E) Bar plots comparing fly Ovo ChIP-seq signals (input-normalized; RPM) between the promoters of the germline-specific and the somatic piRNA pathway genes. (F) The conservation of Ovo motifs within Ovo ChIP-seq peak summits at the germline piRNA pathway gene promoters is shown by multiple sequence alignments across six *Drosophila* species (27 way insect Multiz alignments and the phylogenetic tree from the UCSC genome browser conservation tracks; Ovo motif coordinates for the dm6 genome are shown below).

CCGTT core of the motif was the top hit within the ectopic Ovo ChIP-seq peaks following *ovo* overexpression in OSCs (MEME-ChIP, e -value = 7.8^{-061}) (Fig. 6B), and CNGTT was the fourth strongest hit in vivo (Supplemental Fig. S6C). These results suggest that CCGTTA is the most preferred Ovo binding site in vivo, with changes to the second and last nucleotides often tolerated. The majority of Ovo binding events (~79%) were proximal to promoters (<1 kb to the TSS), while only 1.12% were distal or intergenic (Fig. 6C). Approximately 95% of Ovo binding events at promoter-proximal regions corresponded to open chromatin in ovaries (based on ATAC-seq peaks) (Supplemental Fig. S6A,B). The genes upregulated in response to ectopic *ovo* expression in OSCs showed more than twofold stronger Ovo ChIP-seq binding signals and peak enrichments in their promoters (± 1 kb of TSS, $P < 0.01$) compared with the genes that did not change in expression (Fig. 6D; Supplemental Fig. S6D). Correspondingly, higher numbers of Ovo motifs were observed at upregulated Ovo target gene promoters when compared with genes that did not change in expression (~1.6-fold, $P < 0.01$) (Supplemental Fig. S6E,F).

Transcriptional activation by Ovo-B has been reported previously to occur via direct promoter binding at target genes (Bielinska et al. 2005; Benner et al. 2024). We therefore analyzed Ovo binding within promoters of the piRNA pathway genes (Fig. 6E,F). The average Ovo binding signal (input-normalized ChIP-seq) was more than twofold stronger ($P < 0.05$) at promoters of the germline-specific piRNA pathway genes when compared with somatic or shared piRNA factors (Fig. 6E). This was generally true for all germline genes, which showed on average approximately twofold stronger Ovo ChIP-seq signal ($P < 4.1 \times 1.0^{-8}$) at their promoters (± 0.5 kb of TSS) when compared with the somatic or shared genes. We further validated these results by reanalyzing recently published Ovo-GFP ChIP-seq in ovaries (Benner et al. 2024), which also showed the strongest Ovo binding signals at germline piRNA pathway genes (Supplemental Fig. S6G).

Moreover, ~90% of the germline-specific piRNA pathway genes and ~30% of the germline-somatic shared piRNA pathway genes had an Ovo motif ± 0.5 kb of their TSS overlapping an Ovo ChIP-seq peak summit (Supplemental Fig. S4C) compared with ~22% of all germline-enriched genes. Multiple sequence alignments of Ovo binding sites within promoters of germline piRNA pathway genes showed high conservation of the CNGTTA motifs across six *Drosophila* species (*D. melanogaster*, *D. simulans*, *D. sechellia*, *D. yakuba*, *D. erecta*, and *D. suzuki*) sharing a common ancestor ~13 million years ago (Suvorov et al. 2022), suggesting an evolutionarily conserved gene-regulatory mechanism behind Ovo interaction with CNGTTA motifs (Fig. 6F). These results indicate that direct binding to CNGTTA motifs is a major conserved mechanism through which Ovo controls expression of germline piRNA pathway genes in the *Drosophila* species.

Vertebrate homologs of Ovo bind to ovarian piRNA pathway components

Most vertebrates have three homologs of the fly *ovo* gene (e.g., mouse *Ovol1*, *Ovol2*, and *Ovol3*), with the *Ovol2*

paralog involved in the development of primordial germ cells (PGCs) (Hayashi et al. 2017; Naitou et al. 2022). Clustering of all vertebrate TF motifs in the JASPAR database (2022 vertebrate CORE) revealed high similarities between the motifs of Ovo family TFs and, remarkably, A-MYB (Fig. 7A), a testis-specific TF previously reported to control transcription of both the germline piRNA pathway genes and piRNA clusters in male mice. We hypothesized that Ovo homologs could provide the ovary-specific equivalent of this regulation in females, controlling germline piRNA clusters in ovaries in addition to piRNA pathway genes via the same CCGTTA motifs.

To test this hypothesis, we used mouse OVOL2 ChIP-seq data from female mouse primordial germ cell-like cells (PGCLCs; day 2 of induction) under transgenic expression of mouse *Ovol2* (Naitou et al. 2022). We observed OVOL2 binding to CCGTTA motifs within or near promoters of the piRNA pathway genes (e.g., *Tdrkh*, *Tdrd7*, and *Pld6*) (Fig. 7B; Supplemental Fig. S7B), which also showed upregulation in response to *Ovol2* expression in mouse ESCs or PGCLCs (Supplemental Fig. S7B,C). To find out whether mouse OVOL2 also binds to ovary piRNA clusters, we defined ovary-specific, testis-specific, and shared piRNA clusters using small RNA-seq data from mouse ovaries and testes (more than one RPKM-normalized count mapping to the piRNA cluster regions [data from Aravin et al. 2008]; coordinates were obtained via proTRAC [Rosenkranz and Zischler 2012]) and checked OVOL2 binding at these regions (Fig. 7C; Supplemental Fig. S7A). Our analysis revealed strong OVOL2 binding events to CCGTTA motifs within regions corresponding to ovary-specific piRNA clusters (Fig. 7C; Supplemental Fig. S7A). These OVOL2 binding events within motifs were found near the ends or the center of ovary piRNA clusters (Fig. 7C; Supplemental Fig. S7A). OVOL2 ChIP-seq peaks showed a significant enrichment at ovarian piRNA clusters but were absent from testis-specific piRNA clusters (Fig. 7C; Supplemental Fig. S7A,D).

Next, we asked whether there was evidence of human OVOL2 engagement at human ovarian piRNA clusters. We exploited small RNA-seq data from human fetal ovaries (data from Williams et al. 2015) and analyzed OVOL2 motif occurrences within human ovarian piRNA clusters (Fig. 7D). Ranking of human ovary piRNA clusters with small RNA-seq expression using data and coordinates defined in Figure 7D and Supplemental Figure S7E and by Williams et al. (2015) revealed that the most highly expressed piRNA cluster in human fetal ovaries was cluster 71, which accounted for 54% of all piRNA reads in ovaries (Fig. 7D; Supplemental Fig. S7E). piRNA cluster 71 showed 2.5-fold enrichment for OVOL2 motifs over the genomic background ($P < 0.01$) (Fig. 7D). The enriched OVOL2 motifs were highly concentrated near the 5' end of cluster 71, suggesting a potential impact on promoter activity (Fig. 7D). To determine whether OVOL2 was physically associated with these motifs, we used OVOL2 ChIP-seq data from human induced pluripotent stem cells (iPSCs; WTC11) (data from ENCODE; The ENCODE Project Consortium 2012). Indeed, OVOL2 ChIP-seq showed strong binding to the promoter region of

ovarian piRNA cluster 71 (Supplemental Fig. S7F). Interestingly, piRNA cluster 14, the second highest ranking ovarian piRNA cluster, showed several strong OVOL2 binding regions across the cluster locus (Supplemental Fig. S7F).

Similar to the binding pattern observed for mouse OVOL2 ChIP-seq, human OVOL2 ChIP-seq showed binding near promoters of the piRNA pathway genes (e.g., *TDRKH*, *TDRD7*, *TDRD3*, and *PIWIL4*) when overexpressed in the human iPSCs (Fig. 7B; Supplemental Fig. S7B). Moreover, human OVOL2 binding near promoters of piRNA pathway genes (e.g., *TDRKH*, *PIWIL4*, and *TDRD7*) occurred at orthologous regions in the mouse genome where mouse OVOL2 showed corresponding binding events (Fig. 7B). These orthologous OVOL2 binding events occurred at Ovo motifs that were highly conserved across the mammalian species based on multiple sequence alignments (Fig. 7B). Overall, our results indicate that the Ovo TF family interactions with CCGTTA motifs in the regulation of ovarian piRNA pathway components is a gene-regulatory feature that is conserved from insects to vertebrates.

Ovo motifs are hallmarks of ovarian piRNA clusters in metazoans

To test whether binding of Ovo to piRNA clusters in ovaries is a conserved feature of female germ cells, we analyzed Ovo motif enrichment within ovary piRNA clusters across multiple metazoan species. We characterized ovary piRNA clusters in species by their RPKM-normalized counts (more than one) mapping to the piRNA cluster regions (proTRAC piRNA cluster coordinates) (Rosenkranz and Zischler 2012; Rosenkranz et al. 2022) using ovary small RNA-seq data sets from humans (*Homo sapiens*), crab-eating macaques (*Macaca fascicularis*), mice (*Mus musculus*), golden hamsters (*Mesocricetus auratus*), cows (*Bos taurus*), zebrafish (*Danio rerio*), buff-tailed bumblebees (*Bombus terrestris*), African malaria mosquitoes (*Anopheles gambiae*), Arizona bark scorpions (*Centruroides sculpturatus*), and Pacific oysters (*Crassostrea gigas*). We then calculated the enrichments of OVOL2 motifs within the ovary piRNA clusters of each species. Our results revealed significant enrichment of OVOL2 motifs at ovary piRNA clusters in all 10 species when compared with genomic backgrounds ($P < 0.05$, Wilcoxon signed-rank test) (Fig. 7D).

Moreover, the Ovo motif enrichments at piRNA clusters were comparable with enrichments observed at promoter regions of the Ovo-upregulated target genes and germline-enriched genes and were significantly higher than the average signals observed at somatic-enriched genes or all genes in flies (Supplemental Fig. S6H). Overall, our results suggest a conserved regulatory mechanism using CCGTTA motifs underpinning the expression patterns of the germline piRNA pathway in metazoan species, where Ovo/OVOL2 interaction with the motifs in ovaries serves as a female counterpart of the male-specific regulation by A-MYB in testes (Fig. 7E).

Discussion

Transcription factors and coactivators controlling the male germline piRNA pathway have been previously described in vertebrate testes (Bolcun-Filas et al. 2011; Li et al. 2013; Cecchini et al. 2023; Yu et al. 2023); however, a female-specific counterpart of such a regulatory network controlling the expression of the female germline piRNA pathway in insect and vertebrate ovaries has remained an enigma. The identification of ovary-specific TFs that control the germline-specific piRNA pathway in female germ cells provides crucial insights into the regulation of transposon repression during oogenesis in animals.

In this study, we uncovered Ovo as the principal regulator of ~70% of the germline-specific piRNA pathway genes in *Drosophila*. The *ovo* locus encodes both somatic and germline isoforms that are driven by distinct promoters and were once thought to be two distinct genes: the somatic *shavenbaby* (*svb*) and the germline *ovo* (isoforms *ovo-A* and *ovo-B*) (Fig. 5E). The somatic *svb* is expressed only in embryonic, larval, and pupal epidermis cells, while the germline *ovo-B* is the major and the only essential isoform required for the viability of female *Drosophila* germ cells (Hayashi et al. 2017; Benner et al. 2024). Our results indicate that overexpression of *ovo-B* activates the germline piRNA pathway components in the ovarian somatic cells (Fig. 4), and *l(3)mbt* specifically represses the expression of *ovo-B* in the somatic cells by occupying the promoter region responsible for germline expression (Fig. 5E).

Our data point toward a gene-regulatory model of Ovo in female germ cells where transcription of germline piRNA pathway genes (e.g., *aub*, *ago3*, and *vas*) could be directed by Ovo family TFs, analogous to the model previously suggested for A-MYB in the regulation of the male germline piRNA pathway in mouse testes (Fig. 7F; Li et al. 2013). Similar to A-MYB, Ovo and its homologs autoregulate their own expression (Lü et al. 1998; Nair et al. 2007; Dragan et al. 2023) and, as our motif and ChIP-seq results reveal, show strong binding to germline piRNA pathway components. Our experiments show that Ovo is indeed able to activate expression of germline piRNA pathway components when ectopically expressed in fly somatic cells. This includes the nuage components Aub, Ago3, and Vas, which are activated in expression in somatic cells in the presence of ectopic Ovo and assemble to form cellular structures resembling the nuage bodies of germ cells (Fig. 4D–F). Of note, we found that upon ectopic Ovo expression, the germline-specific factors Rhino and Cuff are not upregulated in somatic cells, suggesting the involvement of additional regulatory mechanisms such as soma-expressed repressors or the lack of specific germline transcriptional coactivators in the somatic cells; however, these hypotheses require further experimental interrogation.

Our model is further supported by the multispecies analyses of OVOL2 ChIP-seq in mouse PGCLCs and human iPSCs revealing strong OVOL2 binding to the ovary piRNA clusters (Fig. 7), thus indicating a high degree of conservation of this gene-regulatory mechanism across

the animal kingdom. Previous work analyzing ovarian piRNA clusters observed strong A-MYB motif enrichments at ovarian piRNA clusters in macaques (Roovers et al. 2015), leading the investigators to suggest a paradoxical involvement of the testis-specific A-MYB in driving transcription of ovarian clusters in ovaries. This enrichment is compatible with our findings given that OVOL2 and A-MYB use the same motifs, and these motifs would recruit OVOL2 in ovaries.

Ovo's conserved role in germline development in animals ranging from flies to mice is well known (Hayashi et al. 2017) despite germ cell development in flies and mammals following distinct pathways (Richardson and Lehmann 2010). In flies, the preformation model states that the germline is established through maternally deposited germ cell determinants within oocytes, whereas mammalian PGCs develop according to epigenetic/inductive mechanisms (epigenesis) where external cues dictate their developmental trajectory (Extavour and Akam 2003; Richardson and Lehmann 2010). Importantly, Ovo in flies is maternally deposited as a component of the germplasm and later pole cells, which establish the fly PGCs (Yatsu et al. 2008), whereas mammalian OVOL2 acts downstream from BMP signaling to control cell fate decisions during mammalian PGC specification in the epiblast (Zhang et al. 2013; Hayashi et al. 2017), thus suggesting that Ovo family TFs control animal PGC development and germline piRNA pathway expression via both intrinsic and inductive mechanisms.

Interestingly, ~36% of the mouse OVOL2 binding sites were also occupied by A-MYB in testes, as both OVOL2 and A-MYB recognize the same core CCGTTA motif sequence (Fig. 7A). This could be a mechanism driving the expression of the shared germline piRNA clusters in both ovaries and testes (Supplemental Fig. S7A); however, other mechanisms such as motif affinity, chromatin accessibility, or methylation must account for expression of sex-specific piRNA clusters. In our motif analysis, we noticed a pseudopalindromic extension to the GTT core in the human A-MYB motif (Fig. 6A) that could prefer A-MYB over OVOL2 at testis-specific clusters. Additionally, we observed that ~52% of the OVOL2 binding events within the mouse ovary piRNA clusters occurred at the CpG islands (Fig. 7C; Supplemental Fig. S7A), which could be linked to sex-specific DNA methylation patterns in developing PGCs. TF binding sites are enriched at hypomethylated regions that evade the first wave of default de novo DNA methylation (Molaro et al. 2011), which starts from day 13 of embryonic development (E13.5) in male mouse PGCs (Molaro et al. 2014) and coincides with the appearance of the prepachytene piRNAs (Aravin et al. 2008); however, female PGCs undergo de novo methylation only later after birth (Schaefer et al. 2007). Temporal differences in DNA methylation and binding patterns by sex-specific TFs that direct female PGC development such as OVOL2 could therefore protect against methylation to control sex-specific piRNA cluster transcription. Of note, the mouse *Ovol2* gene encodes both the repressor isoform *Ovol2a* and the activator isoform *Ovol2b*. The binding patterns between OVOL2A and OVOL2B

ChIP-seqs were indistinguishable from each other at piRNA pathway components in mice; thus, the mechanism driving the ovary-specific piRNA pathway could also depend on the interplay of these isoforms.

In fly ovaries, Ovo is continuously present in the nucleus of germline cells, and maternal Ovo persists in the embryo until zygotic Ovo is expressed; thus, Ovo binding sites could be potentially marking genomic locations important during the transition from one generation to the next (Benner et al. 2024). Ovo persists in ovarian germ cells and plays a role in female germline sex determination by controlling expression of *Otu* and *Sxl* (Hempel et al. 2008). Our findings reveal its concurrent role in regulation of the female germline piRNA pathway. However, the testis-specific regulator controlling the male germline piRNA pathway in fly testes remains to be identified. The fly Myb TF is unlikely to act analogously to the vertebrate A-MYB, as it does not recognize the CCGTTA motifs (based on de novo motifs in ENCODE ChIP-seq data) and was reported to function as a weak repressor of piRNA factors in OSCs (Yamamoto-Matsuda et al. 2022); therefore, a different testis-specific TF candidate concurrently controlling spermatogenesis and the male piRNA pathway likely exists in fly testes. Interestingly, recent work identified the TF Trailblazer as an important regulator of the PIWI clade proteins *Aub* and *Ago3* in male germ cells (Chen et al. 2024); however, it does not seem to control other piRNA genes or general germ cell factors.

Intriguingly, OVO and MYB family TFs belong to different lineages and harbor distinct DNA-binding domains yet are capable of binding and competing for the same genomic CCGTTA motifs (Nair et al. 2007), thus illustrating an example of convergent evolution where unrelated classes of DNA-binding domains evolve to bind to the same DNA elements. Under this model, the CCGTTA motifs represent *cis*-regulatory elements that interact with conserved female germline-specific Ovo family TFs in ovaries and male germline-specific A-MYB in testes of animals that, in combination with cofactors (such as TCFL5) (Cecchini et al. 2023; Yu et al. 2023) and other epigenetic mechanisms [e.g., chromatin accessibility, DNA methylation, histone marks, and L(3)mbt], control the germline piRNA pathway. Moreover, Ovo and Ovo-like TFs are comprised of both repressor and activator isoforms that could interplay to control the piRNA pathway in a sex- and tissue-specific manner. Overall, our results reveal a conserved gene-regulatory mechanism involving interactions of the ovary- and testis-specific TFs with the CCGTTA *cis*-regulatory elements behind the regulation of the germline piRNA pathway in animal ovaries and testes.

Materials and methods

Cell culture and treatments

Wild-type OSCs (*Drosophila* Genomics Resource Center [DGRC]288) (Saito et al. 2009), Δ L(3)mbt-OSCs (DGRC 289) (Sumiyoshi et al. 2016), and fGS/OSS (DGRC 191)

(Niki et al. 2006) were purchased from the *Drosophila* Genomics Resource Center and 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 1645670), 0.6 mg/mL glutathione (Sigma-Aldrich G6013-25G), and 10 mU/mL human insulin (Sigma-Aldrich I9278-5ML). The fGS/OSS cells were cultured using 25% conditioned medium and passaged before reaching 50% confluency.

Fly stocks and handling

All flies were kept at 25°C on standard cornmeal or propionic food. Control *w¹¹¹⁸* flies were available in the Hanon laboratory. *Vas*-GFP flies were reported by Bence et al. (2017) and obtained from the Bloomington *Drosophila* Stock Center (BDSC; 76126).

Ovary dissections and cell dissociation

The ovary dissociation workflow was adapted from a previously published protocol (Slaidina et al. 2020). In brief, flies were fed with yeast extract 2–3 days before the dissections. Ovaries were dissected into ice-cold PBS, centrifuged at 500g for 5 min at 4°C, dissociated using 5 mg/mL trypsin (Sigma-Aldrich T1426-50MG) and 2.5 mg/mL collagenase-A (Sigma-Aldrich 10103586001) in Ringer's solution (600 µL for 148 ovaries) at 800 rpm for 1 h at 30°C in a thermomixer, passed through a 100 µm cell strainer followed by the addition of 400 µL of Schneider's medium (Thermo Fisher Scientific 21720001) containing 10% FBS, and centrifuged at 500g for 5 min at 4°C. The cell pellet was washed with 600 µL of PBS by centrifugation at 500g for 5 min at 4°C, and cells were resuspended in 500 µL of PBS. Dissociation of ~30 ovaries resulted in ~500,000 cells.

FACS sorting

The *w¹¹¹⁸* control flies and transgenic *vas*-GFP flies were fed with yeast paste 3 days prior to dissections. Ovaries were dissected in ice-cold PBS, and cells dissociated as described above. Cells were resuspended in PBS and sorted into GFP-positive and GFP-negative populations using a FACS Aria cell sorter instrument (BD Biosciences). One-hundred-forty-eight ovaries from *vas*-GFP flies resulted in ~36,000 *vas*-GFP-positive cells after FACS sorting, and 19 ovaries from the *w¹¹¹⁸* strain were dissociated and used as controls.

RNA isolation and RT-qPCR

RNA was isolated using RNeasy the minikit (Qiagen 74106) with an RNase-free DNase set (Qiagen 79254) for DNase digestion during RNA purification as per the manufacturer's instructions. Reverse transcription was performed with 100 ng to 1 µg of RNA using SuperScript IV reverse transcriptase (Thermo Fisher Scientific 18090010) and oligo-dT primers. RT-qPCR was performed on 1:10 diluted cDNA using Fast SYBR Green master mix

(Thermo Fisher Scientific 4385610) following the manufacturer's instructions on a Bio-Rad C1000 thermal cycler instrument. Primers were designed against exon–exon junctions for genes with introns, and *rpL32* was used as an internal standard (Supplemental Data Set S3). Relative expression was analyzed using the $\Delta\Delta C_t$ method (Livak and Schmittgen 2001).

RNA-seq

For total RNA-seq, ribosomal RNA depletion was performed using riboPOOLS for *Drosophila melanogaster* (Cambridge Bioscience) from 500 ng of total RNA input, as described previously (Munafò et al. 2021). For mRNA-seq, poly(A) selection was performed using the poly(A) mRNA magnetic isolation module (NEB E7490) with 500 ng of total RNA as input, as described previously (Fabry et al. 2019). Libraries were prepared using the NEB-Next Ultra directional RNA library preparation kit for Illumina (NEB E7420L). Library quality was assessed using an Agilent 2100 bioanalyzer (high-sensitivity DNA kit). Libraries were quantified with a KAPA library quantification kit for Illumina (Kapa Biosystems KK4873) and sequenced on a NovaSeq6000 system with two 50 bp reads on an SP flow cell.

ATAC-seq

For cell lines, 50,000 cells were used as described in the omni-ATAC-seq protocol (Corces et al. 2017). In brief, cells were lysed in ATAC resuspension buffer (RSB) containing 0.1% NP40, 0.1% Tween-20, and 0.01% digitonin and washed out with cold ATAC RSB containing 0.1% Tween-20. The transposition reaction was performed with 25 µL of 2× TD buffer and 2.5 µL of transposase (100 nM final) from the Illumina Tagment DNA enzyme and buffer small kit (Illumina 20034197), 16.5 µL of PBS, 0.5 µL of 1% digitonin, 0.5 µL of 10% Tween-20, and 5 µL of water and mixed at 1000 RPM in a thermomixer for 30 min at 37°C. For the whole-ovary ATAC-seq, the omni-ATAC-seq protocol was optimized by adjusting the numbers of ovaries used (six, 10, 20, and 30) and tagmentation time (30 min to 1 h). The Zymo DNA Clean and Concentrator-5 kit (Zymo D4014) was used for cleanup, and transposed fragments were amplified for five cycles [5 min at 72°C and 30 sec at 98°C (10 sec at 98°C, 30 sec at 63°C, and 1 min at 72°C for) × 5] using NEBNext 2× MasterMix (NEB M0541S) with Ad1_noMX and v2_Ad2.* indexing primers followed by qPCR amplification to determine the required additional cycle numbers. Amplified DNA was purified using the Zymo DNA Clean and Concentrator-5 kit (Zymo D4014) and eluted in 20 µL of water. AMPure XP beads (Beckman Coulter A63881) were used for double-sided bead purification. Fragments (100–600 bp) were selected using Blue Pippin (Sage Science NC1025035). The library fragment size distribution was assessed using an Agilent 2100 bioanalyzer on a high-sensitivity DNA kit. Libraries were quantified with a KAPA library quantification kit for Illumina (Kapa

Biosystems KK4873) and sequenced on a NovaSeq6000 system with two 50 bp reads on an SP flow cell.

ChIP-seq

For ectopic Ovo ChIP-seq, $>30 \times 10^6$ OSCs ($n=2$) were nucleofected with an *ovo-B-FLAG* construct using 0.5–0.8 μg of plasmid per million cells, the nucleofector kit V (Lonza VVCA-1003), and the T-029 program on a Nucleofector 2b device (Lonza) and cross-linked 48 h later by covering the cells in 1% formaldehyde (FA) solution (50 mM HEPES-KOH, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% FA) for 10 min at room temperature. Cell lysis was performed as described previously (Schmidt et al. 2009). Sonication was performed with 15–18 cycles of 10 sec on and 60 sec off (Fisherbrand Q705 Sonicator and 418-21 probe) in 3 mL of LB3 buffer (10 mM Tris-HCl at pH 8, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-deoxycholate, 0.5% N-lauroylsarcosine) in a 7 mL conical polypropylene tube to shear DNA into 100–500 bp fragments. ChIP was performed using 5–10 μg of primary mouse monoclonal anti-FLAG M2 antibody (Merck F1804-200UG) in 250 μL of LB3 with 1% Triton X-100 overnight at 4°C. Cross-links were reversed by incubation for 16 h at 65°C. Proteins and RNA were enzymatically digested. DNA purification was performed using phenol-chloroform extraction and ethanol precipitation. All of the purified DNA and 220 ng of whole-cell extract DNA were used to prepare ChIP-seq libraries using the NEBNext Ultra II DNA library preparation kit for Illumina (NEB E7645S) with PCR amplifications carried out for 16 cycles. The libraries were assessed for fragment size distribution using an Agilent 2100 bioanalyzer on a high-sensitivity DNA kit, quantified with the KAPA library quantification kit for Illumina (Kapa Biosystems KK4873), and sequenced on a NovaSeq6000 system with two 150 bp reads on an SP flow cell.

Overexpression screen

Candidate genes were amplified from the *Drosophila melanogaster* ovary cDNA library using KOD Hot Start DNA polymerase (Merck Millipore 71086-4) and cloned into overexpression vectors under the *act5C* promoter with either N-terminal or C-terminal FLAG tags using Gibson assembly master mix (NEB E2611L) for 1 h at 50°C. Nine microliters of Mix & Go competent cells (strain DH5 α) were thawed on ice and transformed with 1 μL of diluted (2.5 \times in water) Gibson assembly reaction by incubation for 5 min on ice before plating on LB plates containing the appropriate antibiotic. Colony PCR was performed to identify the colonies harboring the ligated constructs, followed by inoculation of the colonies into 3 mL of broth containing 100 $\mu\text{g}/\text{mL}$ ampicillin and shaking overnight at 37°C. The constructs were purified using Qiagen Plasmid Plus kits and verified by Sanger sequencing. Constructs were then transfected into OSCs and $\Delta(3)mbt$ OSCs with the nucleofector kit V (Lonza VVCA-1003) using the T-029 program and a Nucleofector 2b device (Lonza). Cells were passaged 1 day prior to transfections and were

70%–80% confluent at the time of transfections. The ectopic *ovo* overexpression vector carrying the FLAG-tagged *ovo-B* isoform matching NM_080338 transcript was used.

siRNA knockdown experiments

Sense and antisense 21 nt siRNA sequences for target genes were designed using the DSIR tool (<http://biodev.ccea.fr/DSIR/DSIR.html>; Supplemental Data Set S3). The designed sequences were ordered from IDT and resuspended in 400 μL of RNase-free water. Equal volumes of the resuspended sense and antisense siRNA were mixed and then added to an equal amount of 2 \times annealing buffer (60 mM potassium acetate, 200 mM HEPES at pH 7.5). The mix was boiled for 5 min at 75°C and then ramped down to 25°C (–0.1°C/sec) to anneal the siRNA sequences into the siRNA duplex (100 μM final concentration = 200 pmol). Two microliters of the final 100 μM siRNA duplex was mixed with 100 μL of nucleofection solution V (Lonza VVCA-1003) and transfected into 10 million cells using a Nucleofector 2b device and program T-029 (Lonza). Cells were plated into 12 well plates, and the media was changed after 24 h. Either RNA was harvested at 48 h or nucleofection was repeated after 48 h and RNA was harvested at 96 h.

Western blots

Proteins were extracted from 1×10^6 to 3×10^6 cells, washed in PBS at 300g for 5 min, and lysed in ice-cold RIPA buffer (Thermo Scientific 89900) containing cOmplete protease inhibitor cocktail EDTA-free tablets (Roche; one tablet per 20 mL of RIPA) under rocking conditions for 30 min at 4°C. The lysate was centrifuged 20,000g for 20 min at 4°C, and supernatant was transferred to a new tube. The protein concentrations were measured using a Direct Detect spectrometer. Protein (15 mg/mL) was used with NuPAGE 4 \times LDS sample buffer and NuPAGE 10 \times sample-reducing agent (Thermo Fisher Scientific) in 20 μL of total volume and denatured for 10 min at 70°C followed by protein gel electrophoresis using 4%–12% NuPAGE and Bis-Tris in 1.5 mm 10 well miniprotein gels and XCell SureLock minicell electrophoresis system (Thermo Fisher Scientific) at 180 V for 1 h in a 4°C cold room. Precision Plus protein All Blue prestained protein standard (Bio-Rad 1610373) was run in parallel as a ladder. Blotting was performed for 7 min with iBlot 2 transfer stacks, nitrocellulose, and mini using an iBlot 2 gel transfer device. The nitrocellulose membrane was blocked for 1 h at room temperature under gentle agitation using TBS (LI-COR Biosciences). The membrane was rinsed with TBS buffer and incubated with primary mouse monoclonal anti-FLAG M2 antibody (1:2500; Merck F1804-200UG) and primary rabbit polyclonal anti-Tubulin (DM1A+DM1B; 1:5000; Abcam ab18251) in TBS with 0.1% Tween 20 (TBST) overnight at 4°C under gentle agitation. The membrane was washed three times for 5–10 min with TBST and incubated with the secondary antibodies IRDye 800CW donkey antimouse IgG secondary antibody (1:5000; LI-COR Biosciences

926-32212) and IRDye 680RD donkey antirabbit IgG secondary antibody (1:20,000; LI-COR Biosciences 926-68073) in TBST for 1 h room temperature under gentle agitation. The membrane was washed three times for 10 min with TBST and with TBS before image acquisition on an Odyssey CLx infrared imaging system (LI-COR Biosciences).

Immunofluorescence (IF) and microscopy

OSC immunostainings were performed as described previously (Eastwood et al. 2021). Coverslips were coated with fibronectin (1:50 in PBS; Sigma-Aldrich F0895-1MG) for 1 h at 26°C in 6 well plates, and cells were plated on top. After 48 h, cells attached to coverslips were fixed in 4% PFA in PBS (500 μ L per well) for 15 min at room temperature and rinsed three times with PBS. Cells were permeabilized with 0.2% Triton-X-100 in PBS for 10 min at room temperature and rinsed three times with PBS. Blocking was performed with 0.1% Tween-20 and 1% BSA in PBS. Primary mouse monoclonal anti-FLAG M2 (1:500; Merck F1804-200UG); primary rabbit monoclonal anti-FLAG (Cell Signaling Technology 14793S); primary rabbit polyclonal anti-Piwi, anti-Aub, or anti-Ago3 (1:500; available in-house; described by Brennecke et al. (2007); primary mouse monoclonal anti-Aub (1:500; Brennecke laboratory 8A3-D7)) (Senti et al. 2015); primary mouse monoclonal anti-Ago3 (1:500; Brennecke laboratory 7B4-C2); primary mouse monoclonal anti-Yb (Siomi laboratory, 8H12B12 and 180803); primary rat monoclonal anti-Vasa (Developmental Studies Hybridoma Bank [DSHB] AB_760351); and primary mouse monoclonal anti-Lamin Dm0 (DSHB adl67.10) antibodies were diluted in 0.1% Tween-20, 0.2% BSA, and 1 \times PBS (500 μ L/well) and added onto coverslips in each well. Incubation with primary antibodies was performed overnight at 4°C under gentle agitation. Coverslips were washed three times for 5 min with PBST (1 \times PBS with 0.1% Tween-20) under gentle agitation. Goat antimouse IgG (H + L) highly cross-adsorbed secondary antibody AlexaFluor-647 (Thermo Fisher Scientific A-21236) and goat antirabbit IgG (H + L) cross-adsorbed secondary antibody AlexaFluor-488 (Thermo Fisher Scientific, A-11008) were diluted 1:500 in 0.1% Tween-20, 0.2% BSA, and 1 \times PBS and added onto coverslips (500 μ L/well) for 1 h at room temperature under gentle agitation (covered with aluminum foil) followed by three 5 min washes with PBST. Coverslips were incubated with DAPI at 1:10,000 dilution in PBST for 10 min at room temperature (covered with aluminum foil) followed by two 5 min washes with PBST under gentle agitation. Coverslips were mounted on glass slides using a drop of ProLong Diamond antifade mountant (Thermo Fisher Scientific P36961). Imaging was performed on a Leica SP8 confocal laser scanning system, and image analysis was performed using Leica application suite X (v3.5.7.23225) and Huygens Professional (v20.04).

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 the Cutadapt tool (v1.18; default parameters). Burrows–Wheeler aligner (BWA; v0.7.17, bwa mem -M -t 4) (Li and Durbin 2009) was used to align the trimmed paired reads to the *D. melanogaster* genome (BDGP release 6+ISO1 MT/dm6). Duplicates were marked using Picard tool (v2.9.0; MarkDuplicates, validation stringency=lenient). SAMtools (v1.9) was used for indexing and filtering (Li et al. 2009). The quality metrics for the aligned ATAC-seq reads were assessed using ataqv (v1.0.0) (<https://github.com/ParkerLab/ataqv>; Orchard et al. 2020). The ATAC-seq peaks were called with MACS2 (v2.1.1.20160309) using “--nomodel --shift -37 --extsize 73” parameters and FDR cutoff of $q \leq 0.05$ (Zhang et al. 2008). Differential accessibility analysis was performed using the DiffBind package (v3.0.15). RPKM-normalized bigWig files were generated using deepTools (v3.5.1; bamCoverage) (Ramírez et al. 2014). Peak intersections were performed using bedtools (v2.30.0) (Quinlan and Hall 2010). ATAC-seq heat maps and profiles were generated using the plotHeatmap and plotProfile tools in deepTools (v3.5.1). Genome browser visualizations were generated using the UCSC genome browser (Kent et al. 2002).

RNA-seq data analysis

The reads were trimmed of adapter sequences using the Cutadapt tool (v1.18; -m 1 specified to not have reads trimmed to zero). The trimmed reads were aligned to the genome assemblies using the RNA-seq aligner STAR (v2.7.3a) (Dobin et al. 2013). Gene counts were calculated with the featureCounts tool (Subread package v1.5.3; -s 2 for stranded libraries prepared by the dUTP method) (Liao et al. 2014). *D. melanogaster* gene annotations were taken from *Drosophila melanogaster*.BDGP6.28.100.gtf in the Ensembl genome database (Martin et al. 2023). SAMtools (v1.9) was used to merge bam files from replicates. Differential gene expression analysis was performed using DESeq2 (v1.30.1) (Love et al. 2014). The RPKM-normalized bigWig files were generated for each strand using the bamCoverage tool (--filterRNAstrand specified for dUTP stranded libraries) in deepTools (v3.5.1). Raw data for publicly available RNA-seq reads were downloaded from the sources listed in Supplemental Data Set S3 and processed as above. Genome browser visualizations were generated using the UCSC genome browser.

Single-cell RNA-seq analysis

The single-cell RNA-seq data matrices were downloaded from the sources listed in Supplemental Data Set S3 and clustered using the Seurat 4.0 toolkit (v4.3.0.1) (Hao et al. 2021). Normalizations were performed with the SCTransform() function. Differential expression analysis was performed using the FindMarkers() function. Pseudotime analysis was carried out with the Monocle 3 tool (v3.0) (Cao et al. 2019). Correlations were made with the Pearson correlation coefficients, and *P*-values were

corrected for multiple testing using Bonferroni correction. The custom code for single-cell RNA-seq coexpression analysis is available at GitHub (https://github.com/alizadaa/Single-cell_RNA-seq_Co-expression_Analysis_Ovary).

ChIP-seq data analysis

The reads were trimmed of adapter sequences using the Cutadapt tool (v1.18). The trimmed reads were aligned to the genome assemblies using Burrows–Wheeler aligner (BWA; v0.7.17, `bwa aln`). SAMtools (v1.9) was used for sorting, merging, and indexing. ChIP-seq peaks were called using the MACS3 (v3.0.0a6) `callpeak` command (`-q 0.01`), 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) (Li et al. 2011). The RPKM-normalized bigWig files were generated using the `bamCoverage` tool in `deepTools` (v3.5.1) with the parameter `--extendReads 120` specified. The input-normalized bigWig files were generated using the `bamCompare` tool in `deepTools` (v3.5.1). Signals from bigWig files were quantified using the `multiBigwigSummary` tool in `deepTools` (v3.5.1). ChIP-seq heat maps and profiles were generated using the `plotHeatmap` and `plotProfile` tools in `deepTools` (v3.5.1). Raw data for publicly available ChIP-seq data sets were downloaded from the sources listed in Supplemental Data Set S3 and processed as described above. Genome browser visualizations were generated using the UCSC genome browser.

Small RNA-seq and piRNA cluster analysis

The raw small RNA-seq reads from ovaries and testes of humans (*Homo sapiens*), crab-eating macaques (*Macaca fascicularis*), mice (*Mus musculus*), golden hamsters (*Mesocricetus auratus*), cows (*Bos taurus*), zebrafish (*Danio rerio*), buff-tailed bumblebees (*Bombus terrestris*), African malaria mosquitoes (*Anopheles gambiae*), Arizona bark scorpions (*Centruroides sculpturatus*), and Pacific oysters (*Crassostrea gigas*) were downloaded from the sources listed in Supplemental Data Set S3. Reads were trimmed of adapter sequences using Cutadapt tool (v1.18) and aligned to the respective genome assemblies using the RNA-seq aligner STAR (v2.7.3a). SAMtools (v1.9) was used to merge bam files from replicates. The RPKM-normalized bigWig files were generated using the `bamCoverage` tool in `deepTools` (v3.5.1). The piRNA cluster coordinates were taken from the piRNA cluster database (<https://www.smallrnagroup.uni-mainz.de/piRNAclusterDB>; Rosenkranz et al. 2022). Ovary and testis piRNA clusters were defined and ranked using the RPKM-normalized counts (more than one) mapping to the piRNA cluster regions as calculated from bigWig files of the ovary and testis small RNA-seq data sets using `multiBigwigSummary` tool in `deepTools` (v3.5.1). Genome browser visualizations were generated using the UCSC genome browser.

Motif analysis, peak enrichments, and clustering

Motif scanning was performed using the FIMO tool in MEME suite (v5.4.1) (<https://meme-suite.org/meme/index.html>; Grant et al. 2011). De novo ChIP-seq motifs were generated using the MEME-ChIP tool in MEME suite (v5.4.1) (Ma et al. 2014). A list of *Drosophila* motifs was downloaded from the FlyFactorSurvey (<http://pgfe.umassmed.edu/TFDBS>) database (Zhu et al. 2011). `fastaFromBed` (v2.26.0) was used for conversion of bed file coordinates into fasta format using *D.mel* BDGP6.28.dna.toplevel genome (Bailey et al. 2015). Data were visualized with UCSC genome browser. ChIPseeker (v1.36.0) (Yu et al. 2015) was used to annotate the genomic features and distances based on the TxDb.Dmelanogaster.UCSC.dm6.ensGene annotation package. The radial tree clustering of human TF motifs was adapted from the JASPAR database (<https://jaspar.elixir.no/matrix-clusters>, JASPAR 2022 vertebrates CORE, RSAT matrix-clustering; Castro-Mondragon et al. 2017; Rauluseviciute et al. 2024). Orthologous regions between the species genomes were derived using the UCSC LiftOver tool (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>; Hinrichs et al. 2006). The species classification tree was generated using the NCBI taxonomy database (Schoch et al. 2020) and visualized in TreeViewer (Bianchini and Sánchez-Baracaldo 2024). Multiple sequence alignments of motifs and phylogenetic trees were based on Multiz alignments (27 way insect alignment and 46 way vertebrates) (Blanchette et al. 2004) from the UCSC conservation tracks (Murphy et al. 2001; Kent et al. 2002; Hubisz et al. 2011).

Data availability

RNA-seq, ATAC-seq, and ChIP-seq data generated in this study have been deposited to the GEO database under series GSE233246.

Code availability

Custom code is available at GitHub (https://github.com/alizadaa/Single-cell_RNA-seq_Co-expression_Analysis_Ovary).

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We thank Anna Sobieszek for preparing RNA-seq libraries from OSC knockdown samples. We are grateful to Susanne Bornelöv for advice on computational analyses. We thank members of the Hannon laboratory for fruitful discussions. We thank Marie Bao (Life Science Editors) for advice, and Anna Nicholson for comments on the manuscript. We thank the Scientific Computing Core at the Cancer Research UK Cambridge Institute for HPC resources, the Flow Cytometry Core for FACS services,

Research Instrumentation and Cell Services for liquid nitrogen storage and mycoplasma testing, the Light Microscopy Core for training and advise, and the Genomics Core for sequencing services. G.J.H. 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).

Author contributions: A.A., B.C.N., and G.J.H. conceived the study. A.A. and B.C.N. designed the experiments and interpreted the results. A.A. performed all wet-laboratory experiments and computational analyses and wrote the first manuscript draft. B.C.N. and G.J.H. supervised the project and wrote the manuscript. All authors read and approved the final version.

References

- Andersen PR, Tirian L, Vunjak M, Brennecke J. 2017. A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* **549**: 54–59. doi:10.1038/nature23482
- Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ. 2008. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol Cell* **31**: 785–799. doi:10.1016/j.molcel.2008.09.003
- Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME suite. *Nucleic Acids Res* **43**: W39–W49. doi:10.1093/nar/gkv416
- Bastock R, St Johnston D. 2008. *Drosophila* oogenesis. *Curr Biol* **18**: R1082–R1087. doi:10.1016/j.cub.2008.09.011
- Baumgartner L, Handler D, Platzer SW, Yu C, Duchek P, Brennecke J. 2022. The *Drosophila* ZAD zinc finger protein Kipferl guides Rhino to piRNA clusters. *eLife* **11**: e80067. doi:10.7554/eLife.80067
- Bence M, Jankovics F, Lukácsovich T, Erdelyi M. 2017. Combining the auxin-inducible degradation system with CRISPR/Cas9-based genome editing for the conditional depletion of endogenous *Drosophila melanogaster* proteins. *FEBS J* **284**: 1056–1069. doi:10.1111/febs.14042
- Benner L, Muron S, Gomez JG, Oliver B. 2024. OVO positively regulates essential maternal pathways by binding near the transcriptional start sites in the *Drosophila* female germline. *eLife* **13**: RP94631. doi:10.7554/eLife.94631
- Bianchini G, Sánchez-Baracaldo P. 2024. Treeviewer: flexible, modular software to visualise and manipulate phylogenetic trees. *Ecol Evol* **14**: e10873. doi:10.1002/ece3.10873
- Bielinska B, Lü J, Sturgill D, Oliver B. 2005. Core promoter sequences contribute to ovo-B regulation in the *Drosophila melanogaster* germline. *Genetics* **169**: 161–172. doi:10.1534/genetics.104.033118
- Blanchard DP, Georgette D, Antoszewski L, Botchan MR. 2014. Chromatin reader L(3)mbt requires the Myb-MuvB/DREAM transcriptional regulatory complex for chromosomal recruitment. *Proc Natl Acad Sci* **111**: E4234–E4243. doi:10.1073/pnas.1416321111
- Blanchette M, Kent WJ, Riemer C, Elnitski L, Smit AF, Roskin KM, Baertsch R, Rosenbloom K, Clawson H, Green ED, et al. 2004. Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res* **14**: 708–715. doi:10.1101/gr.1933104
- Bolcun-Filas E, Bannister LA, Barash A, Schimenti KJ, Hartford SA, Eppig JJ, Handel MA, Shen L, Schimenti JC. 2011. A MYB (MYBL1) transcription factor is a master regulator of male meiosis. *Development* **138**: 3319–3330. doi:10.1242/dev.067645
- Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128**: 1089–1103. doi:10.1016/j.cell.2007.01.043
- Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, et al. 2019. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**: 496–502. doi:10.1038/s41586-019-0969-x
- Carmell MA, Girard A, van de Kant HJ, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ. 2007. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev Cell* **12**: 503–514. doi:10.1016/j.devcel.2007.03.001
- Castro-Mondragon JA, Jaeger S, Thieffry D, Thomas-Chollier M, van Helden J. 2017. RSAT matrix-clustering: dynamic exploration and redundancy reduction of transcription factor binding motif collections. *Nucleic Acids Res* **45**: e119. doi:10.1093/nar/gkx314
- Cecchini K, Biasini A, Yu T, Saflund M, Mou H, Arif A, Eghbali A, Colpan C, Gainetdinov I, de Rooij DG, et al. 2023. The transcription factor TCFL5 responds to A-MYB to elaborate the male meiotic program in mice. *Reproduction* **165**: 183–196. doi:10.1530/REP-22-0355
- Chen D, McKearin DM. 2003. A discrete transcriptional silencer in the *bam* gene determines asymmetric division of the *Drosophila* germline stem cell. *Development* **130**: 1159–1170. doi:10.1242/dev.00325
- Chen Y, Pane A, Schübach T. 2007. Cutoff and aubergine mutations result in retrotransposon upregulation and checkpoint activation in *Drosophila*. *Curr Biol* **17**: 637–642. doi:10.1016/j.cub.2007.02.027
- Chen P, Pan KC, Park EH, Luo Y, Lee YCG, Aravin AA. 2024. Escalation of genome defense capacity enables control of an expanding meiotic driver. bioRxiv doi:10.1101/2024.06.12.598716
- Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B, et al. 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* **14**: 959–962. doi:10.1038/nmeth.4396
- Coux RX, Teixeira FK, Lehmann R. 2018. L(3)mbt and the LINT complex safeguard cellular identity in the *Drosophila* ovary. *Development* **145**: dev160721. doi:10.1242/dev.160721
- Czech B, Preall JB, McGinn J, Hannon GJ. 2013. A transcriptome-wide RNAi screen in the *Drosophila* ovary reveals factors of the germline piRNA pathway. *Mol Cell* **50**: 749–761. doi:10.1016/j.molcel.2013.04.007
- Czech B, Munafò M, Ciabrelli F, Eastwood EL, Fabry MH, Kneuss E, Hannon GJ. 2018. piRNA-guided genome defense: from biogenesis to silencing. *Annu Rev Genet* **52**: 131–157. doi:10.1146/annurev-genet-120417-031441
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21. doi:10.1093/bioinformatics/bts635
- Dragan M, Chen Z, Li Y, Le J, Sun P, Haensel D, Sureshchandra S, Pham A, Lu E, Pham KT, et al. 2023. *Ovo1/2* loss-induced epidermal defects elicit skin immune activation and alter global metabolism. *EMBO Rep* **24**: e56214. doi:10.15252/embr.202256214
- Eastwood EL, Jara KA, Bornelöv S, Munafò M, Frantzis V, Kneuss E, Barbar EJ, Czech B, Hannon GJ. 2021. Dimerisation of the

- PICTS complex via LC8/Cut-up drives co-transcriptional transposon silencing in *Drosophila*. *eLife* **10**: e65557. doi:10.7554/eLife.65557
- ElMaghraby MF, Andersen PR, Pühringer F, Hohmann U, Meixner K, Lendl T, Tirian L, Brennecke J. 2019. A heterochromatin-specific RNA export pathway facilitates piRNA production. *Cell* **178**: 964–979.e20. doi:10.1016/j.cell.2019.07.007
- The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57–74. doi:10.1038/nature11247
- Extavour CG, Akam M. 2003. Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. *Development* **130**: 5869–5884. doi:10.1242/dev.00804
- Fabry MH, Ciabrelli F, Munafò M, Eastwood EL, Kneuss E, Falcatori I, Falconio FA, Hannon GJ, Czech B. 2019. piRNA-guided co-transcriptional silencing coopts nuclear export factors. *eLife* **8**: e47999. doi:10.7554/eLife.47999
- Forbes A, Lehmann R. 1998. Nanos and Pumilio have critical roles in the development and function of *Drosophila* germline stem cells. *Development* **125**: 679–690. doi:10.1242/dev.125.4.679
- Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**: 1017–1018. doi:10.1093/bioinformatics/btr064
- Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC. 2007. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* **315**: 1587–1590. doi:10.1126/science.1140494
- Handler D, Meixner K, Pizka M, Lauss K, Schmied C, Gruber FS, Brennecke J. 2013. The genetic makeup of the *Drosophila* piRNA pathway. *Mol Cell* **50**: 762–777. doi:10.1016/j.molcel.2013.04.031
- Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, et al. 2021. Integrated analysis of multimodal single-cell data. *Cell* **184**: 3573–3587.e3529. doi:10.1016/j.cell.2021.04.048
- Hayashi M, Shinozuka Y, Shigenobu S, Sato M, Sugimoto M, Ito S, Abe K, Kobayashi S. 2017. Conserved role of Ovo in germline development in mouse and *Drosophila*. *Sci Rep* **7**: 40056. doi:10.1038/srep40056
- Hempel LU, Kalamegham R, Smith JE III, Oliver B. 2008. *Drosophila* germline sex determination: integration of germline autonomous cues and somatic signals. *Curr Top Dev Biol* **83**: 109–150. doi:10.1016/S0070-2153(08)00404-3
- Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey TS, Harte RA, Hsu F, et al. 2006. The UCSC genome browser database: update 2006. *Nucleic Acids Res* **34**: D590–D598. doi:10.1093/nar/gkj144
- Hubisz MJ, Pollard KS, Siepel A. 2011. PHAST and RPHAST: phylogenetic analysis with space/time models. *Brief Bioinform* **12**: 41–51. doi:10.1093/bib/bbq072
- Hur JK, Luo Y, Moon S, Ninova M, Marinov GK, Chung YD, Aravin AA. 2016. Splicing-independent loading of TREX on nascent RNA is required for efficient expression of dual-strand piRNA clusters in *Drosophila*. *Genes Dev* **30**: 840–855. doi:10.1101/gad.276030.115
- Janic A, Mendizabal L, Llamazares S, Rossell D, Gonzalez C. 2010. Ectopic expression of germline genes drives malignant brain tumor growth in *Drosophila*. *Science* **330**: 1824–1827. doi:10.1126/science.1195481
- Jevitt A, Chatterjee D, Xie G, Wang XF, Otwell T, Huang YC, Deng WM. 2020. A single-cell atlas of adult *Drosophila* ovary identifies transcriptional programs and somatic cell lineage regulating oogenesis. *PLoS Biol* **18**: e3000538. doi:10.1371/journal.pbio.3000538
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. *Genome Res* **12**: 996–1006. doi:10.1101/gr.229102
- Klattenhoff C, Bratu DP, McGinnis-Schultz N, Koppetsch BS, Cook HA, Theurkauf WE. 2007. *Drosophila* rasiRNA pathway mutations disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response. *Dev Cell* **12**: 45–55. doi:10.1016/j.devcel.2006.12.001
- Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A, et al. 2009. The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* **138**: 1137–1149. doi:10.1016/j.cell.2009.07.014
- Kneuss E, Munafò M, Eastwood EL, Deumer US, Preall JB, Hannon GJ, Czech B. 2019. Specialization of the *Drosophila* nuclear export family protein Nxf3 for piRNA precursor export. *Genes Dev* **33**: 1208–1220. doi:10.1101/gad.328690.119
- Kobayashi S, Yamada M, Asaoka M, Kitamura T. 1996. Essential role of the posterior morphogen nanos for germline development in *Drosophila*. *Nature* **380**: 708–711. doi:10.1038/380708a0
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754–1760. doi:10.1093/bioinformatics/btp324
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078–2079. doi:10.1093/bioinformatics/btp352
- Li Q, Brown JB, Huang H, Bickel PJ. 2011. Measuring reproducibility of high-throughput experiments. *Ann Appl Stat* **5**: 1752–1779. doi:10.1214/11-AOAS466
- Li XZ, Roy CK, Dong X, Bolcun-Filas E, Wang J, Han BW, Xu J, Moore MJ, Schimenti JC, Weng Z, et al. 2013. An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. *Mol Cell* **50**: 67–81. doi:10.1016/j.molcel.2013.02.016
- Liao Y, Smyth GK, Shi W. 2014. Featurecounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930. doi:10.1093/bioinformatics/btt656
- Lim AK, Kai T. 2007. Unique germ-line organelle, nuage, functions to repress selfish genetic elements in *Drosophila melanogaster*. *Proc Natl Acad Sci* **104**: 6714–6719. doi:10.1073/pnas.0701920104
- Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* **25**: 402–408. doi:10.1006/meth.2001.1262
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550. doi:10.1186/s13059-014-0550-8
- Lü J, Andrews J, Pauli D, Oliver B. 1998. *Drosophila* OVO zinc-finger protein regulates ovo and ovarian tumor target promoters. *Dev Genes Evol* **208**: 213–222. doi:10.1007/s004270050175
- Ma W, Noble WS, Bailey TL. 2014. Motif-based analysis of large nucleotide data sets using MEME-ChIP. *Nat Protoc* **9**: 1428–1450. doi:10.1038/nprot.2014.083
- Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ. 2009. Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* **137**: 522–535. doi:10.1016/j.cell.2009.03.040

Alizada et al.

- Martin FJ, Amode MR, Aneja A, Austine-Orimoloye O, Azov AG, Barnes I, Becker A, Bennett R, Berry A, Bhai J, et al. 2023. Ensembl 2023. *Nucleic Acids Res* **51**: D933–D941. doi:10.1093/nar/gkac958
- The modENCODE Consortium, Roy S, Ernst J, Kharchenko PV, Kheradpour P, Negre N, Eaton ML, Landolin JM, Bristow CA, Ma L, et al. 2010. Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* **330**: 1787–1797. doi:10.1126/science.1198374
- Mohn F, Sienski G, Handler D, Brennecke J. 2014. The rhinoceros-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell* **157**: 1364–1379. doi:10.1016/j.cell.2014.04.031
- Molaro A, Hodges E, Fang F, Song Q, McCombie WR, Hannon GJ, Smith AD. 2011. Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates. *Cell* **146**: 1029–1041. doi:10.1016/j.cell.2011.08.016
- Molaro A, Falcatori I, Hodges E, Aravin AA, Marran K, Rafii S, McCombie WR, Smith AD, Hannon GJ. 2014. Two waves of de novo methylation during mouse germ cell development. *Genes Dev* **28**: 1544–1549. doi:10.1101/gad.244350.114
- Muerdter F, Guzzardo PM, Gillis J, Luo Y, Yu Y, Chen C, Fekete R, Hannon GJ. 2013. A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in *Drosophila*. *Mol Cell* **50**: 736–748. doi:10.1016/j.molcel.2013.04.006
- Munafò M, Lawless VR, Passera A, MacMillan S, Bornelöv S, Haussmann IU, Soller M, Hannon GJ, Czech B. 2021. Channel nuclear pore complex subunits are required for transposon silencing in *Drosophila*. *eLife* **10**: e66321. doi:10.7554/eLife.66321
- Murphy WJ, Eizirik E, O'Brien SJ, Madsen O, Scally M, Douady CJ, Teeling E, Ryder OA, Stanhope MJ, de Jong WW, et al. 2001. Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* **294**: 2348–2351. doi:10.1126/science.1067179
- Nair M, Bilanchone V, Ortt K, Sinha S, Dai X. 2007. Ovol1 represses its own transcription by competing with transcription activator c-Myb and by recruiting histone deacetylase activity. *Nucleic Acids Res* **35**: 1687–1697. doi:10.1093/nar/gkl1141
- Naitou Y, Nagamatsu G, Hamazaki N, Shirane K, Hayashi M, Hayashi M, Kobayashi S, Hayashi K. 2022. Dual role of *Ovol2* on the germ cell lineage segregation during gastrulation in mouse embryogenesis. *Development* **149**: dev200319. doi:10.1242/dev.200319
- Niki Y, Yamaguchi T, Mahowald AP. 2006. Establishment of stable cell lines of *Drosophila* germ-line stem cells. *Proc Natl Acad Sci* **103**: 16325–16330. doi:10.1073/pnas.0607435103
- Ohlstein B, Lavoie CA, Vef O, Gateff E, McKearin DM. 2000. The *Drosophila* cystoblast differentiation factor, *benign gonial cell neoplasm*, is related to DExH-box proteins and interacts genetically with *bag-of-marbles*. *Genetics* **155**: 1809–1819. doi:10.1093/genetics/155.4.1809
- Orchard P, Kyono Y, Hensley J, Kitzman JO, Parker SCJ. 2020. Quantification, dynamic visualization, and validation of bias in ATAC-seq data with ataqv. *Cell Syst* **10**: 298–306.e4. doi:10.1016/j.cels.2020.02.009
- Ozata DM, Gainetdinov I, Zoch A, O'Carroll D, Zamore PD. 2019. PIWI-interacting RNAs: small RNAs with big functions. *Nat Rev Genet* **20**: 89–108. doi:10.1038/s41576-018-0073-3
- Pane A, Wehr K, Schüpbach T. 2007. Zucchini and squash encode two putative nucleases required for rasiRNA production in the *Drosophila* germline. *Dev Cell* **12**: 851–862. doi:10.1016/j.devcel.2007.03.022
- Patil VS, Kai T. 2010. Repression of retroelements in *Drosophila* germline via piRNA pathway by the Tudor domain protein Tejas. *Curr Biol* **20**: 724–730. doi:10.1016/j.cub.2010.02.046
- Patil VS, Anand A, Chakrabarti A, Kai T. 2014. The Tudor domain protein Tapas, a homolog of the vertebrate Tdrd7, functions in the piRNA pathway to regulate retrotransposons in germline of *Drosophila melanogaster*. *BMC Biol* **12**: 61. doi:10.1186/s12915-014-0061-9
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842. doi:10.1093/bioinformatics/btq033
- Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. 2014. DeepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**: W187–W191. doi:10.1093/nar/gku365
- Rauluseviciute I, Riudavets-Puig R, Blanc-Mathieu R, Castro-Mondragon JA, Ferenc K, Kumar V, Lemma RB, Lucas J, Cheneby J, Baranasic D, et al. 2024. JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* **52**: D174–D182. doi:10.1093/nar/gkad1059
- Richardson BE, Lehmann R. 2010. Mechanisms guiding primordial germ cell migration: strategies from different organisms. *Nat Rev Mol Cell Biol* **11**: 37–49. doi:10.1038/nrm2815
- Richter C, Oktaba K, Steinmann J, Müller J, Knoblich JA. 2011. The tumour suppressor L(3)mbt inhibits neuroepithelial proliferation and acts on insulator elements. *Nat Cell Biol* **13**: 1029–1039. doi:10.1038/ncb2306
- Roovers EF, Rosenkranz D, Mahdipour M, Han CT, He N, Chuva de Sousa Lopes SM, van der Westerlaken LA, Zischler H, Butcher F, Roelen BA, et al. 2015. Piwi proteins and piRNAs in mammalian oocytes and early embryos. *Cell Rep* **10**: 2069–2082. doi:10.1016/j.celrep.2015.02.062
- Rosenkranz D, Zischler H. 2012. proTRAC - a software for probabilistic piRNA cluster detection, visualization and analysis. *BMC Bioinformatics* **13**: 5. doi:10.1186/1471-2105-13-5
- Rosenkranz D, Zischler H, Gebert D. 2022. piRNAclusterDB 2.0: update and expansion of the piRNA cluster database. *Nucleic Acids Res* **50**: D259–D264. doi:10.1093/nar/gkac622
- Rust K, Byrnes LE, Yu KS, Park JS, Sneddon JB, Tward AD, Nystul TG. 2020. A single-cell atlas and lineage analysis of the adult *Drosophila* ovary. *Nat Commun* **11**: 5628. doi:10.1038/s41467-020-19361-0
- Ryzansky SS, Kotov AA, Kibanov MV, Akulenko NV, Korbut AP, Lavrov SA, Gvozdev VA, Olenina LV. 2016. RNA helicase Spn-E is required to maintain Aub and AGO3 protein levels for piRNA silencing in the germline of *Drosophila*. *Eur J Cell Biol* **95**: 311–322. doi:10.1016/j.ejcb.2016.06.001
- Saito K, Inagaki S, Mituyama T, Kawamura Y, Ono Y, Sakota E, Kotani H, Asai K, Siomi H, Siomi MC. 2009. A regulatory circuit for piwi by the large Maf gene traffic jam in *Drosophila*. *Nature* **461**: 1296–1299. doi:10.1038/nature08501
- Sato K, Iwasaki YW, Shibuya A, Carninci P, Tsuchizawa Y, Ishizu H, Siomi MC, Siomi H. 2015. Krimper enforces an antisense bias on piRNA pools by binding AGO3 in the *Drosophila* germline. *Mol Cell* **59**: 553–563. doi:10.1016/j.molcel.2015.06.024
- Schaefer CB, Ooi SK, Bestor TH, Bourc'his D. 2007. Epigenetic decisions in mammalian germ cells. *Science* **316**: 398–399. doi:10.1126/science.1137544
- Schmidt D, Wilson MD, Spyrou C, Brown GD, Hadfield J, Odom DT. 2009. CHIP-seq: using high-throughput sequencing to discover protein–DNA interactions. *Methods* **48**: 240–248. doi:10.1016/j.ymeth.2009.03.001
- Schoch CL, Ciufo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, Leipe D, McVeigh R, O'Neill K, Robbertse B,

- et al. 2020. NCBI taxonomy: a comprehensive update on curation, resources and tools. *Database* **2020**: baaa062. doi:10.1093/database/baaa062
- Senti KA, Jurczak D, Sachidanandam R, Brennecke J. 2015. piRNA-guided slicing of transposon transcripts enforces their transcriptional silencing via specifying the nuclear piRNA repertoire. *Genes Dev* **29**: 1747–1762. doi:10.1101/gad.267252.115
- Slaidina M, Banisch TU, Gupta S, Lehmann R. 2020. A single-cell atlas of the developing *Drosophila* ovary identifies follicle stem cell progenitors. *Genes Dev* **34**: 239–249. doi:10.1101/gad.330464.119
- Snee MJ, Harrison D, Yan N, Macdonald PM. 2007. A late phase of Oskar accumulation is crucial for posterior patterning of the *Drosophila* embryo, and is blocked by ectopic expression of Bruno. *Differentiation* **75**: 246–255. doi:10.1111/j.1432-0436.2006.00136.x
- Sumiyoshi T, Sato K, Yamamoto H, Iwasaki YW, Siomi H, Siomi MC. 2016. Loss of l(3)mbt leads to acquisition of the ping-pong cycle in *Drosophila* ovarian somatic cells. *Genes Dev* **30**: 1617–1622. doi:10.1101/gad.283929.116
- Suvorov A, Kim BY, Wang J, Armstrong EE, Peede D, D'Agostino ERR, Price DK, Waddell P, Lang M, Courtier-Orgogozo V, et al. 2022. Widespread introgression across a phylogeny of 155 *Drosophila* genomes. *Curr Biol* **32**: 111–123.e5. doi:10.1016/j.cub.2021.10.052
- Trojer P, Li G, Sims RJ, Vaquero A, Kalakonda N, Bocconi P, Lee D, Erdjument-Bromage H, Tempst P, Nimer SD, et al. 2007. L3MBTL1, a histone-methylation-dependent chromatin lock. *Cell* **129**: 915–928. doi:10.1016/j.cell.2007.03.048
- Webster A, Li S, Hur JK, Wachsmuth M, Bois JS, Perkins EM, Patel DJ, Aravin AA. 2015. Aub and Ago3 are recruited to Nuage through two mechanisms to form a ping-pong complex assembled by Krimper. *Mol Cell* **59**: 564–575. doi:10.1016/j.molcel.2015.07.017
- Williams Z, Morozov P, Mihailovic A, Lin C, Puvvula PK, Juranek S, Rosenwaks Z, Tuschl T. 2015. Discovery and characterization of piRNAs in the human fetal ovary. *Cell Rep* **13**: 854–863. doi:10.1016/j.celrep.2015.09.030
- Yamamoto-Matsuda H, Miyoshi K, Moritoh M, Yoshitane H, Fukada Y, Saito K, Yamanaka S, Siomi MC. 2022. Lint-O cooperates with L(3)mbt in target gene suppression to maintain homeostasis in fly ovary and brain. *EMBO Rep* **23**: e53813. doi:10.15252/embr.202153813
- Yatsu J, Hayashi M, Mukai M, Arita K, Shigenobu S, Kobayashi S. 2008. Maternal RNAs encoding transcription factors for germline-specific gene expression in *Drosophila* embryos. *Int J Dev Biol* **52**: 913–923. doi:10.1387/ijdb.082576jy
- Yu G, Wang LG, He QY. 2015. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**: 2382–2383. doi:10.1093/bioinformatics/btv145
- Yu T, Biasini A, Cecchini K, Säflund M, Mou H, Arif A, Eghbali A, de Rooij D, Weng Z, Zamore PD, et al. 2023. A-MYB/TCFL5 regulatory architecture ensures the production of pachytene piRNAs in placental mammals. *RNA* **29**: 30–43. doi:10.1261/rna.079472.122
- Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based analysis of ChIP-seq (MACS). *Genome Biol* **9**: R137. doi:10.1186/gb-2008-9-9-r137
- Zhang Z, Xu J, Koppetsch BS, Wang J, Tipping C, Ma S, Weng Z, Theurkauf WE, Zamore PD. 2011. Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains. *Mol Cell* **44**: 572–584. doi:10.1016/j.molcel.2011.10.011
- Zhang F, Wang J, Xu J, Zhang Z, Koppetsch BS, Schultz N, Vreven T, Meignin C, Davis I, Zamore PD, et al. 2012. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* **151**: 871–884. doi:10.1016/j.cell.2012.09.040
- Zhang T, Zhu Q, Xie Z, Chen Y, Qiao Y, Li L, Jing N. 2013. The zinc finger transcription factor Ovol2 acts downstream of the bone morphogenetic protein pathway to regulate the cell fate decision between neuroectoderm and mesendoderm. *J Biol Chem* **288**: 6166–6177. doi:10.1074/jbc.M112.418376
- Zhu LJ, Christensen RG, Kazemian M, Hull CJ, Enameh MS, Basciotta MD, Brasefield JA, Zhu C, Asriyan Y, Lapointe DS, et al. 2011. Flyfactorsurvey: a database of *Drosophila* transcription factor binding specificities determined using the bacterial one-hybrid system. *Nucleic Acids Res* **39**: D111–D117. doi:10.1093/nar/gkq858



Transcriptional regulation of the piRNA pathway by Ovo in animal ovarian germ cells

Azad Alizada, Gregory J. Hannon and Benjamin Czech Nicholson

Genes Dev. 2025, **39**: originally published online December 20, 2024
Access the most recent version at doi:[10.1101/gad.352120.124](https://doi.org/10.1101/gad.352120.124)

Supplemental Material

<https://genesdev.cshlp.org/content/suppl/2024/12/17/gad.352120.124.DC1>

References

This article cites 108 articles, 26 of which can be accessed free at:
<https://genesdev.cshlp.org/content/39/3-4/221.full.html#ref-list-1>

Creative Commons License

This article, published in *Genes & Development*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

