L(3)mbt safeguards tissue identity by maintaining transcriptional integrity at head-to-head genes

Rémi-Xavier Coux, Felipe Karam Teixeira, Ruth Lehmann

Howard Hughes Medical Institute (HHMI) and Kimmel Center for Biology and Medicine of the Skirball Institute, Department of Cell Biology, New York University School of Medicine, New York, NY 10016, USA

Running Title: L(3)mbt safeguards tissue identity

Keywords: L(3)mbt; Nanos; Drosophila; oogenesis; insulator protein; Topologically Associated Domains (TADs).

Abstract

Maintenance of cellular identity is essential for development and homeostasis and problems in such control can have dire consequences such as organ malformation or cancer. At the molecular level, cell identity is determined by the coordinated activation and repression of defined sets of genes. The tumor suppressor L(3)mbt was shown to secure cellular identity in Drosophila larval brains by repressing germline specific genes, although the mechanisms remain elusive. Here we show that L(3)mbt safeguards the integrity of the somatic and germline tissues forming the Drosophila ovary. *L(3)mbt* mutation causes multiple developmental defects and simultaneous expression of original identity markers and new tissue-specific misexpression signatures. We demonstrate that the ovarian developmental defects are largely due to the ectopic expression of *nanos,* a key regulator of germline fate in the somatic cells of the ovary, rather than to the loss of cellular identity. Importantly, using genome-wide analysis we found that L(3)mbt exerts boundary activity between differentially regulated genes by binding to high occupancy insulator sites enriched at the borders of topological domains. Using *nanos* as an example, we propose that L(3)mbt maintains cell fate and specific gene expression programs by locally insulating repressed genes from closely apposed transcribed neighbors.

Introduction

Development requires tight control of gene expression as differentiating cells must express lineage-specific genes while repressing genes promoting other fates. Mechanisms ensuring the maintenance of cellular identity must be robust, as changes in cell fate are fairly uncommon in wild-type conditions, with only one documented case of regulated transdifferentiation being described in *C. elegans* (Jarriault et al. 2008). Rare cases of transdifferentiation have been observed in mutants affecting chromatin complexes, suggesting a role for chromatin structure in the maintenance of cellular identity (Tursun et al. 2011; Petrella et al. 2011). A notable example is given by mutations affecting the Drosophila *lethal (3) malignant brain tumor* (*l(3)mbt*) gene, which lead to malignant brain tumors that ectopically express germline-specific genes and have been proposed to cause soma-to-germline transformations (Gateff et al. 1993; Janic et al. 2010). Gene expression profiling of *l(3)mbt* brain tumors and L(3)mbt-depleted cultured somatic cells identified a group of upregulated genes known as the Malignant Brain Tumor Signature (MBTS) that is enriched for factors specifically expressed in germ cells (Georlette et al. 2007; Janic et al. 2010; Meier et al. 2012; Sumiyoshi et al. 2016). Mutations of germline-specific genes, including those impairing the piRNA factors *piwi*, *aub*, and *vasa*, as well as the translational repressor *nanos*, were found to suppress the neural overgrowth induced by loss of L(3)mbt (Janic et al. 2010). However, another study suggested that the Hippo growth control pathway is critical for tumor formation in *l(3)mbt* mutant brains, and questioned tissue transformation as the cause of tumorigenesis (Richter et al. 2011). Furthermore, our lab showed that strong *l(3)mbt* mutations cause a maternal, germline autonomous effect phenotype that precludes normal embryonic development including primordial germ cell formation (Yohn et al. 2003), suggesting a broader role for L(3)mbt in regulation of tissue identity.

L(3)mbt encodes a 1477 amino-acid protein that is ubiquitously expressed in Drosophila and is conserved from worms to humans. L(3)mbt harbors three MBT repeats that were shown to bind to mono- and possibly di-methylated lysine 20 of Histone 4 *in vitro* (Bonasio et al. 2010). L(3)mbt also contains a C_2/HC Zinc-Finger domain and a carboxy-terminal region thought to mediate protein-protein interactions (Bonasio et al. 2010). Notably, L(3)mbt was purified in two non-enzymatic repressive chromatin complexes: the Drosophila RBF, E2F2, and Myb-interacting proteins (dREAM complex, also called Myb-Muv B) as well as the L(3)mbt-Interacting complex (LINT) (Lewis et al. 2004; Meier et al. 2012). DREAM is a large multi-subunit complex that controls gene expression throughout the cell cycle but also represses developmental genes. L(3)mbt associates at sub-stoichiometric levels with dREAM and is strictly found in its repressive forms (Georlette et al. 2007; Lewis et al. 2004). The LINT complex is composed of L(3)mbt, the novel transcriptional repressor Lint1, as well as the corepressor CoREST, and was shown to silence developmental genes in cultured cells (Meier et al. 2012). Interestingly, dREAM and LINT complexes repress overlapping sets of genes in somatic cells, including genes normally expressed in the germline.

Despite extensive biochemical studies, we still know little about which chromatin complex mediates L(3)mbt's role in tissue identity. L(3)mbt binding sites were characterized in larval brains and embryonic cells by Chromatin Immuno-purification followed by sequencing (ChIP-seq, (Richter et al. 2011; Van Bortle et al. 2012)). These studies revealed that L(3)mbt binds approximately 3500 regions and is enriched at the promoters of repressed genes, suggesting a direct role in transcriptional repression. On the other hand, a large number of L(3)mbt binding sites

overlaps with regions occupied by the insulator proteins CP190, BEAF-32 and to a lesser extent CTCF as well as Su(Hw), which led to an alternative model in which L(3)mbt works as an insulator accessory factor (Richter et al. 2011; Van Bortle et al. 2012; 2014).

To further understand how L(3)mbt secures tissue identity, we combined genetic and genomic approaches to characterize the functions of L(3)mbt in *Drosophila melanogaster* ovarian development. We find that L(3)mbt affects gene expression in a tissue-specific manner, in somatic cells of the ovary it represses germline genes and in the female germline it targets genes normally expressed in the testis and the nervous system. Mutant somatic ovarian tissues continue to express signatures of the tissue of origin, indicating that loss of L(3)mbt does not induce transdifferentiation. Consistent with a primary role of repressing the "germline program" in the soma, ectopic expression of a single gene, the translational repressor and key regulator of germline fate *nanos*, is largely responsible for aberrant soma development. Analysis of L(3)mbt associated proteins reveals a primary function for the LINT complex in the ovary. We observed that L(3)mbt bound regions required for regulation are high-occupancy insulator binding sites found at the boundaries of chromosomal domains. Finally, we demonstrate that L(3)mbt depletion results in loss of boundary between differentially-regulated head-to-head genes, allowing expression of genes incompatible with normal ovarian development.

Results

Sterility in *l(3)mbt* mutant females is associated with aberrant ovarian development

L(3)mbt was previously shown to be required for the development of the nervous system, and *l(3)mbt* mutant flies grown at restrictive temperatures (29˚C) develop malignant brain tumors and

die at larval stages (Gateff et al. 1993; Janic et al. 2010; Richter et al. 2011). We observed that, when grown at lower temperature (25˚C), null *l(3)mbt* mutants were viable but fully sterile, indicating that L(3)mbt is critically required for germline development. In the wild type, each ovary is composed of 16 to 20 egg assembly chains called ovarioles (Fig. 1A). At the tip of each ovariole in a region called the germarium, germline stem cells (GSCs) divide asymmetrically to generate a new GSC and a differentiating daughter cell. The differentiating GSC daughter undergoes four rounds of mitosis with incomplete cytokinesis to form a sixteen-cell germline cyst. Sibling germ cells remain interconnected through cytoplasmic bridges called ring canals. Only one of the cyst germ cells develops into an oocyte, while the other 15 cells become supportive nurse cells. The somatic cells of the ovary play important roles in supporting oogenesis: they compose the GSC niche, promote cyst divisions and differentiation, and the follicle cells enclose and individualize egg chambers, contributing to oocyte-nurse cell differentiation. At the macroscopic level, *l(3)mbt* mutant ovaries were atrophied and adult females did not lay eggs. To characterize the ovarian phenotype in detail, we used antibodies against the germline marker Vasa and α-Spectrin (α-Spe), which labels the membranes of somatic cells and distinguishes between undifferentiated germ cells and connected cysts. Similar to the wild-type, *l(3)mbt* mutant germaria contained GSCs adjacent to the somatic niche (Supplemental Fig. S1A-B). In contrast to the wild type, mutant ovarioles contained fewer individualized egg chambers (1.35 egg chambers/ovariole in average vs. 6 in WT, $p<10^{-4}$), which were highly abnormal and displayed several defects (Fig. 1C, quantified in J). These included extra-numerous germ / nurse cells (Fig. 1C), somatic follicle cells that failed to fully enclose germ cells (Fig. 1E, arrowhead), and accumulation of undifferentiated germ cells marked by branched fusomes, a structure stained by α-Spe that connects sibling germ-cells within early cysts (Fig. 1G, arrow). The majority of mutant egg

chambers contained multiple oocytes, as indicated by oo18 RNA-binding protein (Orb, (Christerson and McKearin 1994)) staining and ring canal numbers, suggesting that *l(3)mbt* mutant ovarioles undergo abnormal egg chamber separation or fusion (Fig. 1H-I, arrowheads; Supplemental Fig. S1C-D). In rare cases (2%), we observed multiple germaria connected to the same aberrant egg chamber (Supplemental Fig. S1E-F), suggesting that ovariole fusion can also occur. Taken together, our results indicate that in addition to its previously reported requirement for brain and embryonic development, L(3)mbt is critical for proper ovary development.

L(3)mbt functions in ovarian somatic cells to safeguard ovary development

Previous experiments had shown that loss of *l(3)mbt* in germ cells caused developmental defects in the resulting embryos, but allowed oocytes to mature (Yohn et al. 2003). Thus, we wondered whether the gross abnormalities of mutant ovaries were indicative of a role for L(3)mbt in somatic cells of the ovary. We induced homozygous $l(3)mbt^{GM76}$ mutant clones in the somatic ovarian cells by using the FRT-FLP system (Harrison and Perrimon 1993) under the transcriptional control of the *Ptc-Gal4* or *c587-Gal4* drivers, which drive expression in the somatic cells of the germarium (Hinz et al. 1994; Zhu and Xie 2003). *L(3)mbt* mutation in a subset of somatic cells was sufficient to perturb germline development (Fig. 2A) and led to egg chambers containing multiple oocytes (Fig. 2B). To further test a requirement for L(3)mbt in somatic cells, we generated an inducible *UAS-l(3)mbt::myc* transgene. Its expression, specifically in somatic ovarian cells using the tj-Gal4 driver, was sufficient to rescue the aberrant morphology of mutant ovaries (Fig. 2C-D and Supplemental Fig. S2A-C). Embryos laid by mutant females somatically expressing *l(3)mbt::myc* failed to hatch, in agreement with the described maternal requirement for L(3)mbt (Yohn et al. 2003). While most egg chambers developed normally in somatically-complemented *l(3)mbt*

mutant ovaries, we noticed that 69% of ovarioles contained apoptotic egg chambers, identifying a germline autonomous role for L(3)mbt in egg chamber development (Supplemental Fig. S2C-D). Our results demonstrate that in addition to its germ cell autonomous role, L(3)mbt is required in the somatic tissues of the ovary to support normal oogenesis.

L(3)mbt mutant somatic cells simultaneously express somatic gonad and germline-specific genes

Previous studies suggested that loss of L(3)mbt resulted in a soma to germline transformation in the larval brain and cultured somatic cell (Janic et al. 2010; Georlette et al. 2007; Meier et al. 2012; Sumiyoshi et al. 2016). Therefore, we performed RNA-sequencing (RNA-seq) analysis to gain a genome-wide view of gene expression changes induced by loss of L(3)mbt in ovarian somatic cells *in vivo*. To distinguish between germline and somatic tissues of the ovary, we took advantage of *tud* maternal mutations (tud^{M}), which give rise to progeny that lack germ cells and develop into adults devoid of germline ((Arkov et al. 2006; Smendziuk et al. 2015), see Supplemental File S1 and Materials and Methods). Comparisons between tud^{M}; $l(3)mbt^{GM76}/l(3)mbt^{Df}$ and tud^{M}; $l(3)mbt^{GM76}/+$ control ovaries identified 600 genes differentially expressed (*p-value*<0.05) in somatic cells of the ovary. Of these 459 were upregulated and 141 downregulated in mutant somatic ovarian tissues (Supplemental Table S1). 44 upregulated genes were shared with the 101 MBTS genes (Janic et al. 2010), and 115 out of the 681 genes found upregulated in *Δ–l(3)mbt* Ovarian Somatic Cells (OSC; Sumiyoshi et al., 2016). These derepressed genes include piRNA pathway components and germline-specific genes such as *nos, Pxt, vas, aub, tej, krimp, AGO3,* and *CG9925* (Fig. 3A, Supplemental Fig. S3A). The effect of *l(3)mbt* mutation on gene expression was more pronounced for repressed genes, while genes normally expressed in the soma showed

only low fold-changes in the mutant (124/141 had a Log_2 fold-change between -0.3 and -1). This suggested that multiple cell identities may coexist in *l(3)mbt* mutant tissues. Indeed, key somatic ovarian transcription factors such as Traffic Jam (TJ) and Zfh1, which are essential for gonad development and exclusively expressed in somatic cells (Leatherman and Dinardo 2008; Li et al. 2003), were still expressed in *l(3)mbt* mutants (Fig. 3B-C and Supplemental Fig. S3B-C). Despite the gross morphological abnormalities, we observed TJ expressing cells with epithelial character, which resembled wild-type follicle cells, surrounding germ cells in mutant ovaries. However, individual TJ-positive cells occasionally expressed the germline marker Vasa (Fig. 3C, arrows). As TJ and Vasa expressions are mutually exclusive in the wild type, our results indicate that *l(3)mbt* mutant somatic cells retain somatic and epithelial/follicular characteristics while ectopically expressing hallmark genes of the germline fate.

Expression of Nos is necessary and sufficient to cause ovarian defects

Aberrant growth of *l(3)mbt* brain tumors was shown to rely on the ectopic expression of *nos, aub,* and *vasa* (Janic et al. 2010). We therefore asked whether misexpression of any of these genes contributed to the ovarian defects observed in *l(3)mbt* mutants. We found that *aub, l(3)mbt* double mutant ovaries were phenotypically similar to *l(3)mbt* single mutant, containing many apoptotic cells and aberrant egg chambers with more than 16 germ cells (Supplemental Fig. S4A) (we were unable to test *vas, l(3)mbt* double mutants because of lethality). In contrast, nos^{BN}/nos^{L7} mutations dramatically suppressed the *l(3)mbt* ovarian defects: *nos, l(3)mbt* double mutant ovaries contained late stages egg chambers, their ovarioles were phenotypically indistinguishable from wild type and 85% of double mutant egg chambers contained 16 germ cells and only one oocyte (Fig. 4C-E, n=88). In agreement with this, depletion of Nanos' co-factor Pumilio in *l(3)mbt* mutant ovaries

also suppressed the mutant phenotypes, with 81% of *pum*^{680}, *l(3)mbt* double mutant ovarioles resembling wild-type morphology (Fig. 4F-H, n=181). To test if *nos* misexpression in somatic ovarian cells was sufficient to cause defects similar to those observed in *l(3)mbt* mutants, we expressed *nos* under tj-Gal4-Gal80^{ts} control (McGuire et al. 2004) to restrict expression in the soma to adult stages. *Nos* somatic expression did not fully recapitulate the *l(3)mbt* phenotype but perturbed ovarian morphology resulting in poorly individualized egg chambers (Supplemental Fig. S4B). In contrast, ectopic expression of *aub* or *vas* in somatic ovarian cells did not yield a morphologically significant phenotype (Supplemental Fig. S4C, D). Together, these results suggest that ectopic expression of *nos*, but not *aub or vas*, is necessary and sufficient to perturb somatic ovarian development.

L(3)mbt functions through the LINT complex to secure ovary development

L(3)mbt has been associated with two chromatin complexes that repress developmental and germline genes: the dREAM and LINT complexes (Fig. 5A-B; (Georlette et al. 2007; Meier et al. 2012)). To determine the function of these complexes in ovarian development, we depleted somatic cells from E2f2 and Mip120, two core repressors of the dREAM complex. Mutant somatic clones for *E2f2* (Fig. 5C) or ovaries deficient for *mip120* (Fig. 5E) did not result in phenotypic aberrations reminiscent of those observed in *l(3)mbt* mutants. We also asked whether *mip120* mutation affected L(3)mbt nuclear localization in the ovary, as previously described in salivary glands (Blanchard et al. 2014). Using the TdTomato::L(3)mbt fusion protein we observed L(3)mbt nuclear localization comparable to wild type in *mip120*$^{67.9A.9}$ mutants (Supplemental Fig. S5B). Together, these results suggest that L(3)mbt's critical function in somatic ovarian cells is independent of the dREAM complex.

Next we examined the role of the LINT complex as a mediator of L(3)mbt function. Since mutations in Lint1 had not been identified, we generated CRISPR-induced *Lint1* alleles (Gratz et al. 2013; 2014). *Lint1*^{1} deletes two cytosines (350 and 351), creating a premature stop codon at position 66/540 (Lint1-C) or 128/601 (Lint1-A, Fig. 5F). Homozygous mutant flies were viable and fertile at 25˚C, and their ovaries developed normally although 2% of egg chambers contained misplaced or extra-numerous oocytes (Fig. 5G). However, when grown at 29˚C, *lint1*^{1} females were fully sterile, laying eggs that failed to hatch, and 15% of their ovarioles contained aberrant egg chambers reminiscent of those observed in *l(3)mbt* mutant (Fig. 5H). We conclude that L(3)mbt may function via the LINT complex for ovary development.

In the female germline, L(3)mbt is required to represses neuronal- and testis-specific genes

Our data suggest a role for L(3)mbt in suppressing germline specific genes in somatic tissues of the ovary. However, L(3)mbt is also autonomously required in the germline for embryo development (Yohn et al. 2003) and egg chamber survival (Supplemental Fig. S2D). To gain a genome-wide view of changes in gene expression induced by the germline loss of L(3)mbt, we performed RNA-seq analysis on embryos laid by *l(3)mbt* mutant or heterozygous mothers expressing the L(3)mbt::myc fusion in the somatic ovary. We collected early embryos prior to activation of the zygotic genome as the early embryonic RNA pool is exclusively composed of maternally provided transcripts and reflects the germline expression profile during oogenesis (Edgar and Schubiger 1986). Our analysis identified 878 differentially expressed genes (p<0.05), of which 342 were upregulated and 536 downregulated in embryos laid by mutant females (Supplemental Table S2). Interestingly, we observed only a limited overlap between the genes

derepressed in *l(3)mbt* mutant somatic and germline ovarian tissues, with two thirds of genes being specifically upregulated in one of the two tissues (Fig. 6A). To better characterize the group of genes upregulated in the *l(3)mbt* mutant germline, we performed two-way hierarchical clustering to identify any tissue specific expression signatures. Two major groups were readily identified: one composed of genes highly expressed in neuronal tissues (*brain, thoracicoabdominal ganglion (tag), larval CNS, eye,* and *head*; Fig. 6B) and another comprising testis-specific genes. As in mutant somatic ovaries, two-thirds (365/536) of the downregulated genes had their expression level reduced by less than two-fold compared to the heterozygous controls (Supplemental Fig. S6). These results indicate that L(3)mbt's function is not restricted to repressing the germline program, but that L(3)mbt regulates distinct sets of genes, in a tissue-specific manner.

L(3)mbt-bound regions are high-occupancy insulator sites

L(3)mbt has long been thought to act as a transcriptional repressor (Georlette et al. 2007; Meier et al. 2012). However, L(3)mbt-binding sites were recently found to overlap with insulator elements, suggesting that L(3)mbt may harbor insulating function (Richter et al. 2011). These models are not mutually exclusive, so we asked whether L(3)mbt bound regions show characteristics of insulator elements. We observed that in cultured Ovarian Somatic Cells (OSC), L(3)mbt binding sites coincide with open and transposon-accessible chromatin and are devoid of nucleosomes (Supplemental Fig. S7A-B). Using published datasets (Li et al. 2015; Van Bortle et al. 2014), we further determined that L(3)mbt distribution resembles the binding profiles of the canonical insulators CP190 and BEAF32, being enriched upstream of Transcription Start Sites (TSS) (Supplemental Fig. S7C) and associated with 5' Nucleosome Free Regions that are flanked by nucleosomes containing the histone variant H2Av (Supplemental Fig. S7D) (Raisner et al. 2005;

Soboleva et al. 2014). However, gene expression changes after CP190 (Ali et al. 2016) or Beaf32 depletion (Lhoumaud et al. 2014) do not parallel those observed after *l(3)mbt* knockdown (Meier et al. 2012) in Kc167 cells (Supplemental Fig. S7E-F). Further, our ATAC-seq data showed that nucleosome depletion at regions normally bound by L(3)mbt was not affected in *Δ–l(3)mbt* OSCs (Supplemental Fig. 7B, (Sumiyoshi et al. 2016; Schep et al. 2015)). Therefore, the effect of L(3)mbt on gene repression is not solely mediated by the canonical insulators CP190 and BEAF32 and other factors must mediate local chromatin remodeling at L(3)mbt-binding sites independently of L(3)mbt.

In higher eukaryotes, chromosome three-dimensional structure is organized into topologically associated domains (TADs, (Sexton et al. 2012; Hou et al. 2012; Van Bortle et al. 2014)). Recent studies suggest that co-operative binding of insulators and additional architectural factors such as condensin and cohesin proteins defines insulator strength. Specifically, high-occupancy sites (bound by 7 or more factors) often cluster at TAD boundaries and exhibit robust insulator activities, while low-occupancy sites tend to regulate gene expression more locally (Van Bortle et al. 2014). To investigate binding of insulators and associated factors at L(3)mbt sites upstream regulated genes, we compared the previously generated matrix of architectural proteins with the genes derepressed in L(3)mbt depleted Kc167 cells (Van Bortle et al. 2014; Meier et al. 2012). On average, these functionally relevant L(3)mbt sites were occupied by 6.8 factors, notably including the insulators CP190 (88% of overlap), Chromator (83% overlap), and BEAF32 (69% overlap; Fig. 7A). Of note, Lint1 also occupies 60% of these high occupancy sites (Supplemental Table S3). Next, we investigated the distribution of derepressed genes in *l(3)mbt* mutant somatic ovarian cells in relation to TADs identified in OSC cells (Ulianov et al. 2016). We observed a striking

positional enrichment for L(3)mbt-regulated genes at the borders of TADs (Fig. 7B). Notably, *nos* is located at the boundary of a TAD in OSC cells (Fig. 7C). We conclude that L(3)mbt bound regions required for proper gene regulation are preferentially high-occupancy insulator sites enriched at TAD boundaries.

L(3)mbt exerts boundary activity between differentially regulated head-to-head genes

The genome of *Drosophila melanogaster* is compact and around 32% of its genes are divergently-paired, *i.e*. in head-to-head orientation with less than 1 kb separating the adjacent TSS (Yang and Yu 2009). Insulator proteins are often enriched between divergent pairs of genes where they act as boundaries between differentially regulating genes (Nègre et al. 2010; Yang et al. 2012). To investigate an insulating role for L(3)mbt, we examined the orientation of repressed genes relative to their adjacent neighbors. In *l(3)mbt* mutant somatic ovarian cells, 66 out of 97 (69%) upregulated genes (Log_2 fold change >1.5) were in head-to-head (HtH) orientation with the upstream adjacent gene (Fig. 7D, "opposite orientation"), which significantly differs from an expected random distribution ($p<5x10^{-4}$). In contrast, genes directly downstream of derepressed genes were randomly oriented, with 50.6% of the pairs in a tail-to-tail conformation and 49.6% tail-to-head (p=0.99). Interestingly, we observed a comparable orientation bias for genes upregulated upon L(3)mbt depletion in other tissues (germline: 73% HtH, $p<10^{-4}$; brain: 72% HtH, $p<10^{-4}$ (Janic et al. 2010); and Kc167 cells: 68% HtH, $p<10^{-4}$ (Meier et al. 2012)). Next, we asked whether L(3)mbt binds between head-to-head targets. Indeed, 186 out of 235 (79%) regulated head-to-head pairs affected in the somatic ovary contain L(3)mbt binding sites in their Upstream Intergenic Region (UIR) and/or 5'UTRs (Richter et al. 2011). The proportion of pairs containing

L(3)mbt sites was of 37% (60/164) for other conformations, indicating an enrichment for L(3)mbt binding in between head-to-head targets.

Head-to-the genes share their UIRs, which include most of the transcriptional regulatory sequences. Consequently, close HtH genes (UIR<500bp) tend to be co-expressed (Yang et al. 2012), as confirmed by genome-wide analysis using gene expression data across developmental stages and cell lines (all HtH pairs; Fig. 7E). In contrast, we observed that the divergent HtH genes regulated by L(3)mbt in the germline or in somatic tissues showed no correlation of expression (Fig. 7E). This striking difference is not related to the length of UIRs as UIRs between L(3)mbt-regulated divergent pairs were similar to all HtH pairs, peaking around 250 base pairs (Supplemental Fig. S7G). As L(3)mbt binds to the UIRs of differentially-regulated, head-to-head genes and its loss causes mis-expression of such genes, our data favor a model in which L(3)mbt functions as a boundary factor between differentially-expressed closely apposed genes.

To test this model, we analyzed L(3)mbt occupancy in the UIR shared by *nos* and the adjacent HtH gene *CG11779*. Using the tdTomato::L(3)mbt fusion protein, we performed tissue-specific chromatin immunoprecipitation followed by quantitative PCR (ChIP-qPCR, Fig. 7F). In the germline, where both *nos* and *CG11779* are expressed, L(3)mbt signal was not detected throughout the region spanning the shared UIR. However, in ovarian somatic tissues, where *CG11779* is expressed but *nos* is not, L(3)mbt signal was robustly detected over the previously identified L(3)mbt-binding site. Given that loss of L(3)mbt in ovarian somatic cells leads to *nos* ectopic expression, our results support the hypothesis that L(3)mbt operates as an insulator to safeguard tissue identity.

Discussion

By combining developmental and molecular analysis, we show that *l(3)mbt* mutant ovaries develop aberrantly. L(3)mbt depletion does not result in transdifferentiation, but causes derepression of sets of genes incompatible with original cell identities. Notably the translational repressor *nanos* becomes derepressed in the somatic cells of the ovary and causes aberrant growth. Molecularly, we demonstrate that L(3)mbt functions through the LINT complex in the somatic ovary. Finally, we show that L(3)mbt exerts boundary activity between differentially-regulated genes transcribed in opposite orientation. We propose that this function of L(3)mbt mediates its role as guardian of cell identity.

Loss of L(3)mbt causes the ectopic expression of a number of genes including instructive regulators that, despite continued presence of initial cell fate determinants, are sufficient to interfere with original cellular function and affect tissue development and homeostasis. A prominent example is the germline specific gene *nanos*. Ectopic expression of Nanos has been shown to be detrimental in somatic tissues, including the larval brain (Janic et al. 2010) and somatic ovarian cells (this study). Importantly, in contrast to a previously suggested soma-to-germline transdifferentiation, our results favor the hypothesis that *l(3)mbt* mutation imbalances tissue-homeostasis whereby normally mutually exclusive lineage determinants are simultaneously expressed. In support, L(3)mbt depletion in neuronal, somatic ovarian, and germ cells does not lead to loss of original tissue-specific markers, but genes characteristic of other lineages are derepressed (Richter et al. 2011, this study). Moreover, our results suggest that the role of L(3)mbt is not solely restricted to prevent ectopic expression of germline genes, but instead L(3)mbt

represses distinct, broader sets of genes in a tissue-specific manner. Therefore, we propose that L(3)mbt secures tissue identity by stabilizing gene expression profiles established during differentiation.

Our experiments demonstrate that *nanos* ectopic expression is necessary and sufficient to induce aberrant development of *l(3)mbt* mutant ovaries. These defects are likely not due to a direct interference at the transcriptional level but are rather caused by Nos' function as a translational repressor. In support, we find that Pumilio, a sequence-specific translational repressor and co-factor of Nos, is also essential for the *l(3)mbt* ovarian phenotype. Nos was recently shown to modulate Pum RNA-binding and target-specificity in somatic S2 cells (Weidmann et al. 2016). Since Pum is ubiquitously expressed, we propose that ectopic Nos stabilizes Pum binding at target mRNAs essential for somatic functions. Interestingly, ectopic expression of NANOS1 was found to be required for growth of human *pRb* deficient tumor cells. In this case, NANOS1 and PUM repress p53 translation allowing cells to bypass apoptosis (Miles et al. 2014). Thus, ectopic Nos-Pum complexes may alter tissue homeostasis at the post-transcriptional level in other systems as well. We would expect to find key regulators of somatic fate among mRNAs aberrantly targeted by Nos-Pum in somatic Drosophila tissues.

In contrast to the widespread effects of *nos* derepression observed in multiple somatic tissues upon loss of L(3)mbt, ectopic expression of additional genes may define the exact phenotypic consequences, which depend on tissue type. For example, piRNA pathway genes are ectopically expressed in *l(3)mbt* larval brain tumors and somatic ovarian cells, however, depleting them ameliorates the brain tumor but not the ovarian phenotype. This difference may be explained by

the fact that the somatic ovary uses core components of the piRNA pathway to regulate transposable elements (Handler et al. 2013). Consistent with our finding that in *l(3)mbt* mutants new and original tissue identities are co-expressed, these results suggest that the phenotypic consequences of *l(3)mbt* mutation depend on the context of the original tissue identity.

Our genetic analysis revealed that L(3)mbt functions with Lint1 in the somatic ovary. Lint1 likely contributes to L(3)mbt insulating functions. It co-occupies many L(3)mbt-bound high occupancy insulator sites and co-purifies with L(3)mbt. Further, like L(3)mbt, derepressed genes in Lint1 depleted Kc167 cells (Meier et al. 2012) show a bias towards head-to-head conformation. We propose that Lint1 functions as an architectural protein and contributes to L(3)mbt boundary properties in the somatic ovary. However, the phenotype of *lint1* loss of function is weaker than the *l(3)mbt* phenotype suggesting that other protein complexes contribute to L(3)mbt boundary function. L(3)mbt also associates with the dREAM complex, some components of which have been shown to cooperate with insulator proteins to maintain transcriptional integrity at a subset of divergently-paired genes in Kc167 cells (Bohla et al. 2014; Korenjak et al. 2014). Further, E2f2 binding sites that possess enhancer-blocking properties are all co-occupied by L(3)mbt in Kc cells (Korenjak et al. 2014). Nevertheless, we find that the dREAM complex is dispensable for L(3)mbt boundary activities in the ovary. Thus composition and functionality of complexes associated with L(3)mbt vary in different cell types. It will be interesting to determine how L(3)mbt is recruited to different insulator complexes and how functionality of these complexes is regulated site- and tissue-specifically.

Our results support the idea that L(3)mbt functions as an insulator-associated protein and further show that it exerts insulating functions, in particular at differentially expressed, juxtaposed genes. However, L(3)mbt properties differ from canonical insulator proteins. For example, L(3)mbt does not mediate nucleosome remodeling at its binding sites, indicating that additional factors independently organize nucleosomes at these regions. These factors are not sufficient to maintain gene expression status in the absence of L(3)mbt. Further, the insulator proteins BEAF-32 and Cp190 co-occupy most of L(3)mbt's binding sites but they do not share transcriptional targets. This suggests that L(3)mbt has insulating properties that might be independent of these insulator complexes. While the principles governing which genes are regulated by L(3)mbt in each tissue remain elusive, we find that L(3)mbt-regulated genes are preferentially found near TAD boundaries. Notably, the L(3)mbt targets *nanos* and *vasa* are localized at the border of TADs in multiple cell types (Ulianov et al. 2016). Interestingly, housekeeping genes cluster at TAD boundaries, and *TfIIs* and *CG11779* – the HtH genes upstream of *nos* and *vas* fall in that category (Hou et al. 2012; Hug et al. 2017). We therefore propose that L(3)mbt function is to locally insulate developmental genes from adjacent, highly and promiscuously expressed genes. Considering that CTCF is the only insulator protein identified in mammalian genomes to this date, it will be interesting to investigate whether L(3)mbt mammalian homologues also have the ability to mediate insulating functions.

Figure 1

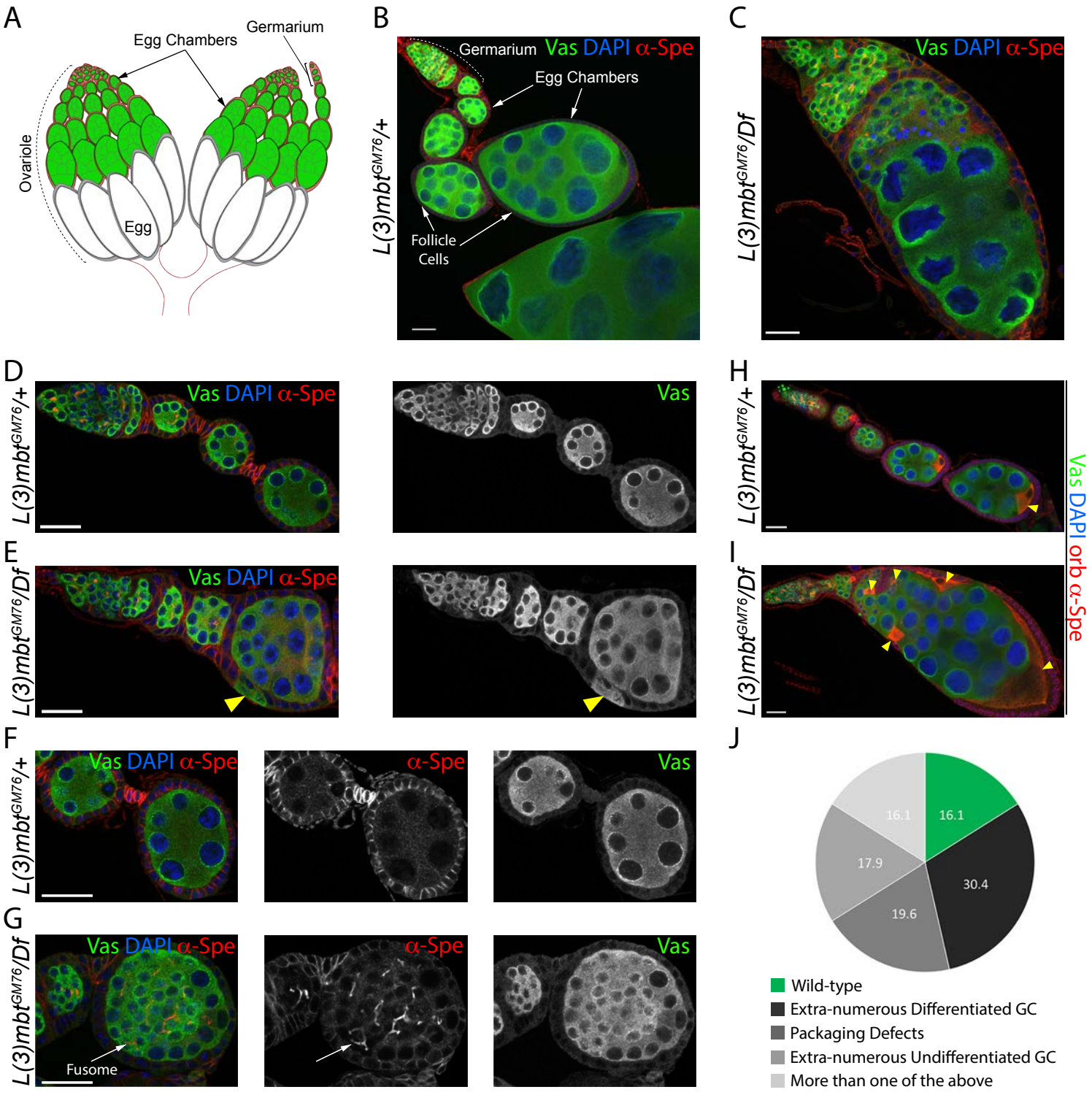
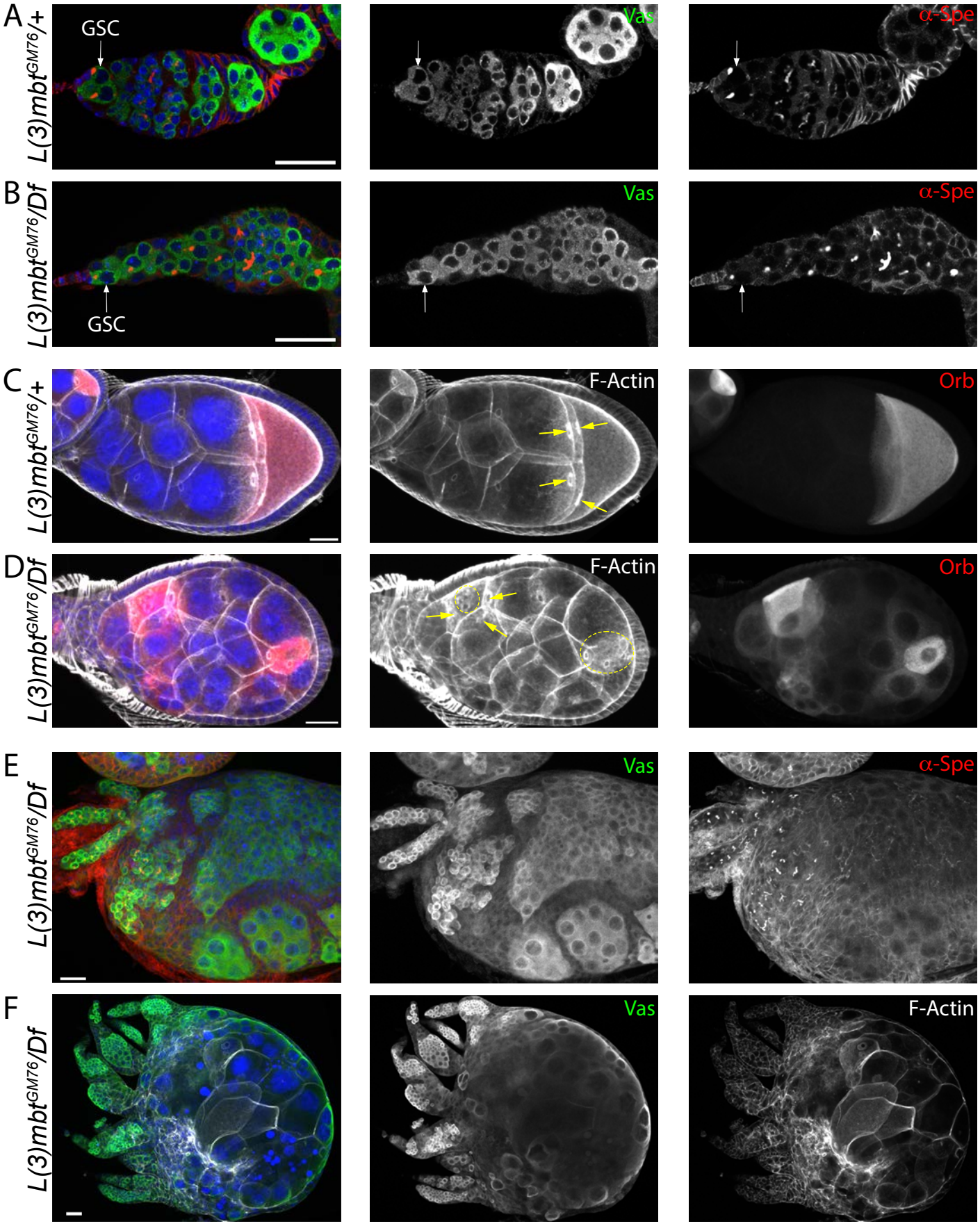


Figure 1. Developmental defects of *l(3)mbt* mutant ovaries. *(A)* Schematic representation of a wild-type ovary composed of ovarioles. *(B-G)* Confocal images of control and *l(3)mbt* mutant ovarioles stained for germ cells (Vasa, green), α-Spectrin (red), and DAPI (blue) for DNA. All ovarioles images are displayed with anterior oriented to the top-left corner. *(B)* Heterozygous control ovariole. *(C)* Representative *l(3)mbt* mutant ovariole with undifferentiated and differentiated germ cells surrounded by follicle cells. *(D)* Tip of wild-type ovariole with germarium and early egg chambers. *(E)* Mutant ovariole with germline packaging defects showing Vasa-expressing germ cells intercalated between follicle cells (arrowhead). *(F)* Wild-type stage 3 and 4 egg chambers. Germ cells within egg chamber are no longer connected by fusomes. *(G)* Similarly staged mutant egg chamber filled with fusome-containing undifferentiated germ cells (arrow). *(H-I)* Confocal images of control and mutant ovarioles stained for Vasa (green), Orb (oocyte marker) and α–Spectrin (red) and DAPI (blue). *(I)* In control ovarioles, Orb is restricted to the developing oocyte at the posterior of egg chambers. *(J) l(3)mbt* mutant ovariole with egg chamber containing more than 16 germ cells and multiple oocytes, as revealed by Orb staining (arrowheads). *(J)* Quantification of phenotypes observed in *L(3)mbt* mutant ovarioles as illustrated in (*C, E and G*). Scale bars, 25 μm.

Supplemental figure S1



Supplemental figure S1. L(3)mbt loss-of-function leads to egg chamber and ovariole fusions. *(A-B)* Confocal images of wild-type *(A)* and mutant *(B)* germaria, stained for Vasa (green), α-Spectrin (red) and DAPI (blue). Wild-type and mutant germaria contain Germline Stem Cells marked by punctuated fusomes (α-Spectrin), in contact with the somatic niche. *(C-D)* Confocal projections of (*C)* wild-type and (*D)* mutant egg chambers stained for F-Actin (grey), Orb (red), and DAPI (blue). Anterior is oriented to the left. (C) a wild-type oocyte is connected to nurse cells by four ring canals (yellow arrows). (*D*) a mutant egg chamber containing multiple oocytes with four or more ring canals (arrows and dotted circles). (*E)* Confocal image of *l(3)mbt* mutant ovarioles stained for Vasa (green), α-Spectrin (red) and DAPI (blue), medium magnification. Three germaria (top left) fused into an aberrant ovariole with intermingled differentiated and undifferentiated germ cells. *(E) l(3)mbt* mutant ovarioles with multiple germaria connected to the same giant egg chamber. Vasa (green), F-Actin (grey) and DAPI (blue), low magnification. Scale bars, 25 μm.

Figure 2

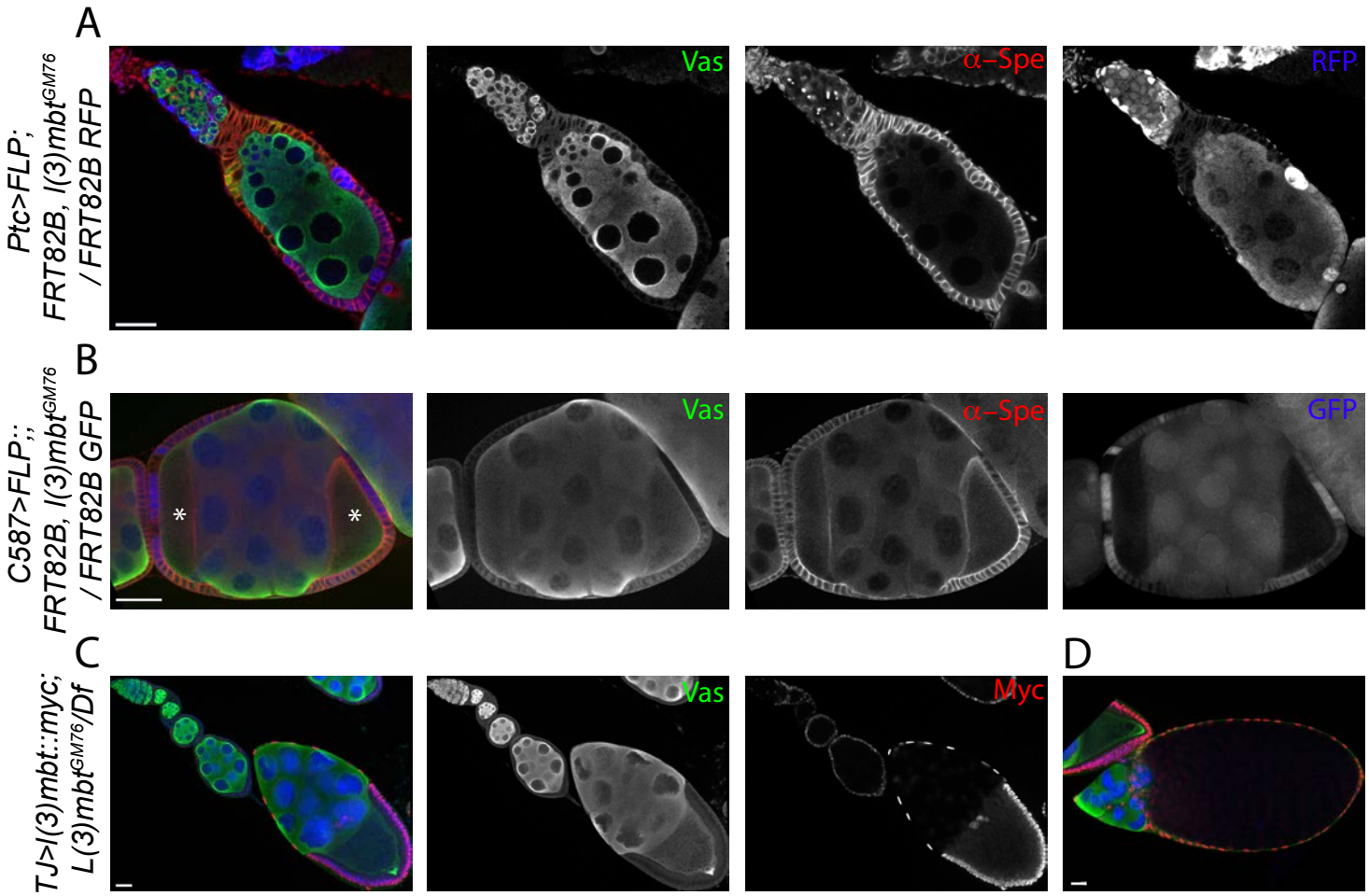
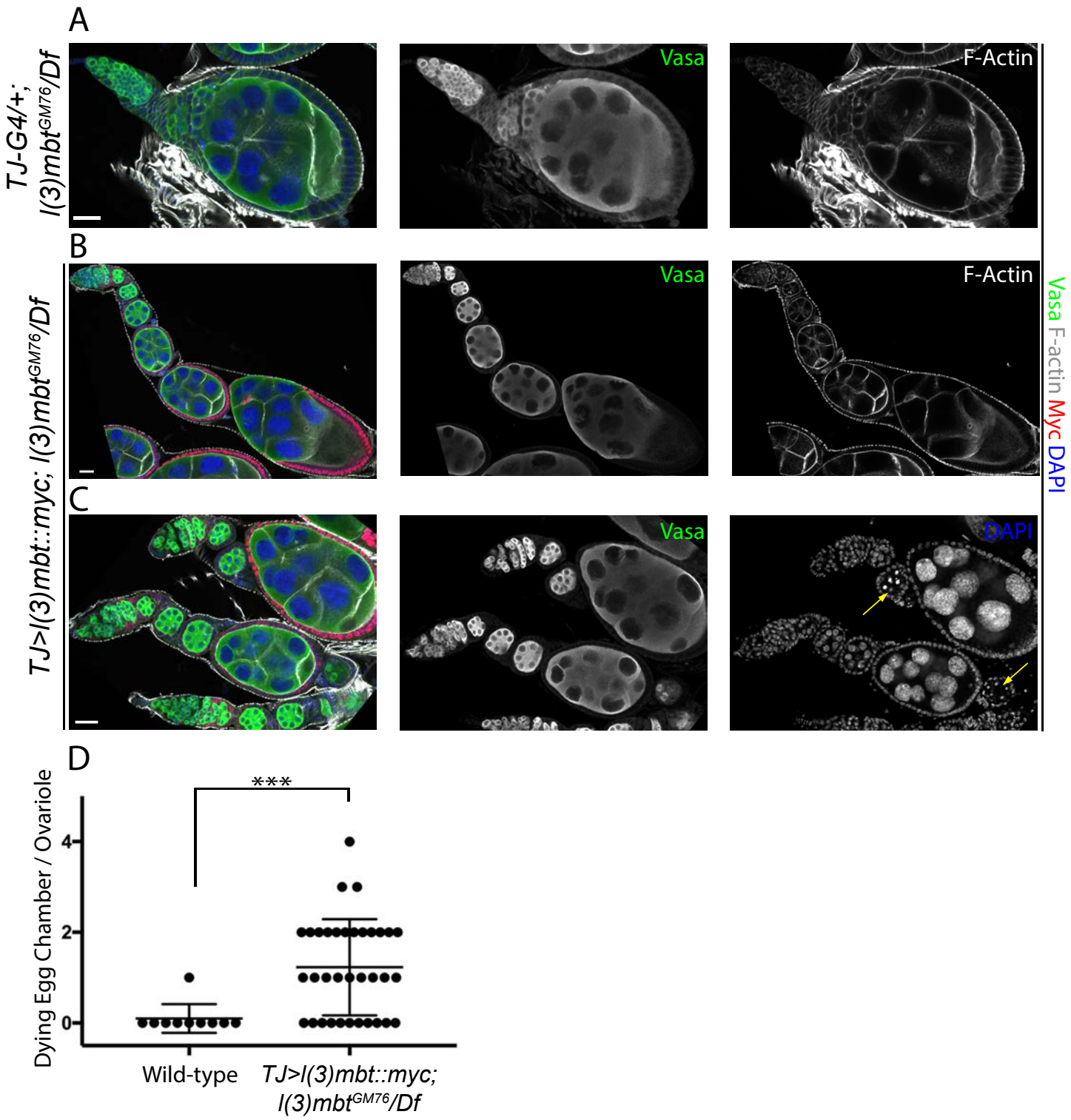


Figure 2. L(3)mbt functions in somatic cells for ovary development. *(A-B)* Confocal images of representative ovarioles with mutant *l(3)mbt* follicle cell clones marked by absence of RFP (A) or GFP (B) (blue), Vasa (green) and α-Spectrin (red). Egg chambers surrounded by numerous *l(3)mbt* mutant follicle cells exhibits aberrant phenotypes. Oocytes are marked by asterisks in (*B*). (*C-D)* Confocal images of mutant ovaries expressing L(3)mbt::myc in somatic cells stained for Vasa (green), Myc (red), and DAPI (blue). *(C)* Ovarian morphology and germ cell number are fully rescued by expression of the transgene in escort and follicle cells. (*D)* Rescued late stage oocyte. Scale bars, 25 μm.

Supplemental figure S2



Supplemental figure S2. L(3)mbt somatic expression rescues *l(3)mbt* mutant ovarian morphology. (*A-C*) Confocal images of ovarioles from (*A*) *l(3)mbt* mutant and (*B-C*) *l(3)mbt* mutant expressing the *l(3)mbt::myc* transgene in somatic cells, stained for Vasa (green), Myc (red), F-Actin (grey), and DAPI (blue). *(B-C)* Complemented ovarioles show wild-type morphology, including proper oocyte specification and ring canals numbers. *(C)* Some *l(3)mbt* mutant egg chambers surrounded by somatic cells expressing the L(3)mbt::myc fusion undergo cell death (yellow arrows). *(D)* Quantification of egg chambers undergoing cell death. Scale bars, 25 μm.

Figure 3

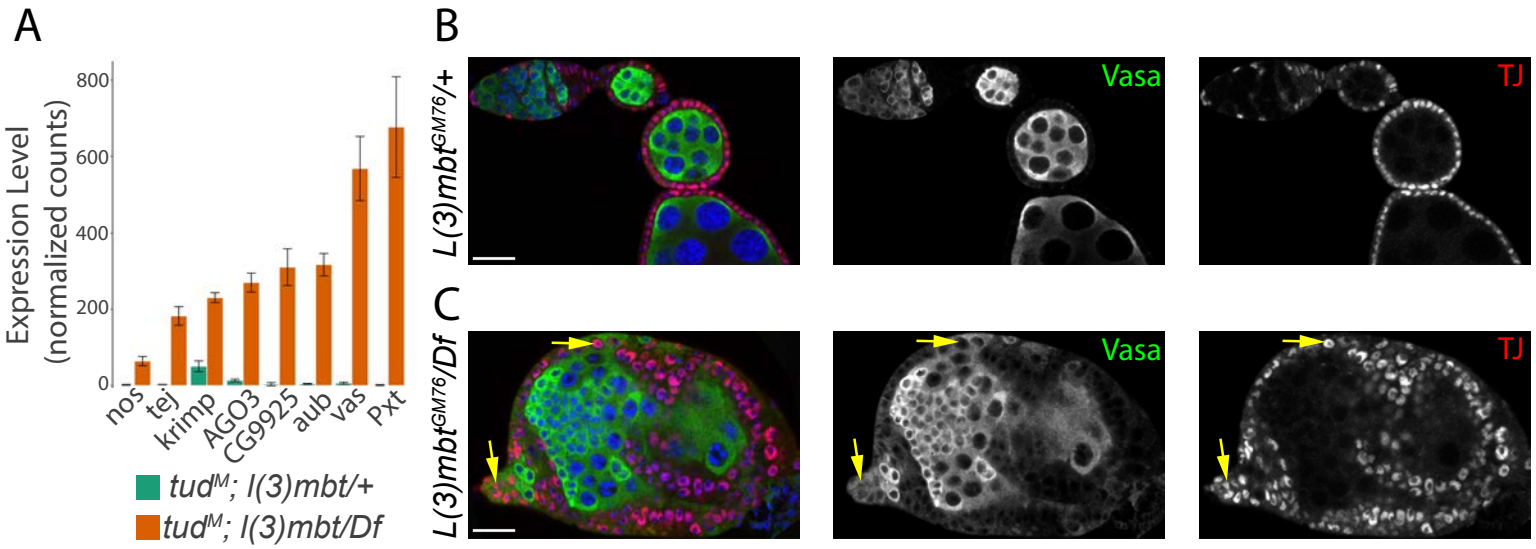
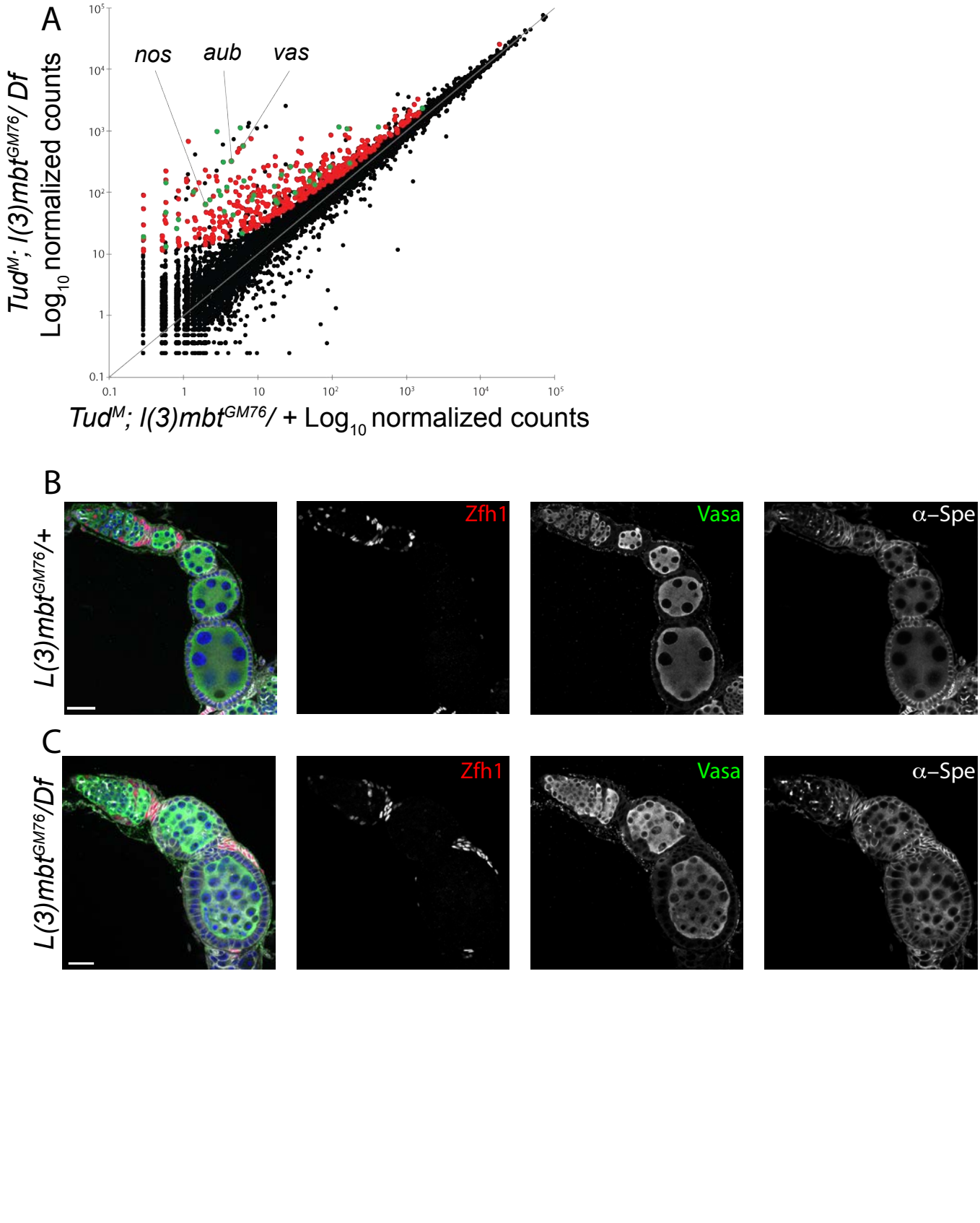


Figure 3. *l(3)mbt* mutant somatic cells are properly specified but ectopically express germline genes. *(A)* Expression level of the germline-specific genes *nos, tej, krimp, AGO3, CG9925, aub, vas* and *Pxt* in tud^{M} ovaries heterozygous and homozygous mutant for *l(3)mbt*, as measured by RNA-seq analysis (expressed in normalized counts). *(B-C)* Confocal images of control and *l(3)mbt* mutant ovarioles stained for Vasa (green), Traffic-Jam (TJ; red) and DAPI (blue). TJ is expressed in all somatic cells of the adult ovary. *(C)* Some TJ-positive somatic cells express the germline marker Vasa (arrows). Scale bars, 25 μm.

Supplemental figure S3



Supplemental figure S3. *L(3)mbt* mutant somatic cells are properly specified but ectopically express germline genes. (*A)* Scatterplot showing the expression of genes in tud^{M}; $l(3)mbt^{GM76}$/+ and tud^{M}; $l(3)mbt^{GM76}$ mutant ovaries, as measured by RNA-seq (normalized counts, $\log_{10}$). De-repressed genes are shown in red and de-repressed MBTS genes in green. (*B-C*) Confocal images of (*B*) control and (*C*) *l(3)mbt* mutant ovarioles stained for Vasa (green), Zfh1 (red), α-Spectrin (grey), and DAPI (blue). (*A*) In control ovaries, Zfh1 is expressed in escort cells, pre-follicle cells as well as stalk cells, which separate egg chambers. *(B) L(3)mbt* mutant ovariole showing normal Zfh1 expression and stalk cells on top of follicle cells. Scale bars, 25 μm.

Figure 4

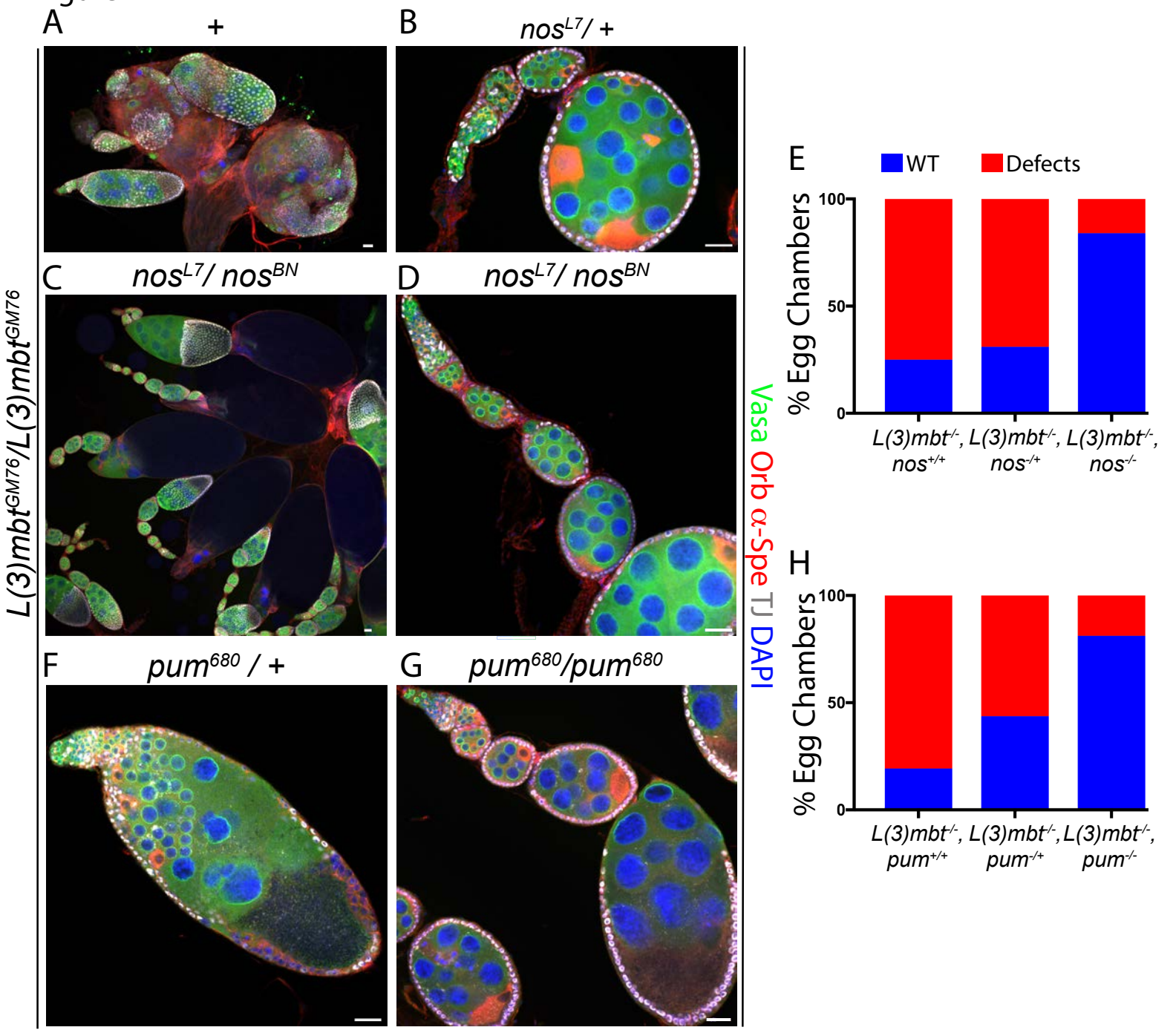
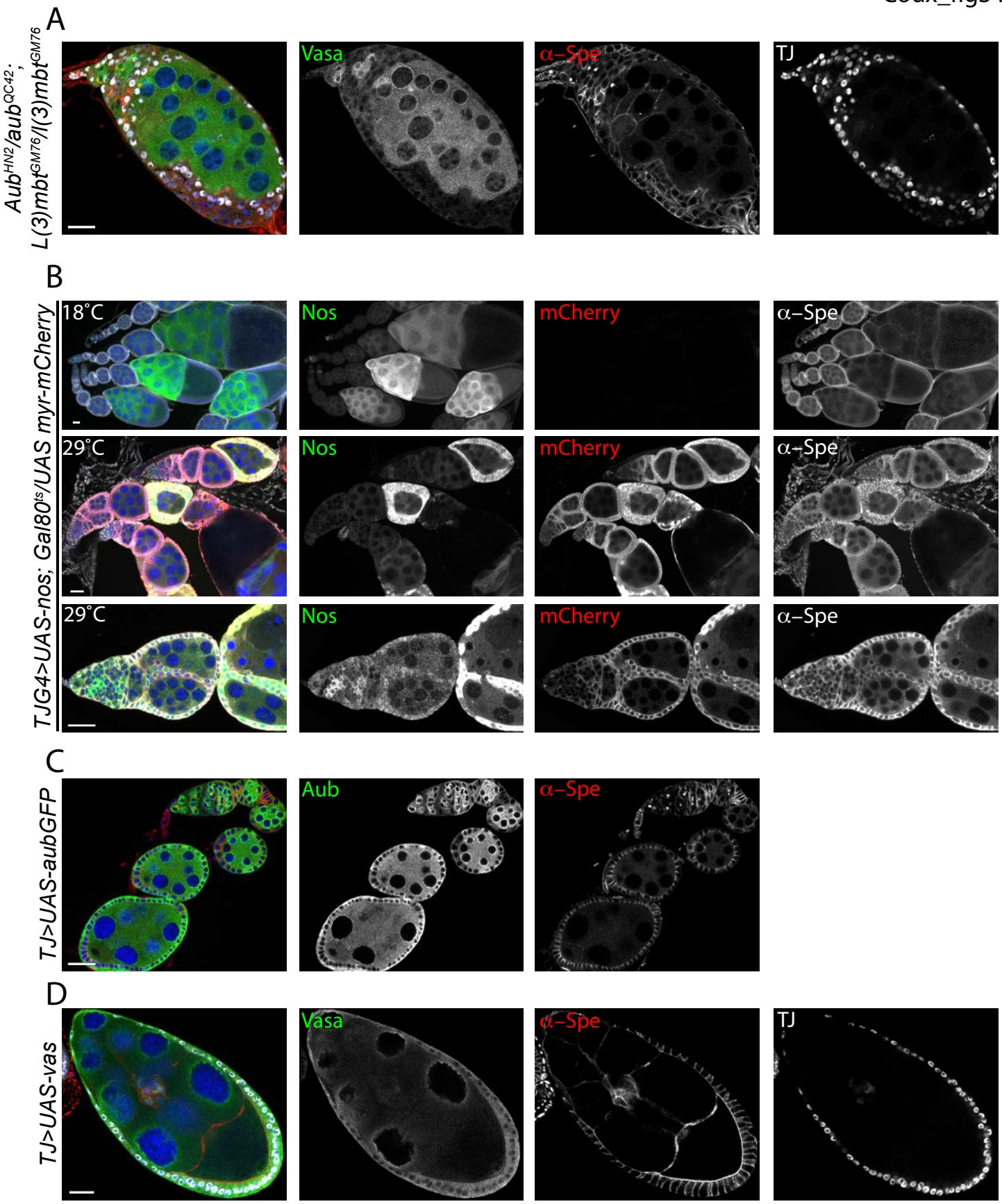


Figure 4. Nanos and its cofactor Pumilio mediate developmental phenotypes in *l(3)mbt* ovaries. (*A-D*, *F-G)* Confocal images of ovaries stained for Vasa (green), Orb and α-Spectrin (red), TJ (grey), and DAPI (blue), scale bars, 25 μm. *(A-D)* Representative confocal images of (A) $l(3)mbt^{GM76}$, (B) $l(3)mbt^{GM76}$, nos^{L7}/ $l(3)mbt^{GM76}$, +, and (C-D) $l(3)mbt^{GM76}$, nos^{L7} / $l(3)mbt^{GM76}$, nos^{BN} double mutant ovarioles. *(A, C)* Low magnification and *(B, D)* medium magnification images. (*E*) Quantification of phenotypes observed in the genotypes described in (*A-D*). *(F-G)* Confocal images of (F) $l(3)mbt^{GM76}$, pum^{680} / $l(3)mbt^{GM76}$, + and (G) $l(3)mbt^{GM76}$, pum^{680} homozygous ovarioles. *(H)* Quantification of phenotypes observed in the genotypes described in (*F-G*).

Coux_figS4



Supplemental figure S4. *Nos* but not *aub* or *vas* ectopic expression interferes with normal ovarian development. *(A)* Confocal image of *aub; l(3)mbt* double mutant ovariole stained for Vasa (green), α-Spectrin (red), TJ (grey), and DAPI (blue). The phenotype is similar to single *l(3)mbt* mutant. *(B)* Confocal images of ovaries expressing *UAS-nos; UAS-myr-mCherry* transgenes in somatic cells using a temperature sensitive system ($Gal80^{ts}$) to express *nos* only in the adult. Ovaries were stained for Nos (green), mCherry (red), α-Spectrin (grey), and DAPI (blue). At 18˚C, the transgenes are not expressed and ovarioles develop normally. When shifted at 29˚C after eclosion, ectopic Nos expression in somatic cells perturbs egg chamber individualization and causes cell death. *(C-D)* Confocal images of ovaries misexpressing Aub or Vasa in somatic cells using the TJ-Gal4 driver and the (*C*) UAS-aubGFP and (*D*) UAS-vas transgenes. Ovaries were stained for (*A*) Aub (green), α-Spectrin (red), and DAPI (blue), or (*D*) Vasa (green), α-Spectrin (red), and TJ (grey). Scale bars, 25 μm.

Coux_Fig5

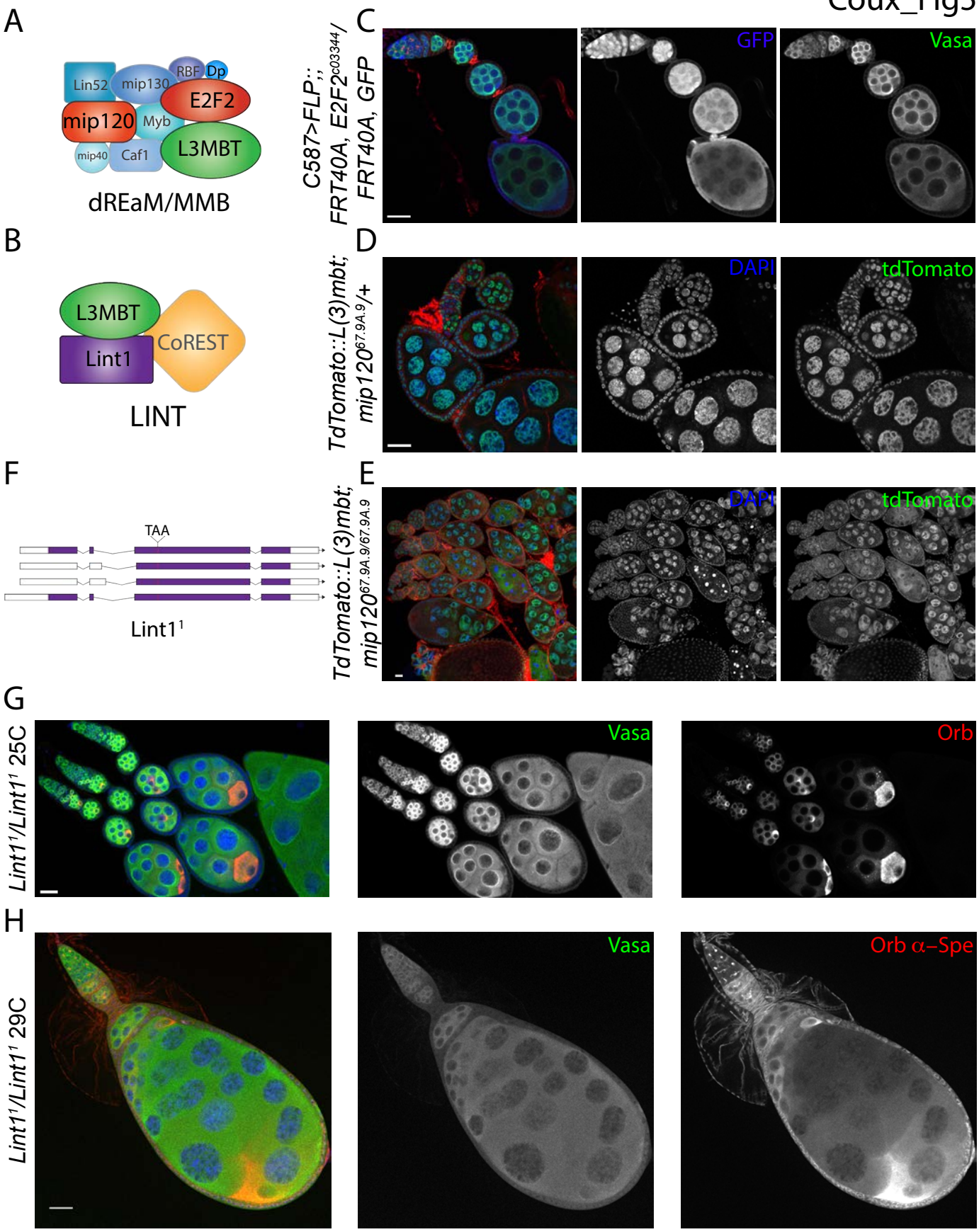
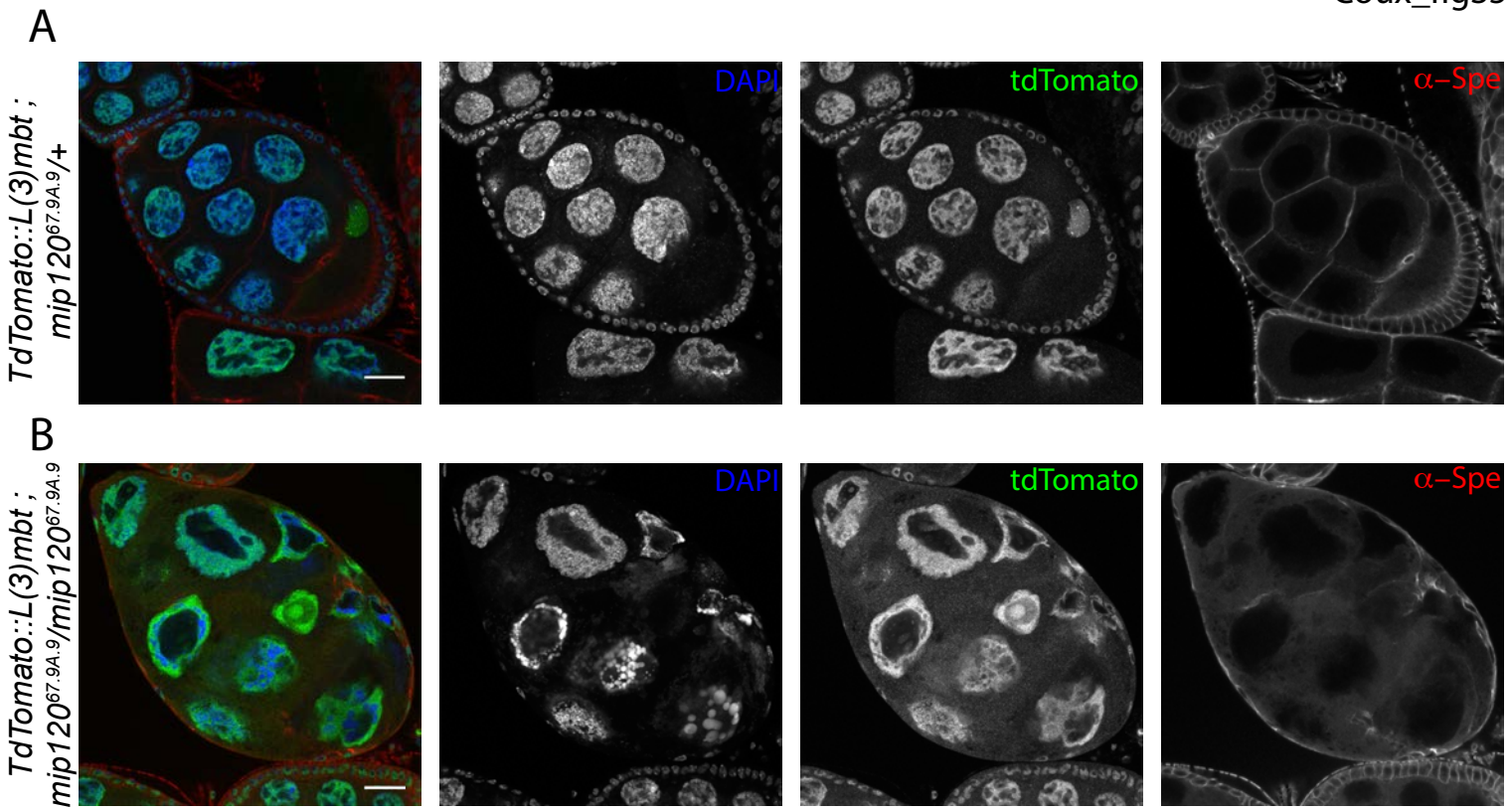


Figure 5. LINT complex mutants have ovarian defects similar to *l(3)mbt*. *(A-B)* Schematic representation of (*A*) the dREAM/MMB and (*B*) LINT complexes. (*C*) Confocal image of ovariole with *E2f2*^{c03344} mutant clones marked by absence of GFP (blue), Vasa (green), α-Spectrin (red). *(D-E)* Confocal images of control and *mip120* mutant ovaries expressing the tdTomato::L(3)mbt fusion and stained for α-Spectrin (red), tdTomato (green), and DAPI (blue). *(D)* *mip120*$^{67.A9.9}$ heterozygous control ovarioles. *(E)* Low magnification confocal image of homozygous mutant *mip120*$^{67.A9.9}$ ovarioles accumulating stage 7-9 egg chambers that eventually undergo apoptosis. *(F)* Schematic representation of the *Lint1*^{1} allele. *(G-H)* Confocal images of *Lint1* mutant ovaries stained for Vasa (green), Orb (and α-Spectrin in *(H)* - red), and DAPI (blue). *(G)* At 25˚C few mutant egg chambers contain two Orb-positive cells or mis-positioned oocyte. *(H)* At 29°C *Lint1*^{1} egg chambers can contain more than 16 germ cells and multiple oocytes similar to defects observed in *l(3)mbt* mutants. Scale bars, 25 μm.

Coux_figS5



Supplemental figure S5. L(3)mbt nuclear localization is not affected in in *mip120*$^{67.9A.9}$ mutant ovaries. *(A-B)* Confocal images of control and *mip120* mutant ovaries expressing the tdTomato::L(3)mbt fusion and stained for α-Spectrin (red), tdTomato (green), and DAPI (blue). TdTomato::L(3)mbt is nuclear and colocalizes with DAPI in both control and mutant ovaries. Mutant nurse cells nuclei are highly vacuolated. Scale bars, 25 μm.

Figure 6

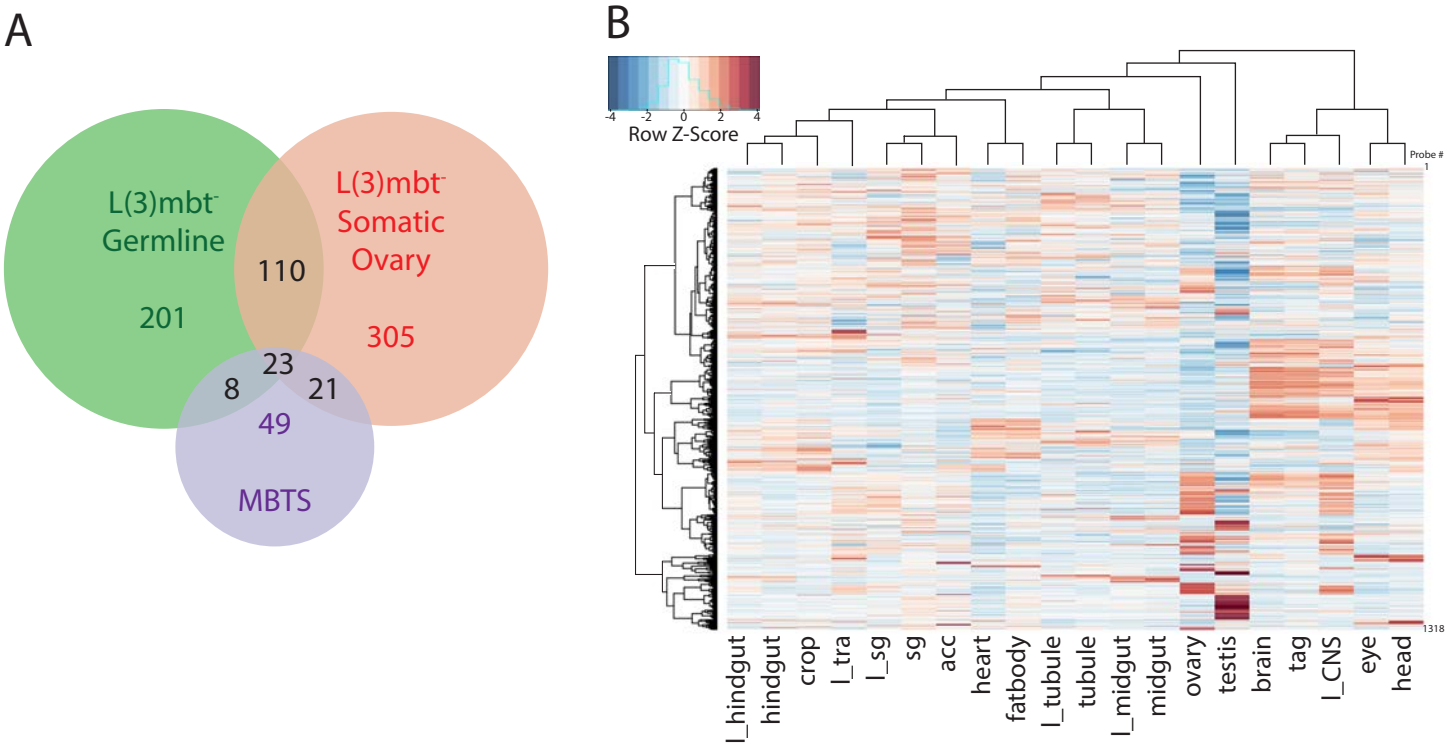
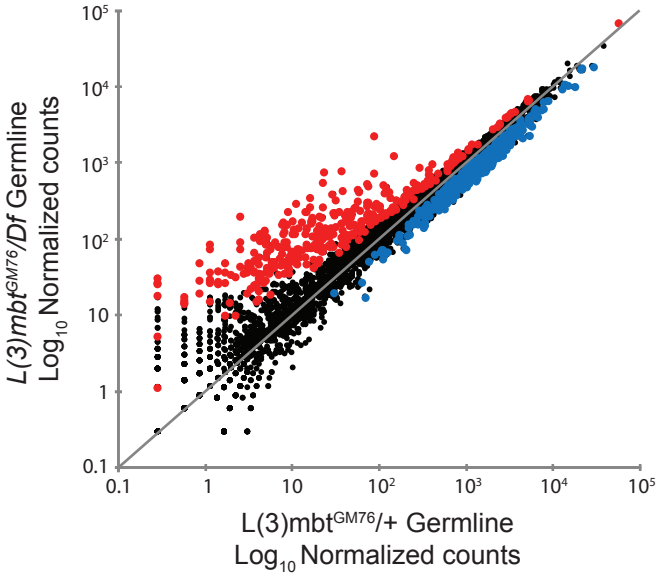


Figure 6. L(3)mbt-mediated gene regulation is tissue specific. *(A)* Venn Diagram showing genes upregulated in *l(3)mbt* mutant ovarian soma (red), female germline (green), and larval brain tumors (MBTS, purple (Janic et al. 2010)). *(B)* Hierarchical clustering of tissue expression profile of genes repressed by L(3)mbt in the female germline. Gene expression per tissue (normalized to fly average) is shown as a Z-Score heatmap.

Supplemental figure S6



Supplemental figure S6. Scatterplots showing expression of genes in control and *l(3)mbt*^{GM76} mutant early embryos, as measured by RNA-seq (normalized counts, $\log_{10}$). Up-regulated genes are shown in red, down-regulated genes are shown in blue.

Figure 7

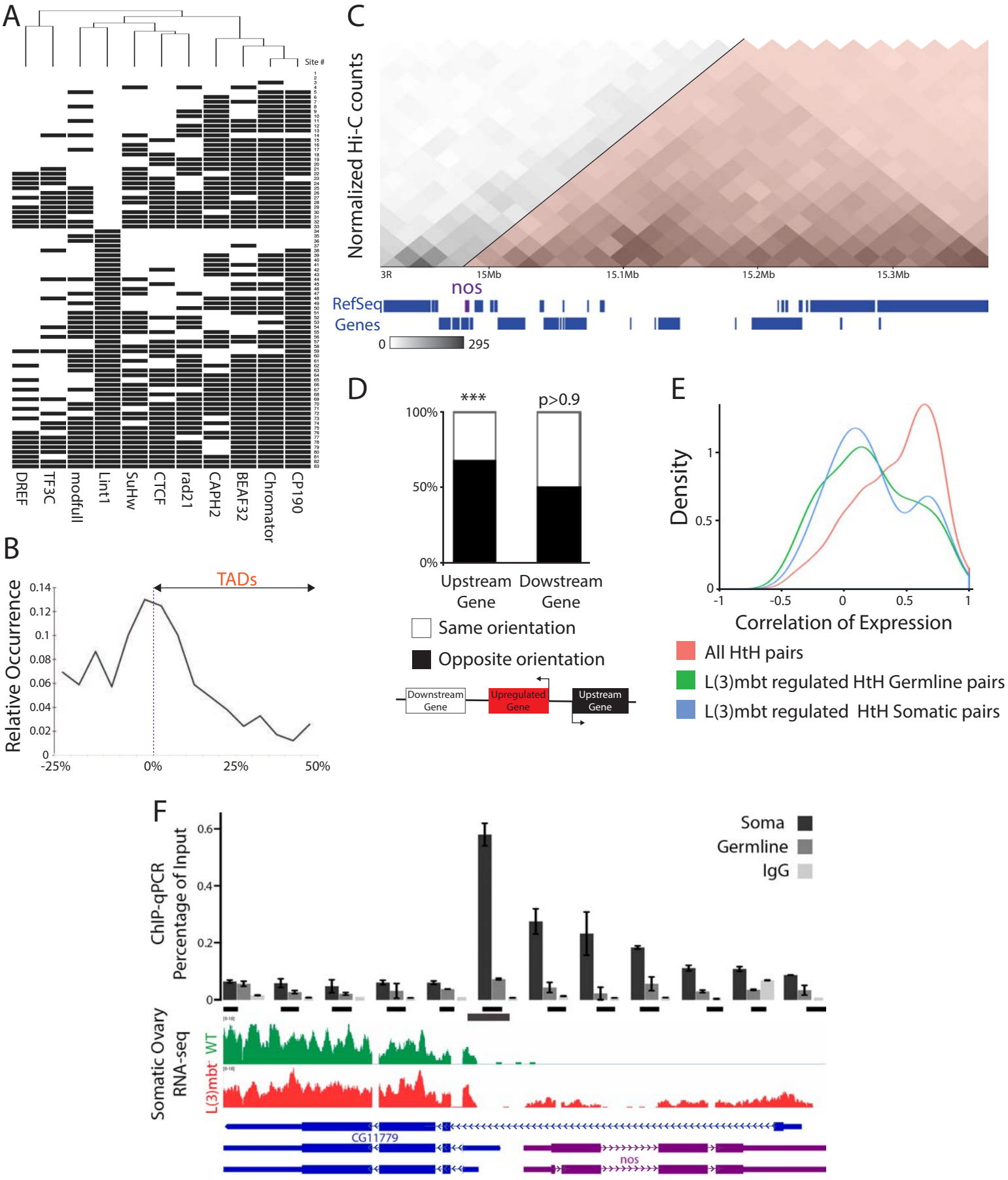
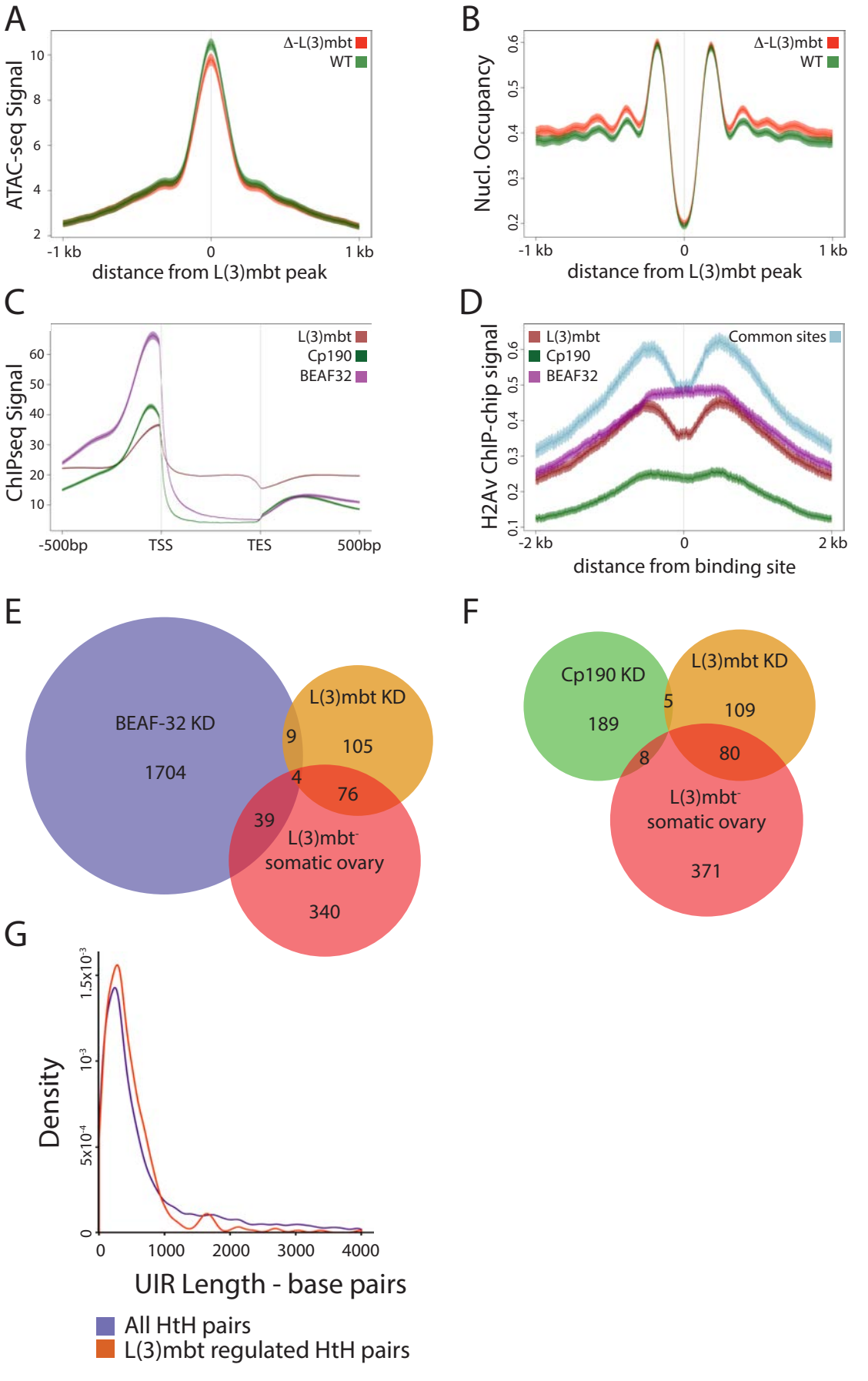


Figure 7. L(3)mbt exerts boundary activity between differentially-regulated, head-to-head genes. *(A)* Hierarchical clustering of architectural proteins binding at L(3)mbt sites required for gene regulation in Kc167 cells, (Van Bortle et al. 2014; Meier et al. 2012). *(B)* Position of L(3)mbt repressed genes with respect to TADs in OSC cells (Ulianov et al. 2016). *(C) Nos* is located at the boundary of a TAD (shown in orange). Hi-C contact probability map (20kb resolution) and TAD position in the vicinity of *nos* locus in OSC cells (Ulianov et al. 2016). *(D)* Orientation of genes flanking L(3)mbt-repressed genes in the somatic ovary. *(E)* Distribution of the correlation for the two genes of HtH pairs (UIRs<500bp). All HtH pairs in the Drosophila genome (red), as well as those pairs containing genes regulated by L(3)mbt in the female germline (green) or somatic ovarian cells (blue). *(F)* L(3)mbt binds between *nos* and the upstream HtH gene *CG11779* in somatic ovarian cells but not in the germline. L(3)mbt ChIP-qPCR results are shown in % of input. Grey rectangle corresponds to the L(3)mbt binding site identified in larval brains. Lower panel: *CG11779* and *nos* expression levels in wild-type and *l(3)mbt* mutant somatic ovarian cells as measured by RNA-seq analysis (normalized counts).

Coux_figS7



Supplemental figure S7. Analysis of L(3)mbt insulator function. *(A)* ATAC-seq signal at L(3)mbt binding sites in control and Δ-*L(3)mbt* OSC cells. L(3)mbt bound regions coincide with peaks of open chromatin. *(B)* Nucleosome occupancy at L(3)mbt sites in control and Δ-*L(3)mbt* OSC cells as calculated based on ATAC-seq data. Loss of *l(3)mbt* does not affect nucleosome positioning at the probed regions. *(C)* Metagene profiles of L(3)mbt, Cp190, and BEAF32 ChIP-seq signal in Kc cells from (Li et al. 2015). TSS: Transcriptional Start Site; TES: Transcriptional End Site. *(D)* H2Av ChIP-chip average profiles in S2 cells at L(3)mbt, Cp190, and BEAF32 bound regions (ModENCODE) and (Richter et al. 2011). L(3)mbt binding sites are flanked by H2Av peaks. Common sites regroup regions bound by the three proteins. The solid line corresponds to the average signal, standard error is depicted by the dark colored area, and light color represents 95% CI. *(E-F)* Venn Diagrams showing genes up-regulated in *l(3)mbt* mutant somatic ovarian cells (red), *l(3)mbt* knocked-down Kc cells (orange), and *Beaf-32* (*E),* purple (Lhoumaud et al. 2014) or *Cp190 (F),* green (Ali et al. 2016) knock-downs. *(G)* UIR Length distribution of all as well as L(3)mbt-regulated head-to-head gene pairs.

Materials and methods

Crosses schemes and genotypes are detailed in Supplemental File S1

Fly stocks

FRT82B, *l(3)mbt*^{GM76}, *e/TM6b* was generated in the Lehmann lab (Yohn et al. 2003) and secondary mutations removed (Richter et al. 2011). The following stocks were obtained from the Kyoto Stock Center: *w*^{1118};; *Df(3R)ED19066/TM6c* (#150208) and y* *w***;P{GawB}NP1624/CyO, P{UAS-lacZ.UW14}UW14* (*Tj-Gal4,* #104055); from Bloomington Drosophila Stock Center: *w1118; P{neoFRT}82B P{Ubi-GFP(S65T)nls}3R/TM6B, Tb*^{1} (BDSC #32655), *y*^{1} *w***; P{UAS-FLP.D}JD1 (BDSC #4539), P{otu-GAL4::VP16.R}1, w***; P{GAL4-nos.NGT}40; P{GAL4::VP16-nos.UTR}CG6325*^{MVD1} (Maternal Triple Driver, BDSC #31777), and *w*; P{tubP-GAL80*^{ts}*}2/TM2* (BDSC #7017). The following transgenes *C587-Gal4*, *UAS-nos-tub* (Clark et al. 2002; Ye et al. 2004), *UAS-vas* (Sengoku et al. 2006), *UASp-mCherry-myr, UAS-Aub-GFP* (Harris and Macdonald 2001), and *TdTomato::l(3)mbt* (Blanchard et al. 2014) were obtained from the Xie, Jan, Nakamura, Zallen, Macdonald, and Botchan labs, respectively. The *mip120*$^{67.9A.9}$ (Beall et al. 2007) and *FRT40A e2f2*^{c03344} (Ambrus et al. 2007) mutations were generated by the Botchan and Frolov labs. The following mutations are from the Lehmann lab stocks: *tud*^{1}, *tud*^{B42} (Arkov et al. 2006), *aub*^{HN2}, *aub*^{QC42} (Schupbach and Wieschaus 1991), *nos*^{L7} (Wang and Lehmann 1991), and *nos*^{BN} (Wang et al. 1994). All stocks were maintained at 18˚C and crosses were performed at 25˚C unless otherwise stated.

Generation of transgenic lines

To generate *UASp-l(3)mbt::*myc transgenic flies, *l(3)mbt* coding sequence was amplified from LD05287 gold cDNA (Drosophila Gemonics Resource Center), cloned using the p-ENTR/D-TOPO system and recombined into the pPWM destination vector (Drosophila Gateway Vector Collection) using Gateway technology (Invitrogen). *pPWM-l(3)mbt* was randomly inserted on the 2nd chromosome through P-transposition. The *UAS-tdTomato* short hairpin RNA (*tdTomato*^{shRNA}) was generated by Transgenic RNAi Project (SH08523.N) and inserted in the attP2 site (BDSC #8622). To generate the *lint1*^{1} mutation, the target sequence (chrX:11044844-11044866) was identified by the flyCRISPR Optimal Target Finder tool (Gratz et al. 2014), amplified from genomic DNA using *lint-CRISPR* oligos (Supplemental Table S4) and ligated in the pU6-BbsI-gRNA plasmid (Gratz et al. 2013). The resulting construct was injected in *FRT19A;; vas-Cas9* embryos and progeny were screened by PCR and sequencing.

Immunofluorescence

Adult ovaries from 2-3 day-old fatten females were dissected and stained following standard procedures (Arkov et al. 2006). Ovaries were mounted in SlowFade® Gold mountant (Invitrogen) and imaged on Zeiss LSM780 or LSM800 confocal microscopes, using 10, 20 or 43x objectives. The following primary antibodies were used: rabbit α-Vasa (1:5000, Lehmann lab); goat α-Vasa (1:200, Santa Cruz Biotechnology sc-26877); mouse α-spectrin (1:200, DSHB); mouse α-Orb (4H8, 1:200, DHSB); chicken α-GFP (1:1000, Aves GFP-1020); goat α-tj (1:7000, kind gift of Dorothea Godt (Li et al. 2003)); rabbit α-zfh1 (1:5000, Lehmann lab); rat α-RFP (1:500, Chromotek 5F8); rabbit α-Nanos (1:200, kind gift of Prof. Nakamura); rabbit α-Aub (1:1000, Lehmann lab); rabbit α-DsRed (1:500, Living Colors #632496). Alexa Fluor 647 Phalloidin (1:500 Life Technologies) and rabbit α-myc Alexa fluor 555 conjugated (Millipore 16-225) were

used as secondary antibodies. Alexa Fluor 488- (Life Technologies), Cy3- or Alexa Fluor 647-conjugated (Jackson Immunoresearch) secondary antibodies were used at a 1:1000 dilution.

RNA sequencing

60-70 ovaries from females of maternal *tud*^{1}/*tud*^{B42} and zygotic *l(3)mbt*^{GM76}/+ or *l(3)mbt*^{GM76}/*Df* genotypes (see Supplemental File 1) were dissected in cold PBS and RNA was extracted using TRIzol (Invitrogen) following the manufacturer's protocol. For early embryos, *TJ>UAS-l(3)mbt::myc;l(3)mbt*^{GM76}/+ and *TJ>UAS-l(3)mbt::myc;l(3)mbt*^{GM76}/*Df* females were allowed to lay for 30 minutes to 1 hour on agar plates. Embryos were dechorionated in 50% bleach for 5 minutes, rinsed with PBS, and then lysed in TRIzol. Libraries were generated from 1μg of total RNA using the NEBNext Poly(A) magnetic Isolation Module (NEB #7490) and the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB # E7420). Libraries from biological replicates (three for ovaries, two for embryos) were sequenced on an Illumina Hi-Seq2000, single-end 50 run.

RNA-seq data analysis

Sequencing results were demultiplexed and converted to FASTQ format using Illumina Bcl2FastQ software (Illumina). Reads were aligned to the fly genome (build dm3/BDGP5) using the splice-aware STAR aligner (Dobin et al. 2013). PCR duplicates were removed using the Picard toolkit (https://broadinstitute.github.io/picard/). HTSeq package was used to generate counts for each gene based on how many aligned reads overlap its exons (Anders et al. 2015). These counts were then used to test for differential expression using negative binomial generalized linear models implemented by the DESeq2 R package.

Tissue expression clustering

Expression of deregulated genes was extracted from FlyAtlas (Chintapalli et al. 2007) using FlyBaseIDs, normalized to fly average and log2 transformed. Distance matrix was calculated using the "Manhattan" method and data clustered using "ward.D2". Heatmap was generated using the heatmap.2 function of the gplots R package.

Tissue-specific Chromatin Immuno-Precipitation followed by quantitative PCR (ChIP-qPCR)

The *tdTomato::l(3)mbt* transgene (in a heterozygous *l(3)mbt* background) was knocked-down in the soma using *TJ>UAS-tdTomato*^{shRNA} for germline-specific L(3)mbt ChIP. Reciprocally, we used *MTD>UAS-tdTomato*^{shRNA} to perform L(3)mbt somatic ChIP. Chromatin Immuno-Precipitation was carried as previously described (Soshnev et al. 2008) with modifications: 60-70 1-to-2 day-old *MTD>UAS-tdTomato*^{shRNA} or 45-50 *TJ>UAS-tdTomato*^{shRNA} females were fattened overnight, dissected in cold PBS and their ovaries were fixed in 1.8% formaldehyde in PBS for 20 minutes at room temperature. Chromatin was sheared at 4˚C using 16 cycles of sonication (30s ON – 30s OFF) in a Bioruptor Pico (Diagenode). 150µg of chromatin was used per IP reaction with 15μg of goat anti-tdTomato polyclonal antibody (SICGEN AB8181-200). QPCRs were performed on a Roche LightCycler 480 machine using standard procedures, primers described in Supplemental Table S4. Binding was calculated as a percentage of Input: %input = 100*2^(Ct(input) – Ct(IP))* 1/Dilution Factor).

Expression correlation of divergently paired genes

Pearson correlation was calculated for the expression of both genes of head-to-head pairs (UIRs<500 bp) across multiple cell lines and developmental stages (modENCODE dataset #3305, http://www.modencode.org/) using R correlation function and "pairwise.complete.obs" method. Correlations of expression were plotted as probability density for all, L(3)mbt-regulated somatic, and germline HtH gene pairs using the ggplot2 R package.

Cell culture

Ovarian Somatic Cell (OSCs), kind gifts of the Siomi and Brennecke labs, were cultured using standard procedures (Sumiyoshi et al. 2016).

ATAC-seq

ATAC-seq libraries were generated from $1x10^6$ WT and *Δ–l(3)mbt* OSC cells using the Greenleaf lab protocol (Buenrostro et al. 2013) modified as in (Iwasaki et al. 2016). Experiment was performed in triplicates, and libraries were sequenced on Illumina Hiseq 2000 using Paired-End 50 runs.

ATAC-seq data analysis

Sequencing results were demultiplexed and converted to FASTQ format using Illumina Bcl2FastQ software. Reads were aligned to the Drosophila melanogaster genome (build dm3/BDGP5) with Bowtie2 (Langmead and Salzberg 2012) using local alignment and allowing maximum fragment length up to 2 kb. Mitochondrial and poorly mapped reads (mapping quality <30) were discarded. Duplicate reads were removed using Sambamba (Tarasov et al. 2015). Broad open chromatin regions were detected with MACS2 (Zhang et al. 2008). Triplicate libraries yielding around $42x10^7$

demultiplexed reads were pooled, and nucleosome positioning was generated using NucleoATAC (Schep et al. 2015) based on the broad peaks extended by 100 bp.

Data processing and plotting

Metaplots were generated using SeqPlots (https://github.com/Przemol/seqplots). L(3)mbt, CP190 and BEAF32 ChIP-seq profiles used to generate metagene plots originate from ((Li et al. 2015), GEO GSE63518). H2Av (modENCODE dataset #2991) ChIP-chip data in S2 cells was used to plot profiles across L(3)mbt (larval brain ChIPseq, (Richter et al. 2011), GEO GSE29206), and Cp190 and BEAF32 sites (0-12h embryos, ChIP-chip, modENCODE datasets #21 and #22). The coordinates of the Upstream Intergenic Regions (+/-1kb to include UTRs) of genes upregulated in *l(3)mbt* depleted Kc167 cells ((Meier et al. 2012), ArrayExpress MTAB-849) were intersected with the Architectural Protein Binding Sites matrix generated by ChIP-seq in Kc167 cells (Additional File 3 from (Van Bortle et al. 2014)) using Bedtools (Quinlan 2014). Hi-C and TADs predictions in OSC cells were generated by ((Ulianov et al. 2016), GEO GSE69013) and plotted using the WashU EpiGenome Browser (Zhou et al. 2011). Occurrence of derepressed genes with respect to TADs was calculated using a custom R script. Briefly, TADs were divided in 20 windows and occurrence of derepressed genes was calculated for each bin as well as 6 bins up and downstream.

Acknowledgements

We thank the entire Lehmann lab for discussion and input; C. Desplan, D. Keefe, E. Mazzoni, N. Dyson and J. Treisman for discussions and advice; A. Soshnev for advice on ChIP, P. Rocha, R. Raviram and I. Dogalev for bioinformatics input. We would like to thank J. Knoblich and C.

Richter, M. Siomi, J. Brennecke, N. Dyson, M. Botchan and D. Blanchard, M.V. Frolov, D. Godt, S. Kobayashi and A. Nakamura for sharing reagents. We are grateful to the Bloomington Drosophila Stock Center (NIH P40OD018537) for fly strains; the Developmental Studies Hybridoma Bank, created by the NICHD of the NIH and maintained at The University of Iowa for antibodies and the NYUMC Genome Technology Center (NIH P30CA016087), for sequencing. R.X.C was supported by the New York State Stem Cell Program (NYSTEM) of the New York State Health Department (#C026880) and FKT was funded by EMBO and HFSP fellowships. RL is supported by NIH R01/R37HD41900 and is an HHMI investigator.

Author Contributions

R.X.C. and R.L. conceived the project. R.X.C. F.K.T. and R.L. designed experiments and wrote the manuscript. R.X.C. performed the experiments and analyzed the data.

References

Ali T, Krüger M, Bhuju S, Jarek M, Bartkuhn M, Renkawitz R. 2016. Chromatin binding of Gcn5 in Drosophila is largely mediated by CP190. *Nucleic Acids Research* **45**: 2384–2395.

Ambrus AM, Nicolay BN, Rasheva VI, Suckling RJ, Frolov MV. 2007. dE2F2-independent rescue of proliferation in cells lacking an activator dE2F1. *Mol Cell Biol* **27**: 8561–8570.

Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169.

Arkov AL, Wang J-YS, Ramos A, Lehmann R. 2006. The role of Tudor domains in germline development and polar granule architecture. *Development* **133**: 4053–4062.

Beall EL, Lewis PW, Bell M, Rocha M, Jones DL, Botchan MR. 2007. Discovery of tMAC: a Drosophila testis-specific meiotic arrest complex paralogous to Myb-Muv B. *Genes & Development* **21**: 904–919.

Blanchard DP, Georlette 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 USA*.

Bohla D, Herold M, Panzer I, Buxa MK, Ali T, Demmers J, Krüger M, Scharfe M, Jarek M, Bartkuhn M, et al. 2014. A Functional Insulator Screen Identifies NURF and dREAM Components to Be Required for Enhancer-Blocking. *PLoS ONE* **9**: e107765.

Bonasio R, Lecona E, Reinberg D. 2010. MBT domain proteins in development and disease. *Semin Cell Dev Biol* **21**: 221–230.

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**: 1213–1218.

Chintapalli VR, Wang J, Dow JAT. 2007. Using FlyAtlas to identify better Drosophila melanogaster models of human disease. *Nat Genet* **39**: 715–720.

Christerson LB, McKearin DM. 1994. orb is required for anteroposterior and dorsoventral patterning during Drosophila oogenesis. *Genes & Development* **8**: 614–628.

Clark IE, Dobi KC, Duchow HK, Vlasak AN, Gavis ER. 2002. A common translational control mechanism functions in axial patterning and neuroendocrine signaling in Drosophila. *Development* **129**: 3325–3334.

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.

Edgar BA, Schubiger G. 1986. Parameters controlling transcriptional activation during early Drosophila development. *Cell* **44**: 871–877.

Gateff E, Löffler T, Wismar J. 1993. A temperature-sensitive brain tumor suppressor mutation of Drosophila melanogaster: developmental studies and molecular localization of the gene. *Mech Dev* **41**: 15–31.

Georlette D, Ahn S, MacAlpine DM, Cheung E, Lewis PW, Beall EL, Bell SP, Speed T, Manak JR, Botchan MR. 2007. Genomic profiling and expression studies reveal both positive and negative activities for the Drosophila Myb MuvB/dREAM complex in proliferating cells. **21**: 2880–2896.

Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, Harrison MM, Wildonger J, O'Connor-Giles KM. 2013. Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **194**: 1029–1035.

Gratz SJ, Ukken FP, Rubinstein CD, Thiede G, Donohue LK, Cummings AM, O'Connor-Giles KM. 2014. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. *Genetics* **196**: 961–971.

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.

Harris AN, Macdonald PM. 2001. Aubergine encodes a Drosophila polar granule component

required for pole cell formation and related to eIF2C. *Development* **128**: 2823–2832.

Harrison DA, Perrimon N. 1993. Simple and efficient generation of marked clones in Drosophila. *Current Biology* **3**: 424–433.

Hinz U, Giebel B, Campos-Ortega JA. 1994. The basic-helix-loop-helix domain of Drosophila lethal of scute protein is sufficient for proneural function and activates neurogenic genes. *Cell* **76**: 77–87.

Hou C, Li L, Qin ZS, Corces VG. 2012. Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. *Mol Cell* **48**: 471–484.

Hug CB, Grimaldi AG, Kruse K, Vaquerizas JM. 2017. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell* **169**: 216–228.e19.

Iwasaki YW, Murano K, Ishizu H, Shibuya A, Iyoda Y, Siomi MC, Siomi H, Saito K. 2016. Piwi Modulates Chromatin Accessibility by Regulating Multiple Factors Including Histone H1 to Repress Transposons. *Mol Cell* **63**: 408–419.

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.

Jarriault S, Schwab Y, Greenwald I. 2008. A Caenorhabditis elegans model for epithelial-neuronal transdifferentiation. *Proc Natl Acad Sci USA* **105**: 3790–3795.

Korenjak M, Kwon E, Morris RT, Anderssen E, Amzallag A, Ramaswamy S, Dyson NJ. 2014. dREAM co-operates with insulator-binding proteins and regulates expression at divergently paired genes. *Nucleic Acids Research* **42**: 8939–8953.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Publishing Group* **9**: 357–359.

Leatherman JL, Dinardo S. 2008. Zfh-1 controls somatic stem cell self-renewal in the Drosophila testis and nonautonomously influences germline stem cell self-renewal. *Cell Stem Cell* **3**: 44–54.

Lewis PW, Beall EL, Fleischer TC, Georlette D, Link AJ, Botchan MR. 2004. Identification of a Drosophila Myb-E2F2/RBF transcriptional repressor complex. *Genes & Development* **18**: 2929–2940.

Lhoumaud P, Hennion M, Gamot A, Cuddapah S, Queille S, Liang J, Micas G, Morillon P, Urbach S, Bouchez O, et al. 2014. Insulators recruit histone methyltransferase dMes4 to regulate chromatin of flanking genes. **33**: 1599–1613.

Li L, Lyu X, Hou C, Takenaka N, Nguyen HQ, Ong C-T, Cubeñas-Potts C, Hu M, Lei EP, Bosco G, et al. 2015. Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol Cell* **58**: 216–231.

Li MA, Alls JD, Avancini RM, Koo K, Godt D. 2003. The large Maf factor Traffic Jam controls gonad morphogenesis in Drosophila. *Nat Cell Biol* **5**: 994–1000.

McGuire SE, Mao Z, Davis RL. 2004. Spatiotemporal gene expression targeting with the TARGET and gene-switch systems in Drosophila. *Sci STKE* **2004**: pl6–pl6.

Meier K, Mathieu E-L, Finkernagel F, Reuter LM, Scharfe M, Doehlemann G, Jarek M, Brehm A. 2012. LINT, a novel dL(3)mbt-containing complex, represses malignant brain tumour signature genes. ed. A. Akhtar. *PLoS Genet* **8**: e1002676.

Miles WO, Korenjak M, Griffiths LM, Dyer MA, Provero P, Dyson NJ. 2014. Post-transcriptional gene expression control by NANOS is up-regulated and functionally important in pRb-deficient cells. **33**: 2201–2215.

Nègre N, Brown CD, Shah PK, Kheradpour P, Morrison CA, Henikoff JG, Feng X, Ahmad K, Russell S, White RAH, et al. 2010. A comprehensive map of insulator elements for the Drosophila genome. *PLoS Genet* **6**: e1000814.

Petrella LN, Wang W, Spike CA, Rechtsteiner A, Reinke V, Strome S. 2011. synMuv B proteins antagonize germline fate in the intestine and ensure C. elegans survival. *Development* **138**: 1069–1079.

Quinlan AR. 2014. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr Protoc Bioinformatics* **47**: 11.12.1–34.

Raisner RM, Hartley PD, Meneghini MD, Bao MZ, Liu CL, Schreiber SL, Rando OJ, Madhani HD. 2005. Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* **123**: 233–248.

Richter C, Oktaba K, Steinmann J, Knoblich JA. 2011. The tumour suppressor L(3)mbt inhibits neuroepithelial proliferation and acts on insulator elements. *Nat Cell Biol* **13**: 1029–1039.

Schep AN, Buenrostro JD, Denny SK, Schwartz K, Sherlock G, Greenleaf WJ. 2015. Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. *Genome Res* **25**: 1757–1770.

Schupbach T, Wieschaus E. 1991. Female sterile mutations on the second chromosome of Drosophila melanogaster. II. Mutations blocking oogenesis or altering egg morphology. *Genetics* **129**: 1119–1136.

Sengoku T, Nureki O, Nakamura A, Kobayashi S, Yokoyama S. 2006. Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. *Cell* **125**: 287–300.

Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. 2012. Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* **148**: 458–472.

Smendziuk CM, Messenberg A, Vogl AW, Tanentzapf G. 2015. Bi-directional gap junction-

mediated soma-germline communication is essential for spermatogenesis. *Development* **142**: 2598–2609.

Soboleva TA, Nekrasov M, Ryan DP, Tremethick DJ. 2014. Histone variants at the transcription start-site. *Trends Genet* **30**: 199–209.

Soshnev AA, Li X, Wehling MD, Geyer PK. 2008. Context differences reveal insulator and activator functions of a Su(Hw) binding region. *PLoS Genet* **4**: e1000159.

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 & Development* **30**: 1617–1622.

Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. 2015. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**: 2032–2034.

Tursun B, Patel T, Kratsios P, Hobert O. 2011. Direct Conversion of C. elegans Germ Cells into Specific Neuron Types. *Science* **331**: 304–308.

Ulianov SV, Khrameeva EE, Gavrilov AA, Flyamer IM, Kos P, Mikhaleva EA, Penin AA, Logacheva MD, Imakaev MV, Chertovich A, et al. 2016. Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. *Genome Res* **26**: 70–84.

Van Bortle K, Nichols MH, Li L, Ong C-T, Takenaka N, Qin ZS, Corces VG. 2014. Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol* **15**: R82.

Van Bortle K, Ramos E, Takenaka N, Yang J, Wahi JE, Corces VG. 2012. Drosophila CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. *Genome Res* **22**: 2176–2187.

Wang C, Dickinson LK, Lehmann R. 1994. Genetics of nanos localization in Drosophila. *Dev Dyn* **199**: 103–115.

Wang C, Lehmann R. 1991. Nanos is the localized posterior determinant in Drosophila. *Cell* **66**: 637–647.

Weidmann CA, Qiu C, Arvola RM, Lou T-F, Killingsworth J, Campbell ZT, Tanaka Hall TM, Goldstrohm AC. 2016. Drosophila Nanos acts as a molecular clamp that modulates the RNA-binding and repression activities of Pumilio. *Elife* **5**: e17096.

Yang J, Ramos E, Corces VG. 2012. The BEAF-32 insulator coordinates genome organization and function during the evolution of Drosophila species. *Genome Res* **22**: 2199–2207.

Yang L, Yu J. 2009. A comparative analysis of divergently-paired genes (DPGs) among Drosophila and vertebrate genomes. *BMC Evol Biol* **9**: 55.

Ye B, Petritsch C, Clark IE, Gavis ER, Jan LY, Jan YN. 2004. Nanos and Pumilio are essential for dendrite morphogenesis in Drosophila peripheral neurons. *Curr Biol* **14**: 314–321.

Yohn CB, Pusateri L, Barbosa V, Lehmann R. 2003. l(3)malignant brain tumor and three novel genes are required for Drosophila germ-cell formation. *Genetics* **165**: 1889–1900.

Zhang Y, Liu T, Meyer CA, Eeckhoute 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.

Zhou X, Maricque B, Xie M, Li D, Sundaram V, Martin EA, Koebbe BC, Nielsen C, Hirst M, Farnham P, et al. 2011. The Human Epigenome Browser at Washington University. *Nat Methods* **8**: 989–990.

Zhu C-H, Xie T. 2003. Clonal expansion of ovarian germline stem cells during niche formation in Drosophila. *Development* **130**: 2579–2588.

List of Supplemental files:

Supplemental File S1: fly genotypes and crosses

Supplemental Table S1: *tud*^{M}*, l(3)mbt* (somatic) ovaries RNA seq

Supplemental Table S2: *l(3)mbt* (germline) emrbyos RNA seq

Supplemental Table S3: Insulator and accessory proteins occupancy at L(3)mbt binding sites upstream regulated genes in Kc167 cells.

Supplemental Table S4: list of oligos