Multi-tissue proteomics identifies a link between satellite DNA organization and transgenerational transposon repression in *Drosophila*

Ankita Chavan[1,2,3¶], Lena Skrutl[1,2¶], Federico Uliana[1,3], Melanie Pfister[1], David Burke[4], Franziska Brändle[1,2], Pedro Beltrao[5], Madhav Jagannathan[1,3*]

[1]Institute of Biochemistry, ETH Zürich, Zürich, Switzerland

[2]Life Sciences Zürich Graduate School, Zürich, Switzerland

[3]Bringing Materials to Life Consortium, Zürich, Switzerland

[4]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Cambridge, UK

[5]Institute of Molecular Systems Biology, ETH Zürich, Zürich, Switzerland

*Corresponding author

E-mail: madhav.jagannathan@bc.biol.ethz.ch

[¶]These authors contributed equally to this work.

ABSTRACT

The pericentromeric heterochromatin of eukaryotic chromosomes primarily consists of abundant non-coding repeats known as satellite DNA, which promote accurate chromosome segregation and genome stability. During interphase, sequence-specific satellite DNA-binding proteins cluster repeats from multiple chromosomes into foci known as chromocenters, which function to encapsulate the entire genome in a single nucleus. Despite the pivotal role of satellite DNA-binding proteins and chromocenters in cellular function, the proteins associated with these repetitive sequences remains incompletely characterized. Here, we use quantitative mass spectrometry to characterize the chromocenter-associated proteome in *Drosophila* embryos, ovaries and testes using two satellite DNA-binding proteins, D1 and Prod, as baits. We identify nearly 500 interactions, including known heterochromatin-associated proteins as well as proteins previously unlinked to satellite DNA or chromocenters. Among these interactions, we find that multiple components of the transposon-silencing piRNA pathway are associated with chromocenters. Strikingly, we reveal that proper satellite DNA clustering plays a role in transgenerational transposon repression, such that mothers with disrupted chromocenters give rise to progeny that exhibit transposon de-repression, germ cell loss and gonadal atrophy. Overall, our study highlights a novel link between satellite DNA repeats and transposon repression and lays the foundation for a more comprehensive understanding of satellite DNA function across tissues.

INTRODUCTION

Large tracts of eukaryotic genomes remain tightly condensed throughout the cell cycle and are collectively called 'constitutive heterochromatin'[1]. The majority of constitutive heterochromatin is found adjacent to the centromere and is largely comprised of simple non-coding tandem repeats known as satellite DNA [2] and interspersed repeats known as transposable elements [3]. Both types of repeats are thought to be engaged in intragenomic genetic conflicts, often propagating without benefitting overall organismal fitness. As such, transposable elements and satellite DNA are often dismissed as 'junk DNA'[4] and their abundance is typically attributed to their 'selfish' or 'parasitic' nature[5,6].

In contrast, accumulating evidence points to important functions for these repetitive sequences. Satellite DNA repeats at the centromere have been proposed to function in the stable propagation of centromere identity[7,8] while pericentromeric satellite DNA repeats are thought to be important for chromosome pairing[9,10], chromosome segregation[11,12], and genome encapsulation within the nucleus[13,14]. Protein-dependent organization of satellite DNA repeats within the nucleus is central to many of these functions. For example, we have previously demonstrated that the clustering of pericentromeric satellite DNA repeats from multiple chromosomes into nuclear foci known as chromocenters is central to genome encapsulation[13–15]. The formation and maintenance of chromocenters relies on sequence-specific satellite DNA-binding proteins, which include D1 and Prod in *Drosophila*[13,14,16–19] and HMGA1 in mammals[20–22]. In the absence of these proteins, pericentromeric satellite DNA clustering is strongly impaired in a tissue-specific manner, leading to micronuclei formation, defective genome encapsulation and cell death[13,14]. While these findings point to an important cellular role for satellite DNA clustering, our knowledge of the chromocenter-associated proteome is largely limited. To the best of our knowledge, satellite DNA-binding proteins have not been used as baits for *Drosophila* proteomic studies, with most studies rather focusing on constitutive heterochromatin

proteins[23–25] or centromeric proteins[24,26,27]. In addition, these studies have largely been performed in cultured cells and we still lack detailed knowledge of the heterochromatin-associated proteome in intact tissues. As a result, the full scope of satellite DNA function in multi-cellular organisms such as *Drosophila* remains enigmatic.

Transposable elements (TEs) are typically viewed as agents of genomic instability and their unchecked mobilization is usually associated with DNA damage, mutation, cell death and sterility[28]. At the same time, regulated activity of TEs can benefit cellular function[29] through roles in telomere maintenance[30], repeat copy number stability[31,32], genome restructuring[33] and the birth of new genes. In germline tissues from flies, as well as other metazoans, genome defense against transposons requires a dedicated small RNA pathway known as the Piwi-interacting RNA (piRNA) pathway[34,35]. In *Drosophila*, transcription of piRNA precursors is mediated by a protein complex comprised of Rhino, Cutoff and Deadlock (RDC complex) [36–38], followed by nuclear export of precursor transcripts and processing into mature piRNAs[34,35]. Mature piRNAs are loaded onto a family of RNA-binding Argonaute proteins including Piwi (P-element induced wimpy testis), a canonical member of the piRNA pathway [34,35]. piRNA-bound Piwi enters the nucleus, locates genomic TE loci through sequence complementarity and mediates TE repression through local heterochromatin formation and chromatin compaction. Consistently, heterochromatin-associated proteins such as Hybrid male rescue (Hmr)[39], Lethal hybrid rescue (Lhr)[39] and small ovary (sov)[40–42] are implicated in heterochromatin formation at transposon loci[40], leading to TE repression.

In addition to a cell-autonomous role in TE repression in adult gonads, piRNAs are also critical for transgenerational TE repression in both somatic and germ cells [43–45]. Specifically, *Drosophila* females deposit TE-targeting piRNAs in their eggs to ensure effective TE repression in their progeny. In a phenomenon known as hybrid dysgenesis, the

progeny of females and males carrying distinct sets of TEs are sterile with rudimentary gonads[46,47], since maternally deposited piRNAs in the egg are not complementary to the full set of paternal TEs. As a result, certain transposons evade repression and mobilize, trigger a DNA-damage checkpoint and block germline stem cell (GSC) differentiation, leading to a failure in gamete production [48,49]. A similar phenotype is observed in mutants of the heterochromatin-associated H3K9me3 depositing enzyme, dSETDB1, which plays an important role in piRNA biogenesis[50].

In this study, we set out to identify chromocenter-associated proteins using quantitative mass spectrometry. To do so, we used the satellite DNA-binding proteins, D1 and Prod, as baits and identified chromocenter-associated proteins from embryos, adult ovaries and GSC-enriched adult testes. We have identified ~500 proteins significantly enriched after affinity purification of D1 and Prod across all tissues, with a substantial proportion of novel hits. These hits include proteins associated with pericentromeric heterochromatin, chromatin remodeling and the nuclear pore complex. Interestingly, pathway analysis indicated that the chromocenter-associated proteome in embryos was enriched in DNA repair and transposon repression proteins. Following up on these data, we have identified a novel and unanticipated molecular relationship between satellite DNA organization and transgenerational repression of transposable elements in *Drosophila*. Taken together, our study lays the groundwork toward a better understanding of the multi-faceted contributions of satellite DNA repeats to cellular function.

RESULTS

Establishment and validation of strains carrying GFP-tagged satellite DNA-binding proteins

Our previous work has shown that the *Drosophila* DNA-binding proteins, D1 and Prod, cluster their cognate satellite DNA repeats (AATAT for D1, AATAACATAG for Prod) from multiple chromosomes into chromocenters, forming discrete protein-DNA 'modules' that are typically present next to one another within the larger HP1-positive constitutive heterochromatin domain of the nucleus [13–15]. We performed affinity purification coupled to mass spectrometry (AP-MS) to identify chromocenter-associated proteins in *Drosophila* embryos, ovaries and GSC-enriched testes using D1 and Prod as baits (Fig. 1A). In nuclei from these tissues, D1 and Prod foci were often found in proximity to each other (Fig. 1B), and we expected to identify common interactors between the two baits. To minimize variability between baits, we used an affinity purification strategy involving pulldown of GFP-tagged baits using GFP nanobody-coupled beads. For D1, we made use of an available strain where GFP has been inserted upstream of the endogenous locus. To validate this strain, we assessed the number of chromocenters (D1 and Prod foci per cell) in male and female GSCs as well as fertility in both sexes in comparison to a wild type reference strain (*D. melanogaster yw*). We observed that the number of D1 and Prod foci per cell in male and female GSCs was identical to the wild type (Fig. 1C, D), indicating that chromocenter formation was unperturbed in the GFP-D1 strain in these cell types. In addition, we observed that homozygous GFP-D1 males and females were fertile, although the fertility of homozygous GFP-D1 females is reduced compared to the wild type (Fig. 1E, F). For Prod, we first generated a transgenic strain where a GFP-tagged codon-optimized *Prod* was placed downstream of a 400 bp promoter sequence and inserted into the attP40 locus (Fig. S1A). Importantly, the *GFP-Prod* transgene fully rescued the larval lethality associated with loss-of-function *Prod* mutations, $prod^{k01180}$ and $prod^{U}$ [51](Fig. S1B).

Hereafter, GFP-Prod refers to the *GFP-Prod* transgene in a *Prod* mutant background (*p400-GFP-Prod*, $prod^{k01180}/prod^{U}$) (Fig. S1B). We observed that the number of GFP-Prod foci was at wild type levels in male and female GSCs (Fig. 1C, D), suggesting that the transgene does not interfere with chromocenter formation in these cells. Like the GFP-D1 strain, GFP-Prod males were fully fertile, while GFP-Prod females were sub-fertile in comparison to the wild type (Fig. 1E, F). As a negative control for GFP-D1 and GFP-Prod, we used a strain containing ubiquitously expressed nucleoplasmic GFP (NLS-GFP). The NLS-GFP did not enrich at chromocenters, and we observed that the number of chromocenters was comparable to the wild type, GFP-D1 and GFP-Prod strains (Fig. 1C, D). Therefore, the GFP-tagged D1 and Prod strains and the NLS-GFP control strain are an excellent starting point to characterize the chromocenter-associated proteome in intact *Drosophila* tissues.

Multi-tissue characterization of the chromocenter-associated proteome

We next affinity purified GFP-D1, GFP-Prod (and NLS-GFP as a control) from embryo, ovary and testis lysates and identified the associated proteins by quantitative mass spectrometry. While we were unable to analyze Prod pulldowns specifically in GSC-enriched testes due to poor sample quality, we identified ~500 proteins that were significantly enriched in the D1 and Prod pulldowns in comparison to the control pulldown of NLS-GFP across all tissues ($\log_2 FC>1$ and $p<0.05$, Fig. 2A-C, Fig. S2, Table S1-3). Importantly, our hits included known chromocenter-associated or pericentromeric heterochromatin-associated proteins, such as Su(var)3-9[52], ADD1[23,24], HP5[23,24], mei-S332[53], Mes-4[23], Hmr[24,39,54], Lhr[24,39], and members of the chromosome passenger complex, such as borr and Incenp[55] (Fig. 2A-B, Fig. S2, Table S1-3). We also observed a substantial overlap between D1- and Prod-associated proteins in embryos and ovaries (yellow points in Fig. 2A-B, Fig. 2C, Table S1-3) with approximately 20% of hits pulled down by both baits (blue arrowheads, Fig. 2C). A notable protein pulled down by both D1 and Prod in embryos was fizzy, the *Drosophila* orthologue of Cdc20, which is a

substrate adaptor for the APC/C ubiquitin ligase complex[56–58]. This potential interaction hints that APC/C-dependent ubiquitylation of D1/Prod may be important for chromocenter breakdown at the metaphase-to-anaphase transition to enable chromosome segregation. Another notable example in embryos was CG42232, the putative *Drosophila* orthologue of Ki67, which was also pulled down by both D1 and Prod. Ki67 is a DNA-binding protein that localizes to mammalian chromocenters[59,60] and is implicated in the compaction and spatial organization of heterochromatin during interphase[61]. In ovaries, both D1 and Prod were associated with nuclear pore complex (NPC) proteins, such as Nup98-96 and Nup93, likely reflecting the fact that many chromocenters are often present at the nuclear periphery[62].

Surprisingly, pulldown of the same protein (D1 or Prod) across different tissues enriched very few proteins in common (magenta arrowheads, Fig. 2C, Table S1-3). There were only a few exceptions such as the nucleoporin, Gle1, which was D1-associated in embryos and testes, while the uncharacterized DnaJ-like chaperone, CG5504, was associated with both D1 and Prod in embryos and ovaries (Fig. 2D). These results suggest that the protein composition of chromocenters may be tailored for specific cellular and tissue functions in *Drosophila*.

Next, we performed a hypergeometric test for pathway enrichment and found that DNA repair proteins were associated with chromocenters in multiple tissues (Fig. 2E, Table S4). For instance, D1 and Prod pulled down multiple proteins involved in DNA repair, such as Fancd2, Top2, DNA ligase 3, mus101, chrac-16, the RPA complex, Rfc38 and the fly orthologue of Tyrosyl DNA Phosphodiesterase 1 (glaikit/dTdp1), across the different tissues (Fig. 2A-B, Fig. S2, Table S4). These hits suggest that satellite DNA repeats may require specialized DNA repair pathways to counteract their inherent vulnerability to copy number changes[63]. Finally, we also observed that proteins involved in TE repression were highly

enriched in the chromocenter-associated proteome across tissues (Fig. 2E, Table S5). This effect was strongest in embryos, where D1/Prod pulled down canonical piRNA pathway proteins such as Piwi (Prod only) and Mael (D1 and Prod), and heterochromatin-associated TE repression proteins such as Su(var)3-9, sov, Hmr and Lhr (all D1 only). Consistent with our findings, a recent study identified Prod and other pericentromeric heterochromatin proteins as Piwi interactors in cultured fly ovarian somatic cells [64].

TE repression proteins are associated with chromocenters

Due to the importance of TE repression in genome stability as well as readily available tools[65], we focused on the relationship between chromocenters and TE repression. To validate the pulldown data, we used GFP-tagged Piwi and GFP-tagged sov strains to assess co-localization with D1 and Prod in embryos and early germ cells from ovaries and testes. Although Piwi interacts directly with constitutive heterochromatin proteins such as HP1[66], it is not visibly enriched in the heterochromatin domain but is rather present throughout the entire nucleus. Consistently, we observed that GFP-Piwi was nucleoplasmic in most cells and tissues analyzed and overlapped with the heterochromatic D1 and Prod foci (Fig. 3A, Fig. S3A, D). For GFP-sov, which is recruited to transposon loci by piRNA pathway protein Panoramix[40], we observed a punctate signal in embryonic nuclei and in male and female germ cells, which were frequently present at the boundaries of D1/Prod containing chromocenters (Fig. 3A, Fig. S3B, E). Similarly, the D1-interacting heterochromatin protein, Hmr, which contributes to TE repression, was also observed at the chromocenter boundary, in agreement with recent reports suggesting that Hmr may bridge centromeric and pericentromeric heterochromatin[24,27] (Fig. S3C, F). The proximity of sov and Hmr to D1/Prod foci support the notion of an association between piRNA pathway proteins and chromocenters. To further validate these associations, we used embryonic lysates from the GFP-Piwi and the NLS-GFP strains to identify Piwi interactors using quantitative MS (Fig. 3B, Table S6). In addition to D1 ($\log_2$FC=1.8, p=0.05), Piwi also pulled

down proteins identified as D1/Prod interactors in embryonic lysates (red points, Fig. 3B). Strikingly, four proteins (CG14715, fit, CG10208 and Ugt35D1) were high confidence hits from the Piwi, D1 and Prod pulldowns (Fig. 3C). All these proteins are relatively unstudied. CG14715 is an uncharacterized FK506-binding protein, while Ugt35D1 is a putative UDP-glycosyltransferase. Female independent of transformer (fit) is a sexually dimorphic small peptide that has been implicated in feeding control[67], while CG10208 is completely uncharacterized protein with no known domains or function.

We next predicted the likelihood of direct interactions between D1/Prod/Piwi and the four common interactors (Table S7, Supplementary Dataset 1) using AlphaFold2. Recently, this deep-learning approach has been used to predict structural models for over 65,000 pairs of experimentally validated protein-protein interactions [68]. These structural models were used to generate a confidence metric, termed the pDockQ score, which distinguishes between interacting and non-interacting proteins. Specifically, pDockQ > 0.23 indicates a medium confidence model, while pDockQ > 0.5 indicate a high confidence model[68]. Interestingly, 5/24 of our protein pairs had medium confidence (blue points) and 4/24 had high confidence (red points) pDockQ scores, most notably D1-Piwi (Fig. 3D). For the D1-Piwi model, we observed that a C-terminal linear peptide of the completely unstructured D1 protein appeared to thread through a central cleft in the Piwi protein (Fig. 3E), which is normally occupied by the small RNA in Argonaute proteins[69]. While this model suggested a potential interaction interface for D1-Piwi, alternate possibilities include bridging interactions, for example through CG10208, which may facilitate the association of satellite DNA-binding proteins with Piwi. Nevertheless, our data reinforce the conclusion that components of the TE repression machinery are associated with chromocenters and hint at a relationship between the piRNA pathway and satellite DNA organization within nuclei.

D1 and Prod do not contribute to the repression of germline and somatic transposable elements in the adult ovary

Even though we identified that key piRNA pathway components, such as Piwi, Sov and Mael, are associated with chromocenters in embryos, we first assessed the role of D1 and Prod in the repression of transposable elements in the germline and somatic cells of *Drosophila* ovaries, where piRNA pathway function is best characterized[70,71]. We used two established reporters of transposon activity in the germline cells (*GFP-Burdock-lacZ*) and somatic cells (*gypsy-lacZ*) of *Drosophila* ovaries[70]. In the gonadal soma, we used the *tj-gal4* driver to express RNAi against D1 and Prod in a genetic background containing the *gypsy-lacZ* reporter. Despite effective depletion of both D1 and Prod (Fig. S4A-B), we did not observe increased expression of the TE reporter (*gypsy-lacZ*) based on staining for beta-galactosidase (β-Gal) in comparison to the negative control (mCherry RNAi) (Fig. S5A). In contrast, RNAi-mediated depletion of Sov (positive control) resulted in β-Gal signal in the somatic follicle cells across multiple stages of oogenesis, indicating expression of the somatic TE reporter (Fig. S5A). Next, we tested whether loss of D1 or Prod in the germline resulted in increased expression of the germline TE reporter. Since flies that lack Prod fail to develop into adults, we depleted Prod using the germline-specific *nos-gal4* driver in a background containing *GFP-Burdock-lacZ*. Although we observed a strong reduction of Prod levels (Fig. S4C-D), we did not observe GFP signal in the ovaries, suggesting that the *GFP-Burdock-lacZ* reporter is not expressed (Fig. S5B). In contrast, depletion of aubergine (positive control) resulted in a strong germline GFP signal (Fig. S5B). For D1, we used a previously validated transposon insertion allele ($D1^{LL03310}$) in trans to a chromosomal deficiency spanning the D1 locus ($Df(3R)^{Bsc666}$) in a background containing *GFP-Burdock-lacZ*[13]. Here, we did not observe GFP signal in the germline cells of both *D1* mutant ovaries and the heterozygous controls (Fig. S5B). Overall, our data suggest that D1 and Prod do not directly contribute to the repression of transposable elements by the piRNA

pathway in the ovary. Our findings confirm results from previous genome-wide screens in adult gonads[70,71], which also did not identify a role for D1 or Prod in TE repression.

Satellite DNA clustering into chromocenters does not rely on piRNA pathway proteins

Guided by piRNAs, Piwi localizes to the TE loci and represses their transcription through local heterochromatin formation involving repressive epigenetic modifications. Repressive modifications, such as H3K9me2/3, can influence the organization of pericentromeric heterochromatin[12,72] and recent studies identifying piRNAs complementary to complex satellite DNA repeats [73] suggested that Piwi may promote repressive epigenetic modifications at satellite DNA and facilitate chromocenter formation. This idea was supported by prior studies showing that an interaction between Piwi and the heterochromatin protein, HP1, contributes to heterochromatin organization in both somatic and germ cells [66,74,75]. Since Piwi mutation results in a severe agametic phenotype in ovaries and testes, we instead used mutations of aubergine, which mediates piRNA amplification through the ping-pong mechanism[35]. In the absence of aubergine, total piRNA levels are decreased and piRNA-free Piwi fails to enter the nucleus[76]. We observed that aubergine mutation (aub^{HN2}/aub^{QC42}) as well as germline knockdown of aubergine resulted in a mild de-clustering of Prod foci in female GSCs but had no effect on the number of D1 foci in the same cells (Fig. S6A-C). These findings were consistent with our results from the few surviving germ cells in Piwi mutant ovaries ($piwi^{1}/piwi^{2}$, data not shown) suggesting that the piRNA pathway does not have an important role in chromocenter stability in female GSCs.

Mis-localization of the piRNA biogenesis protein Cutoff to satellite DNA in the absence of D1

Our data thus far suggest that D1 and Prod do not have a direct role in transposon repression in the adult germline. However, a recent study[77] led us to consider that the

organization of satellite DNA repeats may have an indirect effect on transposon repression. Specifically, a recent study by Baumgartner et al. identified that a ZAD-zinc finger DNA-binding protein, Kipferl, promotes the localization of the piRNA biogenesis complex consisting of Rhino, Deadlock and Cutoff (RDC complex) to piRNA clusters[77]. In the absence of Kipferl, the RDC complex strikingly mis-localizes to the complex 359bp satellite DNA array with Kipferl mutants exhibiting transposon de-repression and decreased fertility[77]. Although members of the RDC complex were not pulled down by either D1 or Prod, we wondered whether defective satellite DNA clustering may affect RDC localization. We used a strain expressing GFP-tagged Cutoff (EGFP-Cuff) to first assess RDC complex localization in ovarian nurse cells, where it licenses transcription of piRNA precursors. In these cells, pericentromeric satellite DNA repeats from all chromosomes are clustered into a single prominent DAPI-dense chromocenter, which is strongly enriched for the D1 protein (arrowhead, Fig. 4A). We observed that Cuff was localized to the periphery of the nurse cell chromocenter (like Sov and Hmr) and was also present at many sites across the nucleus, consistent with the presence of piRNA clusters at pericentromeric heterochromatin and chromosome arms (Fig. 4A, Fig. S7A-B). Unlike *prod* mutants, which die as larvae, *D1* mutants are viable, and their ovaries contain all developmental stages, including egg chambers with nurse cells, during their early life. Therefore, we examined Cuff localization in *D1* mutant nurse cells in comparison to the heterozygous control. In both the heterozygous control and *D1* mutant, DAPI-dense chromocenters are still detected in nurse cells (yellow arrowheads, Fig. 4B). However, we quantified an approximately 3-fold increase in chromocenter size in nurse cells lacking D1 (Fig. 4B-C), likely due to impaired satellite DNA clustering. Strikingly, this increase in chromocenter size in *D1* mutants was accompanied by a substantial mis-localization of Cuff to satellite DNA (Fig. 4B, D), where we quantified ~6-fold more Cuff at *D1* mutant chromocenters in comparison to the heterozygous control (Fig. 4B, D).

Loss of D1 triggers transgenerational transposon de-repression and gonadal atrophy

Despite the substantial mis-localization of Cuff in ovary nurse cells, we did not observe a strong effect on adult gonad development or fertility in *D1* mutants. However, maternal deposition of piRNAs is critical for repressing transposons in both germline and somatic cells of the next generation[43–45] and we hypothesized that Cuff mis-localization to satellite DNA in F1 *D1* mutants may reduce the pool of maternally deposited piRNAs complementary to transposons. Moreover, since our proteomic screen identified interactions between D1 and piRNA pathway proteins in the embryo (Fig. 2A), we wanted to investigate a context whether these interactions could not take place. We therefore assessed the gonad development in the progeny of F1 *D1* mutants (F2 *D1* mutant, Fig. 5A), which lack both maternal and zygotic D1 protein. In these flies, we observed a dramatic increase in agametic ovaries (Fig. 5B) with the vast majority of ovarioles either lacking any discernible Vasa-positive germ cells or containing only undifferentiated Vasa-positive germ cells, which did not develop into cysts (Fig. 5C, E). This phenotype is highly reminiscent of the one observed in agametic gonads from dysgenic hybrids, which lack maternal piRNAs complementary to invading transposons[46,48]. We therefore tested whether certain transposons escaped repression in F2 *D1* mutant ovaries using single molecule RNA *in situ* hybridization (smFISH). Here, we observed that the *copia* retrotransposon (but not *burdock* or *gypsy*) was highly expressed in ~70% of early germ cells from F2 *D1* mutant ovaries (Fig. 5D, F, G). In comparison, *copia* expression was not observed in the *yw* wild type strain or in the F1 *D1* mutant, which have intact ovaries (Fig. 5B). While ~90% of the *aubergine* mutant germaria (positive control) contained *copia* expressing germ cells (Fig. 5D, F), this expression was mostly restricted to differentiated germ cell cysts. In contrast, undifferentiated *aubergine* mutant germ cells (GSCs and cystoblasts, arrowheads, Fig. 5D) exhibited a markedly lower level of *copia* expression (Fig. 5G). Thus, the high level of transposon expression, occurring in the early germ cells of F2 *D1* mutants, may drive the loss of GSCs or block their differentiation, ultimately leading to agametic gonads.

Discussion

In this study, we have characterized the chromocenter-associated proteome across *Drosophila* embryos, ovaries and GSC-enriched testes using two known satellite DNA-binding proteins, D1 and Prod, as baits. We have identified ~500 interactions across three different tissues encompassing known heterochromatin-associated proteins, DNA repair enzymes, proteins involved in transposon repression and a host of uncharacterized proteins. Surprisingly, we observed a tissue-specific proteomic signature for chromocenters, with both D1 and Prod associating with many common proteins in embryos and ovaries. This unexpected finding suggests that the tissue-context or cellular function may be the primary determinant of the chromocenter proteome, rather than the underlying repeats.

Although we identified that D1 and Prod interact with multiple components of the piRNA pathway in the embryo, we failed to detect a direct role for D1 or Prod in transposon repression in adult gonads. Rather, we observed a striking effect on transposon de-repression and gonad development in the F2 *D1* mutant, which we postulate has at least two likely causes. First, mis-localization of Cuff in *D1* mutant mothers may lead to a reduction of piRNAs targeting transposons in the maternally deposited pool. Second, the absence of maternal and zygotic D1 protein in the embryos of F2 *D1* mutants may prevent interactions with the piRNA pathway proteins (identified in this study) and disrupt heterochromatin formation at transposon loci. A similar observation has been made for an interaction between Piwi and HP1 in early embryos, which was shown to be important for heterochromatin-dependent gene silencing in adult somatic cells[75].

While our study highlights that multiple proteins involved in transposon repression are associated with satellite DNA-binding proteins in embryos, how these interactions occur remains less clear. Here, the structural predictions from Alphafold 2 offer potential clues.

For example, the high pDockQ score of the D1-Piwi pair suggests that these proteins could interact directly. However, the C-terminus of D1 occupies the Piwi central cleft and would likely compete with piRNA binding; this mode of D1-Piwi binding would therefore suggest a regulatory function. Another option comes from the structural predictions of D1, Prod and Piwi and their four common interactors in embryos, CG10208, fit, CG14715 and Ugt35D1, which reveal some high scoring non-overlapping interfaces, which potentially hint at bridging proteins between D1, Prod and Piwi. Future work is needed to address the mechanism of these interactions in embryos and their functional significance in transposon repression in adult gonads.

Finally, our study showcases the underappreciated ability of satellite DNA to alter the distribution of DNA-binding proteins within nuclei. While the concept of repetitive sequences as a 'sink' for heterochromatin proteins has been previously proposed and studied[78–80], our data highlight that the function of satellite DNA-binding proteins is an important parameter that must be considered. Specifically, the extent of satellite DNA clustering may significantly regulate the accessibility of the 'heterochromatin sink' to other DNA-binding proteins. Given the accumulating evidence that chromocenters are phase separated biomolecular condensates[81–83], biophysical properties of chromocenters such as surface tension, viscosity and charge distribution may also play prominent roles in the selective partitioning of DNA-binding proteins to satellite DNA.

In summary, our study has characterized the chromocenter-associated proteome across three *Drosophila* tissues. We believe that the novel hits identified in this study establish a path toward illuminating the full scope of satellite DNA and chromocenter function in *Drosophila* as well as other eukaryotes.

Materials and Methods

Fly husbandry and strains. All fly stocks were raised on standard Bloomington medium at 25°C unless otherwise indicated. *D. melanogaster y w* was used as a wild type stock. *prod*^{U} (BDSC42686), *GFP-D1* (BDSC50850), *Ubi-NLS-GFP* (BDSC5826, BDSC5826), *Df(3R)*^{Bsc666} (BDSC26518), *aub*^{QC42} (BDSC4968), *aub*^{HN2} (BDSC8517), *piwi*^{1} (BDSC43637), *piwi*^{2} (BDSC43319), *UAS-Dcr-2* (BDSC24650), *mCherry* ^{RNAi} (BDSC35785) and *aub* ^{RNAi} (BDSC39026) were obtained from the Bloomington *Drosophila* stock center. *D1*^{LL03310} (DGRC140754) and *FRT42D prod*^{k08810} (DGRC111248) were obtained from the Kyoto stock center. *GFP-Burdock-lacZ* (v313212), *tj-gal4, gypsy-lacZ* (v313222), *GFP-Piwi* (v313321), *EGFP-Cuff* (v313269), *D1* ^{RNAi} (v101946) and *Prod* ^{RNAi} (v106593) were obtained from the Vienna Drosophila Resource Center. *Hmr-HA*[39] was a gift from Daniel Barbash. Sov-GFP (*sov*^{def1}; *Dp(1;2)*^{FF025056})[41] was a gift from Ferenc Jankovics. *nos-gal4* (second chromosome insertion) was a gift from Yukiko Yamashita. *nos-gal4* (third chromosome insertion)[84] and UAS-Upd[85] have been previously described.

Construction of the GFP-Prod strain. First, a 400bp promoter upstream of the Prod start site from *D. melanogaster* was PCR amplified using the following primer pair, GATCAAGCTTCTGTTGTTATGCATATCGTTC and GATCGAATTCCCGGGTATCCTTGCTC and subcloned into the HindIII and EcoRI sites on *pUASt-GFP-attB*, replacing the UAS sequence. Subsequently, a codon-optimized Prod sequence was inserted downstream of the GFP, and transgenic flies were generated using PhiC31 integrase-mediated transgenesis into the *attP40* site (BestGene). The GFP-Prod allele was generated by recombining the *p400-GFP-Prod* transgene with *prod*^{k08810}. Two virgin *prod*^{U} females were crossed to single tester male carrying either the *p400-GFP-Prod*, *prod*^{k08810} allele or the *prod*^{k08810} allele as a control in vials at 25°C. The percent of trans-heterozygous *prod* mutant flies was quantified in each replicate.

Fertility assay. For male fertility assays, a single tester male was maintained with two virgin *y w* females in vials for 7d and the number of resulting progenies was counted until 20 days post setup. For female fertility assays, a single tester female was maintained with two *y w* males in vials for 7d and the number of resulting progenies was counted until 20 days post setup. Vials that contained any deceased parent flies were omitted from the analysis.

Immunofluorescence staining. Immunofluorescence staining of *Drosophila* tissues was performed as follows. For fixation, adult tissues were dissected in PBS, transferred to 4% formaldehyde in PBS and fixed for 30 minutes. Fixation of *Drosophila* embryos was performed using a standard protocol. Briefly, collected embryos were moved to glass scintillation vials and washed three times with 1x PBS, dechorionated using 3 ml of 50% bleach solution for 90 secs by vigorous shaking and washed again in 1x PBS five times to remove the bleach. Dechorionated embryos were fixed using a heptane: 4% paraformaldehyde (PFA) (1:1) solution by vigorously shaking for 20 minutes. Following fixation, the bottom aqueous phase was removed and ice cold methanol was added to devitellinize the fixed embryos. After this step, embryos were washed three times in ice-cold methanol and stored in methanol at -20°C for further use. Prior to immunofluorescence staining, embryos were rehydrated using serial washes of methanol: PBS-T at 4:1, 1:1 and 1:4 volumes and followed by one wash in 1x PBS-T for 5 min each. For immunostaining, fixed tissues were washed in 1x PBS-T (PBS containing 0.1% Triton-X) for at least 60 minutes, followed by incubation with primary antibody in 3% bovine serum albumin (BSA) in PBS-T at 4°C overnight. Samples were washed three times in PBS-T, incubated with secondary antibody in 3% BSA in PBS-T at 4°C overnight, washed as above, and mounted in VECTASHIELD with DAPI (Vector Labs). Fluorescent images were taken using a Leica TCS SP8 confocal microscope with 63x oil-immersion objectives (NA=1.4). Images were processed using Adobe Photoshop or ImageJ.

RNA in situ hybridization. RNA *in situ* hybridization on ovaries was performed as previously described[86]. Briefly, 3-4 ovaries were dissected in RNase-free 1x PBS and fixed in 3.7% formaldehyde in PBS for 30 minutes on a nutator with gentle shaking. Following fixation, samples were washed with RNase-free 1x PBS three times and subsequently incubated in ice-cold ethanol at 4°C overnight on a nutator. The next day, samples were washed with RNase-free wash buffer (2XSSC, 10% formamide) for 3 minutes at room temperature and subsequently incubated with the hybridization mix (50-125 nM probes, 2x SSC, 10% dextran sulfate, 1g/L *E.coli* tRNA, 2mM vanadyl ribonucleoside complex, 0.5% Rnase-free BSA and 10% deionized formamide) at 37°C overnight. Samples were then washed twice (wash buffer) at 37°C for 30 minutes each and mounted in VECTASHIELD with DAPI (Vector Labs). Stellaris RNA FISH probes against the full length Burdock consensus sequence were synthesized (LGC Biosearch technologies). Stellaris RNA FISH probes targeting Het-A and Copia were gifts from Yukiko Yamashita. An oligo-dT probe was used to hybridize to polyadenylated RNA. Fluorescent images were taken using a Leica TCS SP8 confocal microscope with 63x oil-immersion objectives (NA=1.4). Images were processed using Adobe Photoshop.

Image Analysis. All image analysis was performed using ImageJ. For the quantitation of relative chromocenter size, we demarcated a region of interest (ROI^N) encompassing the entire nucleus as well as second ROI encompassing only the chromocenter (ROI^C) for each nurse cell (stages 4-6) analyzed. ROI^C/ROI^N equals the relative chromocenter size. For enrichment of EGFP-Cuff at chromocenters, we estimated the total EGFP-Cuff signal in ROI^C and divided it by the total EGFP-Cuff signal in ROI^N to give the percentage of EGFP-Cuff at chromocenters. For the normalized estimation of Copia expression, we estimated the average intensity of nuclear Copia (ROI^{Copia}) and normalized this value to the three cytoplasmic ROIs that estimated the average intensity of total mRNA based on the oligo-dT probe.

Isolation of nuclei from embryos. Flies reared in population cages were allowed to lay eggs on fresh yeast paste for ~18 hours. Embryos were collected using a set of sieves, washed with 1X PBS and dechorionated using 50% bleach. Next, embryos were moved to a dounce homogenizer containing light nuclear buffer (15 mM Tris-HCl pH 7.4, 60 mM KCl, 5 mM $MgCl_2$, 15 mM NaCl2 and 1 M sucrose) supplemented with 1x protease inhibitor cocktail (PIC, Roche,11697498001)and 1mM PMSF. The homogenate was then incubated on ice for 5 minutes and filtered through 2 layers of Miracloth tissue (Calbiochem). The filtered homogenate was then centrifuged for 20 min at 480 rpm at 4°C to pellet any debris. The supernatant containing nuclei was transferred into a fresh tube and the nuclei were pelleted at 4000 g for 30 min at 4°C. The pellet was then resuspended in 1 ml light nuclear buffer containing 0.1 mM EDTA and the nuclei were again centrifuged at 16000 g for 5 minutes before lysis.

Tissue lysis. Nuclei isolated from embryos were lysed for 90 min at 4°C with extraction buffer (15 mM Tris-HCl pH 7.4, 60 mM KCl, 5 mM $MgCl_2$, 450 mM NaCl2, 1% Triton-X [vol/vol], 500 units of benzonase) 1X PIC and 5 μl/ml PMSF). The resulting lysate was separated from the insoluble chromatin by centrifugation for 20 min at 14000 g at 4°C. For each ovary replicate, ~ 200 pairs were dissected from <1d old flies, flash frozen with liquid nitrogen and stored at -80°C prior to lysis. 500µl of ice-cold lysis buffer (450 mM NaCl, 50mM Tris-HCl (pH-7.5), 5% glycerol, 1mM MgCl2, 1mM EDTA, 0.2% NP40, 0.5mM DTT, 1x PIC, 5 μl/ml PMSF) was added to frozen ovaries, which were then homogenized with a motorized pestle until complete breakdown of the tissue. After homogenization, samples were incubated with benzonase (Sigma Aldrich, 2 µl/500 µl of lysate) for 90 minutes at 4°C on a rotating wheel. Following this, samples were first centrifuged for 3 min at 10000 g at 4°C to precipitate the debris from the tissues, homogenized again for 30 secs and centrifuged for 5 min at ~18700 g at 4°C. The supernatant (without the surface lipid layer) was transferred to a new 1.5 ml tube and centrifuged again for 5 min at ~18700g prior to

use. For each testis replicate, ~100 pairs of GSC-enriched (nos-gal4 > UAS-Upd) testis tumors were dissected, flash frozen with liquid nitrogen and stored at -80°C prior to lysis. 300 μl RIPA lysis buffer [450 mM NaCl, 0.5 % NP-40 [vol/vol], 0.5 % Sodium deoxycholate [mass/vol], 0.1 % SDS, 50 mM Tris pH 7.4] containing 1x protease inhibitor cocktail (PIC) (Roche,11697498001) and 5 μl/ml PMSF (Sigma-Aldrich, P7626) was added to frozen testes for 10 min. Tissue disruption was performed by applying 20 strokes with the tight pestle in a 1 ml dounce homogenizer, the samples were incubated for another 20 minutes on ice following which benzonase was added for 1 hour at 10°C. Homogenized lysates were centrifuged at 14000 g for 20 min at 4°C prior to use. Piwi-GFP and NLS-GFP embryos were collected as indicated in the previous section, dechorionated, flash frozen with liquid nitrogen and stored at -80°C prior to lysis. Embryo lysis for these samples was performed in an identical manner to the ovary lysis above.

Affinity purification and mass spectrometry. Affinity purification was performed using anti-GFP nanobody coupled magnetic beads (~ 30μl slurry) (Chromotek AB_2631358). Briefly, beads were collected using a magnetic rack, washed three times using the lysis buffer and incubated with the lysate for 90 minutes on a rotating wheel at 4°C. After the incubation, beads were washed three times for 5 minutes with a wash buffer (450 mM NaCl, 50mM Tris-HCl (pH-7.5), 1mM MgCl2, 1 mM EDTA, 0.5mM DTT) at room temperature. Washed beads containing the protein complexes were subjected to on-beads proteolysis. Samples were transferred into a 10kDa molecular weight cutoff spin column (Vivacon 500, Sartorious), following the FASP protocol[87]. Beads in solution were dried, denaturated (8M Urea), reduced (5mM TCEP, 30min 37C) and alkylated (10mM Iodoacetamide, 30min 37C). Beads were then washed three times with 50mM ammonium bicarbonate (250μl). During the buffer exchange, samples were centrifuged at 10000g. Subsequently, samples were proteolyzed with 0.5μg of Trypsin (Promega, sequencing grade) for 16h at 37°C. The proteolysis was quenched with 5% formic acid and peptides were subjected to C18

cleanup (BioPureSPN PROTO 300 C18, Nest group), following the manufacturer's procedure. The eluted peptides were then dried using a speedvac and resuspended in 20µl of 2% acetonitrile and 0.1% formic acid.

Mass spectrometry data acquisition. The MS datasets were acquired with two different instrumental setups. The first setup (A) was utilized for the D1, Prod and NLS-GFP samples from embryos, ovaries and testes. The second setup (B) was employed for acquiring the Piwi and NLS-GFP samples from embryos. In setup A, LC-MS/MS acquisition was performed on an Orbitrap QExactive+ mass spectrometer (ThermoFisher) coupled to an EASY-nLC-1000 liquid chromatography system (ThermoFisher). Peptides were separated using a reverse phase column (75µm ID x 400mm New Objective, in-house packed with ReproSil Gold 120 C18, 1.9 µm, Dr. Maisch GmbH) across a gradient from 3 to 25% in 170 minutes and from 25 to 40% in 10 minutes (buffer A: 0.1% (v/v) formic acid; buffer B: 0.1% (v/v) formic acid, 95% (v/v) acetonitrile). The DDA data acquisition mode was set to perform one MS1 scan followed by a maximum of 20 scans for the top 20 most intense peptides with MS1 scans (R=70'000 at 400m/z, AGC=1e6 and maximum IT=100 ms), HCD fragmentation (NCE=25%), isolation windows (1.8 m/z) and MS2 scans (R=17'500 at 400m/z, AGC=1e5 and maximum IT=55ms). A dynamic exclusion of 30s was applied and charge states lower than two and higher than seven were rejected for the isolation. In setup B, LC-MS/MS acquisition was performed on an Orbitrap Exploris480 mass spectrometer (ThermoFisher) coupled to an Vanquish Neo liquid chromatography system (ThermoFisher). Peptides were separated using a reverse phase column (75µm ID x 400mm New Objective, in-house packed with ReproSil Gold 120 C18, 1.9 µm, Dr. Maisch GmbH) across a linear gradient from 7 to 35% in 120 minutes (buffer A: 0.1% (v/v) formic acid; buffer B: 0.1% (v/v) formic acid, 95% (v/v) acetonitrile). The DDA data acquisition mode was set to perform one MS1 scan followed by a maximum of 20 scans for the top 20 most intense peptides with MS1 scans (R=60'000 at 400m/z, AGC=1e6 and auto maximum IT), HCD fragmentation

(NCE=28%), isolation windows (1.4 m/z) and MS2 scans (R=15'000 at 400m/z, AGC=2e5 and auto maximum IT). A dynamic exclusion of 25s was applied and charge states lower than two and higher than six were rejected for the isolation.

Mass spectrometry data analysis. Acquired spectra were analyzed using the MaxQuant software package version 1.5.2.8 embedded with the Andromeda search engine against *Drosophila* proteome reference dataset (http://www.uniprot.org/, downloaded on 18.01.2021, 22'044 proteins including not reviewed proteins) extended with reverse decoy sequences. The search parameters were set to include only full tryptic peptides, maximum two missed cleavage, carbamidomethyl as static peptide modification, oxidation (M) and deamidation (N-ter) as variable modification and "match between runs" option. The MS and MS/MS mass tolerance was set to 10ppm. False discovery rate of <1% was used for protein identification. Protein abundance was determined from the intensity of the top two unique peptides. Intensity values were median normalized and imputed using random sampling from a normal distribution-generated 1% less intense values. Statistical analysis was performed using unpaired two-sided t-test. Hits identified from the differential analysis between the bait pulldown versus the NLS-GFP control, with $\log_2$FC >1 and p-value < 0.05, were considered as interacting proteins. Surprisingly, Prod spectra were not always reliably identified in certain samples, even when used as bait. For pathway enrichment analysis, we first generated a list of *Drosophila* DNA repair proteins (GO:0006281, Table S4) and a custom list of transposon repression proteins based on the literature (Table S5) and used a hypergeometric test to calculate the enrichment of proteins belonging to these lists in the D1/Prod pulldowns.

Structural predictions using Alphafold2. To predict the structures of binary protein-protein interactions, we used version 2.3.2 of Alphafold2, which contains updated parameters for multimeric protein-protein interactions[88]. Default values for all parameters and for generating multiple sequence alignments were used. To rank the models according to a

confidence metric, we used the pDockQ score, which has been show to distinguish interacting from non-interacting proteins[89]. This score is defined by a sigmoidal equation incorporating the predicted accuracy of the model (pLDDT) and the number of amino acids within 10 Å of another chain:

$$\mathrm{pDockQ} = \frac{0.707}{1 + \mathrm{e}^{-0.03148(x-388.06)}} + 0.03138$$

where x = average interface pLDDT * log(number of interface contacts).

References

1. Jost KL, Bertulat B, Cardoso MC. Heterochromatin and gene positioning: inside, outside, any side? Chromosoma. 2012;121: 555–63. doi:10.1007/s00412-012-0389-2

2. Pardue ML, Gall JG. Chromosomal localization of mouse satellite DNA. Science. 1970;168: 1356–8.

3. Pimpinelli S, Berloco M, Fanti L, Dimitri P, Bonaccorsi S, Marchetti E, et al. Transposable elements are stable structural components of Drosophila melanogaster heterochromatin. Proc Natl Acad Sci U S A. 1995;92: 3804–8.

4. Ohno S. So much “junk” DNA in our genome. Brookhaven Symp Biol. 1972;23: 366–70.

5. Orgel LE, Crick FH. Selfish DNA: the ultimate parasite. Nature. 1980;284: 604–7.

6. Doolittle WF, Sapienza C. Selfish genes, the phenotype paradigm and genome evolution. Nature. 1980;284: 601–603. doi:10.1038/284601a0

7. Mellone BG, Fachinetti D. Diverse mechanisms of centromere specification. Current Biology. 2021;31: R1491–R1504. doi:10.1016/j.cub.2021.09.083

8. Kasinathan S, Henikoff S. Non-B-Form DNA Is Enriched at Centromeres. Mol Biol Evol. 2018;35: 949–962. doi:10.1093/molbev/msy010

9. Dernburg AF, Sedat JW, Hawley RS. Direct Evidence of a Role for Heterochromatin in Meiotic Chromosome Segregation. Cell. 1996;86: 135–146. doi:10.1016/S0092-8674(00)80084-7

10. Dernburg AF, Broman KW, Fung JC, Marshall WF, Philips J, Agard DA, et al. Perturbation of Nuclear Architecture by Long-Distance Chromosome Interactions. Cell. 1996;85: 745–759. doi:10.1016/S0092-8674(00)81240-4

11. Bouck DC, Bloom K. Pericentric Chromatin Is an Elastic Component of the Mitotic Spindle. Current Biology. 2007;17: 741–748. doi:10.1016/j.cub.2007.03.033

12. Peters AH, O’Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, et al. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell. 2001;107: 323–37.

13. Jagannathan M, Cummings R, Yamashita YM. A conserved function for pericentromeric satellite DNA. VijayRaghavan K, editor. eLife. 2018;7: e34122. doi:10.7554/eLife.34122

14. Jagannathan M, Cummings R, Yamashita YM. The modular mechanism of chromocenter formation in Drosophila. VijayRaghavan K, Gerbi SA, editors. eLife. 2019;8: e43938. doi:10.7554/eLife.43938

15. Brändle F, Frühbauer B, Jagannathan M. Principles and functions of pericentromeric satellite DNA clustering into chromocenters. Seminars in Cell & Developmental Biology. 2022;128: 26–39. doi:10.1016/j.semcdb.2022.02.005

16. Aulner N, Monod C, Mandicourt G, Jullien D, Cuvier O, Sall A, et al. The AT-Hook Protein D1 Is Essential for Drosophila melanogaster Development and Is Implicated in Position-Effect Variegation. MCB. 2002;22: 1218–1232. doi:10.1128/MCB.22.4.1218-1232.2002

17. Torok T, Gorjanacz M, Bryant PJ, Kiss I. Prod is a novel DNA-binding protein that binds to the 1.686 g/cm(3) 10 bp satellite repeat of Drosophila melanogaster. Nucleic Acids Res. 2000;28: 3551–7.

18. Rodriguez Alfageme C, Rudkin GT, Cohen LH. Isolation, properties and cellular distribution of D1, a chromosomal protein of Drosophila. Chromosoma. 1980;78: 1–31.

19. Levinger L, Varshavsky A. Protein D1 preferentially binds A + T-rich DNA in vitro and is a component of Drosophila melanogaster nucleosomes containing A + T-rich satellite DNA. Proceedings of the National Academy of Sciences. 1982;79: 7152–7156. doi:10.1073/pnas.79.23.7152

20. Strauss F, Varshavsky A. A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell. 1984;37: 889–901. doi:10.1016/0092-8674(84)90424-0

21. Solomon MJ, Strauss F, Varshavsky A. A mammalian high mobility group protein recognizes any stretch of six A.T base pairs in duplex DNA. Proceedings of the National Academy of Sciences. 1986;83: 1276–1280. doi:10.1073/pnas.83.5.1276

22. Radic MZ, Saghbini M, Elton TS, Reeves R, Hamkalo BA. Hoechst 33258, distamycin A, and high mobility group protein I (HMG-I) compete for binding to mouse satellite DNA. Chromosoma. 1992;101: 602–8.

23. Alekseyenko AA, Gorchakov AA, Zee BM, Fuchs SM, Kharchenko PV, Kuroda MI. Heterochromatin-associated interactions of *Drosophila* HP1a with dADD1, HIPP1, and repetitive RNAs. Genes Dev. 2014;28: 1445–1460. doi:10.1101/gad.241950.114

24. Kochanova NY, Schauer T, Mathias GP, Lukacs A, Schmidt A, Flatley A, et al. A multi-layered structure of the interphase chromocenter revealed by proximity-based biotinylation. Nucleic Acids Research. 2020;48: 4161–4178. doi:10.1093/nar/gkaa145

25. Swenson JM, Colmenares SU, Strom AR, Costes SV, Karpen GH. The composition and organization of Drosophila heterochromatin are heterogeneous and dynamic. Shilatifard A, editor. eLife. 2016;5: e16096. doi:10.7554/eLife.16096

26. Thomae AW, Schade GOM, Padeken J, Borath M, Vetter I, Kremmer E, et al. A Pair of Centromeric Proteins Mediates Reproductive Isolation in Drosophila Species. Developmental Cell. 2013;27: 412–424. doi:10.1016/j.devcel.2013.10.001

27. Lukacs A, Thomae AW, Krueger P, Schauer T, Venkatasubramani AV, Kochanova NY, et al. The Integrity of the Speciation Core Complex is necessary for centromeric binding and reproductive isolation in Drosophila. bioRxiv. 2021; 2021.02.05.429932. doi:10.1101/2021.02.05.429932

28. Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, et al. Ten things you should know about transposable elements. Genome Biology. 2018;19: 199. doi:10.1186/s13059-018-1577-z

29. Cosby RL, Chang N-C, Feschotte C. Host–transposon interactions: conflict, cooperation, and cooption. Genes Dev. 2019;33: 1098–1116. doi:10.1101/gad.327312.119

30. Cacchione S, Cenci G, Raffa GD. Silence at the End: How Drosophila Regulates Expression and Transposition of Telomeric Retroelements. Journal of Molecular Biology. 2020;432: 4305–4321. doi:10.1016/j.jmb.2020.06.004

31. Nelson JO, Slicko A, Yamashita YM. The retrotransposon R2 maintains Drosophila ribosomal DNA repeats. Proceedings of the National Academy of Sciences. 2023;120: e2221613120. doi:10.1073/pnas.2221613120

32. Chang C-H, Chavan A, Palladino J, Wei X, Martins NMC, Santinello B, et al. Islands of retroelements are major components of Drosophila centromeres. Becker PB, editor. PLoS Biol. 2019;17: e3000241. doi:10.1371/journal.pbio.3000241

33. Nowacki M, Higgins BP, Maquilan GM, Swart EC, Doak TG, Landweber LF. A Functional Role for Transposases in a Large Eukaryotic Genome. Science. 2009;324: 935–938. doi:10.1126/science.1170023

34. Ozata DM, Gainetdinov I, Zoch A, O'Carroll D, Zamore PD. PIWI-interacting RNAs: small RNAs with big functions. Nat Rev Genet. 2019;20: 89–108. doi:10.1038/s41576-018-0073-3

35. Czech B, Munafò M, Ciabrelli F, Eastwood EL, Fabry MH, Kneuss E, et al. piRNA-Guided Genome Defense: From Biogenesis to Silencing. Annual Review of Genetics. 2018;52: 131–157. doi:10.1146/annurev-genet-120417-031441

36. Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, et al. The Drosophila HP1 Homolog Rhino Is Required for Transposon Silencing and piRNA Production by Dual-Strand Clusters. Cell. 2009;138: 1137–1149. doi:10.1016/j.cell.2009.07.014

37. Zhang Z, Wang J, Schultz N, Zhang F, Parhad SS, Tu S, et al. The HP1 Homolog Rhino Anchors a Nuclear Complex that Suppresses piRNA Precursor Splicing. Cell. 2014;157: 1353–1363. doi:10.1016/j.cell.2014.04.030

38. Mohn F, Sienski G, Handler D, Brennecke J. The Rhino-Deadlock-Cutoff Complex Licenses Noncanonical Transcription of Dual-Strand piRNA Clusters in Drosophila. Cell. 2014;157: 1364–1379. doi:10.1016/j.cell.2014.04.031

39. Satyaki PR, Cuykendall TN, Wei KH, Brideau NJ, Kwak H, Aruna S, et al. The Hmr and Lhr hybrid incompatibility genes suppress a broad range of heterochromatic repeats. PLoS Genet. 2014;10: e1004240. doi:10.1371/journal.pgen.1004240

40. Andreev VI, Yu C, Wang J, Schnabl J, Tirian L, Gehre M, et al. Panoramix SUMOylation on chromatin connects the piRNA pathway to the cellular heterochromatin machinery. Nat Struct Mol Biol. 2022;29: 130–142. doi:10.1038/s41594-022-00721-x

41. Jankovics F, Bence M, Sinka R, Faragó A, Bodai L, Pettkó-Szandtner A, et al. Drosophila small ovary gene is required for transposon silencing and heterochromatin organization, and ensures germline stem cell maintenance and differentiation. Development. 2018;145: dev170639. doi:10.1242/dev.170639

42. Benner L, Castro EA, Whitworth C, Venken KJT, Yang H, Fang J, et al. Drosophila Heterochromatin Stabilization Requires the Zinc-Finger Protein Small Ovary. Genetics. 2019;213: 877–895. doi:10.1534/genetics.119.302590

43. Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. An Epigenetic Role for Maternally Inherited piRNAs in Transposon Silencing. Science. 2008;322: 1387–1392. doi:10.1126/science.1165171

44. Akkouche A, Grentzinger T, Fablet M, Armenise C, Burlet N, Braman V, et al. Maternally deposited germline piRNAs silence the tirant retrotransposon in somatic cells. EMBO reports. 2013;14: 458–464. doi:10.1038/embor.2013.38

45. Fabry MH, Falconio FA, Joud F, Lythgoe EK, Czech B, Hannon GJ. Maternally inherited piRNAs direct transient heterochromatin formation at active transposons during early Drosophila embryogenesis. Struhl K, editor. eLife. 2021;10: e68573. doi:10.7554/eLife.68573

46. Kidwell MG, Kidwell JF, Sved JA. Hybrid Dysgenesis in DROSOPHILA MELANOGASTER: A Syndrome of Aberrant Traits Including Mutation, Sterility and Male Recombination. Genetics. 1977;86: 813–33.

47. Bucheton A. I transposable elements and I-R hybrid dysgenesis in Drosophila. Trends in Genetics. 1990;6: 16–21. doi:10.1016/0168-9525(90)90044-7

48. Khurana JS, Wang J, Xu J, Koppetsch BS, Thomson TC, Nowosielska A, et al. Adaptation to P element transposon invasion in Drosophila melanogaster. Cell. 2011;147: 1551–63. doi:10.1016/j.cell.2011.11.042

49. Moon S, Cassani M, Lin YA, Wang L, Dou K, Zhang ZZ. A Robust Transposon-Endogenizing Response from Germline Stem Cells. Dev Cell. 2018;47: 660-671 e3. doi:10.1016/j.devcel.2018.10.011

50. Rangan P, Malone CD, Navarro C, Newbold SP, Hayes PS, Sachidanandam R, et al. piRNA Production Requires Heterochromatin Formation in Drosophila. Current Biology. 2011;21: 1373–1379. doi:10.1016/j.cub.2011.06.057

51. Torok T, Harvie PD, Buratovich M, Bryant PJ. The product of proliferation disrupter is concentrated at centromeres and required for mitotic chromosome condensation and cell proliferation in Drosophila. Genes & Development. 1997;11: 213–225. doi:10.1101/gad.11.2.213

52. Schotta G, Ebert A, Krauss V, Fischer A, Hoffmann J, Rea S, et al. Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J. 2002;21: 1121–1131. doi:10.1093/emboj/21.5.1121

53. Kerrebrock AW, Moore DP, Wu JS, Orr-Weaver TL. Mei-S332, a drosophila protein required for sister-chromatid cohesion, can localize to meiotic centromere regions. Cell. 1995;83: 247–256. doi:10.1016/0092-8674(95)90166-3

54. Blum JA, Bonaccorsi S, Marzullo M, Palumbo V, Yamashita YM, Barbash DA, et al. The Hybrid Incompatibility Genes Lhr and Hmr Are Required for Sister Chromatid Detachment During Anaphase but Not for Centromere Function. Genetics. 2017;207: 1457–1472. doi:10.1534/genetics.117.300390

55. Ainsztein AM, Kandels-Lewis SE, Mackay AM, Earnshaw WC. INCENP Centromere and Spindle Targeting: Identification of Essential Conserved Motifs and Involvement of Heterochromatin Protein HP1. Journal of Cell Biology. 1998;143: 1763–1774. doi:10.1083/jcb.143.7.1763

56. Sigrist S, Jacobs H, Stratmann R, Lehner CF. Exit from mitosis is regulated by Drosophila fizzy and the sequential destruction of cyclins A, B and B3. EMBO J. 1995;14: 4827–4838. doi:10.1002/j.1460-2075.1995.tb00164.x

57. Dawson IA, Roth S, Artavanis-Tsakonas S. The Drosophila cell cycle gene fizzy is required for normal degradation of cyclins A and B during mitosis and has homology to the CDC20 gene of Saccharomyces cerevisiae. Journal of Cell Biology. 1995;129: 725–737. doi:10.1083/jcb.129.3.725

58. Pfleger CM, Lee E, Kirschner MW. Substrate recognition by the Cdc20 and Cdh1 components of the anaphase-promoting complex. Genes Dev. 2001;15: 2396–2407. doi:10.1101/gad.918201

59. Bridger JM, Kill IR, Lichter P. Association of pKi-67 with satellite DNA of the human genome in early G1 cells. Chromosome Res. 1998;6: 13–24. doi:10.1023/A:1009210206855

60. Starborg M, Gell K, Brundell E, Höög C. The murine Ki-67 cell proliferation antigen accumulates in the nucleolar and heterochromatic regions of interphase cells and at the periphery of the mitotic chromosomes in a process essential for cell cycle progression. Journal of Cell Science. 1996;109: 143–153. doi:10.1242/jcs.109.1.143

61. Sobecki M, Mrouj K, Camasses A, Parisis N, Nicolas E, Llères D, et al. The cell proliferation antigen Ki-67 organises heterochromatin. Watt FM, editor. eLife. 2016;5: e13722. doi:10.7554/eLife.13722

62. Mayer R, Brero A, von Hase J, Schroeder T, Cremer T, Dietzel S. Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. BMC Cell Biol. 2005;6: 44. doi:10.1186/1471-2121-6-44

63. Charlesworth B, Sniegowski P, Stephan W. The evolutionary dynamics of repetitive DNA in eukaryotes. Nature. 1994;371: 215–220. doi:10.1038/371215a0

64. Onishi R, Sato K, Murano K, Negishi L, Siomi H, Siomi MC. Piwi suppresses transcription of Brahma-dependent transposons via Maelstrom in ovarian somatic cells. Science Advances. 2020;6: eaaz7420. doi:10.1126/sciadv.aaz7420

65. ElMaghraby MF, Tirian L, Senti K-A, Meixner K, Brennecke J. A genetic toolkit for studying transposon control in the Drosophila melanogaster ovary. Genetics. 2022;220: iyab179. doi:10.1093/genetics/iyab179

66. Brower-Toland B, Findley SD, Jiang L, Liu L, Yin H, Dus M, et al. Drosophila PIWI associates with chromatin and interacts directly with HP1a. Genes Dev. 2007;21: 2300–2311. doi:10.1101/gad.1564307

67. Sun J, Liu C, Bai X, Li X, Li J, Zhang Z, et al. Drosophila FIT is a protein-specific satiety hormone essential for feeding control. Nat Commun. 2017;8: 14161. doi:10.1038/ncomms14161

68. Burke DF, Bryant P, Barrio-Hernandez I, Memon D, Pozzati G, Shenoy A, et al. Towards a structurally resolved human protein interaction network. Nat Struct Mol Biol. 2023;30: 216–225. doi:10.1038/s41594-022-00910-8

69. Sheu-Gruttadauria J, MacRae IJ. Structural Foundations of RNA Silencing by Argonaute. Journal of Molecular Biology. 2017;429: 2619–2639. doi:10.1016/j.jmb.2017.07.018

70. Handler D, Meixner K, Pizka M, Lauss K, Schmied C, Gruber FS, et al. The Genetic Makeup of the Drosophila piRNA Pathway. Molecular Cell. 2013;50: 762–777. doi:10.1016/j.molcel.2013.04.031

71. Czech B, Preall JB, McGinn J, Hannon GJ. A Transcriptome-wide RNAi Screen in the Drosophila Ovary Reveals Factors of the Germline piRNA Pathway. Molecular Cell. 2013;50: 749–761. doi:10.1016/j.molcel.2013.04.007

72. Montavon T, Shukeir N, Erikson G, Engist B, Onishi-Seebacher M, Ryan D, et al. Complete loss of H3K9 methylation dissolves mouse heterochromatin organization. Nat Commun. 2021;12: 4359. doi:10.1038/s41467-021-24532-8

73. Wei X, Eickbush DG, Speece I, Larracuente AM. Heterochromatin-dependent transcription of satellite DNAs in the Drosophila melanogaster female germline. Weigel D, editor. eLife. 2021;10: e62375. doi:10.7554/eLife.62375

74. Giauque CC, Bickel SE. Heterochromatin-Associated Proteins HP1a and Piwi Collaborate to Maintain the Association of Achiasmate Homologs in Drosophila Oocytes. Genetics. 2016;203: 173–189. doi:10.1534/genetics.115.186460

75. Gu T, Elgin SCR. Maternal Depletion of Piwi, a Component of the RNAi System, Impacts Heterochromatin Formation in Drosophila. PLOS Genetics. 2013;9: e1003780. doi:10.1371/journal.pgen.1003780

76. Ma X, Wang S, Do T, Song X, Inaba M, Nishimoto Y, et al. Piwi Is Required in Multiple Cell Types to Control Germline Stem Cell Lineage Development in the Drosophila Ovary. PLOS ONE. 2014;9: e90267. doi:10.1371/journal.pone.0090267

77. Baumgartner L, Handler D, Platzer SW, Yu C, Duchek P, Brennecke J. The Drosophila ZAD zinc finger protein Kipferl guides Rhino to piRNA clusters. Bourc'his D, Struhl K, Zhang Z, editors. eLife. 2022;11: e80067. doi:10.7554/eLife.80067

78. Henikoff S. Dosage-dependent modification of position-effect variegation in Drosophla. BioEssays. 1996;18: 401–409. doi:10.1002/bies.950180510

79. Lemos B, Araripe LO, Hartl DL. Polymorphic Y Chromosomes Harbor Cryptic Variation with Manifold Functional Consequences. Science. 2008;319: 91–93. doi:10.1126/science.1148861

80. Janssen S, Cuvier O, Müller M, Laemmli UK. Specific Gain- and Loss-of-Function Phenotypes Induced by Satellite-Specific DNA-Binding Drugs Fed to Drosophila melanogaster. Molecular Cell. 2000;6: 1013–1024. doi:10.1016/S1097-2765(00)00100-3

81. Erdel F, Rademacher A, Vlijm R, Tünnermann J, Frank L, Weinmann R, et al. Mouse Heterochromatin Adopts Digital Compaction States without Showing Hallmarks of

HP1-Driven Liquid-Liquid Phase Separation. Molecular Cell. 2020; S1097276520300757. doi:10.1016/j.molcel.2020.02.005

82. Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. Phase separation drives heterochromatin domain formation. Nature. 2017;547: 241–245. doi:10.1038/nature22989

83. Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, et al. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature. 2017;547: 236–240. doi:10.1038/nature22822

84. Van Doren M, Williamson AL, Lehmann R. Regulation of zygotic gene expression in Drosophilaprimordialgerm cells. Current Biology. 1998;8: 243–246. doi:10.1016/S0960-9822(98)70091-0

85. Zeidler MP, Perrimon N, Strutt DI. Polarity determination in the Drosophila eye: a novel role for Unpaired and JAK/STAT signaling. Genes Dev. 1999;13: 1342–1353.

86. Lu KL, Nelson JO, Watase GJ, Warsinger-Pepe N, Yamashita YM. Transgenerational dynamics of rDNA copy number in Drosophila male germline stem cells. Spradling AC, editor. eLife. 2018;7: e32421. doi:10.7554/eLife.32421

87. Wiśniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat Methods. 2009;6: 359–362. doi:10.1038/nmeth.1322

88. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. bioRxiv; 2022. p. 2021.10.04.463034. doi:10.1101/2021.10.04.463034

89. Bryant P, Pozzati G, Elofsson A. Improved prediction of protein-protein interactions using AlphaFold2. Nat Commun. 2022;13: 1265. doi:10.1038/s41467-022-28865-w

Acknowledgements

We thank Daniel Barbash, Ferenc Jankovics, Yukiko Yamashita, the Bloomington Drosophila Stock Center, the Vienna Drosophila Resource Center, the Kyoto Drosophila Stock Center and the Developmental Studies Hybridoma Bank for reagents and resources. We thank members of the Jagannathan lab, Ulrike Kutay, Hugo Stocker and Anna Sintsova for discussion and comments on the manuscript. This work was supported by a project grant from the Swiss National Science Foundation (MJ, 310030_189131).

Legends

Figure 1. Experimental scheme for the characterization of the chromocenter-associated proteome. (A) Experimental scheme for the identification of D1- and Prod-associated proteins across multiple *Drosophila* tissues using quantitative mass spectrometry. (B) D1 and Prod are in proximity within the nucleus in multiple cell types including gastrulating embryos, female germ cells and male germ cells. (C) Bar plot of the total number of D1 foci present in female GSCs and male GSCs in the wild type, *NLS-GFP*, *D1-GFP* and *Prod-GFP* strains. n indicates the number of cells analyzed. (D) Bar plot of the total number of Prod foci present in female GSCs and male GSCs in the wild type, *NLS-GFP*, *D1-GFP* and *Prod-GFP* strains. n indicates the number of cells analyzed. (E, F) Bar plot of the number of progeny sired by a single tester male (E) or female (F) when mated with two wild type flies of the opposite sex over the course of one week. **** represents $p<0.0001$ from an ordinary one-way ANOVA. All error bars are SD and all scale bars are 5μm.

Figure 2. The chromocenter-associated proteome across *Drosophila* tissues. (A-B) Volcano plots of the D1-associated proteome in embryos (A) and ovaries (B). The dashed lines mark $\log_2FC>1$ and $p<0.05$; blue points indicate proteins involved in transposon repression, red points indicate proteins involved in DNA repair and yellow points indicate proteins pulled down by D1 and Prod. Number of biological replicates are indicated in the plot. (C) Intersection plot representing number of overlapping proteins in the D1 and Prod pulldowns from different tissues. Blue arrowheads indicate number of proteins common to D1 and Prod pulldown in the same tissue while magenta arrowheads indicate number of common proteins pulled down by D1 or Prod across different tissues (D) Enrichment of the selected proteins pulled down by D1 and Prod from more than one tissue. The size and colour of the circles are proportional to the $\log_2FC$ while the black outline indicates $p<0.05$.

(E) Hypergeometric test of DNA repair or transposon repression pathway proteins from the multi-tissue D1 and Prod interactome. Dashed line indicates p=0.05.

Figure 3. Molecular relationships between proteins involved in transposon repression and chromocenter formation (A) Piwi-GFP and Sov-GFP embryos were stained with D1/Prod (magenta) and DAPI (blue). Arrowheads indicate sov-GFP foci in proximity to chromocenter boundaries. (B) Volcano plot of the Piwi-associated proteome in embryos. The dashed lines mark $\log_2$FC>1 and p<0.05 while the red points indicate proteins identified in pulldowns of D1 or Prod from embryos. Number of biological replicates are indicated. (C) Enrichment of the selected proteins pulled down by D1, Prod and Piwi from embryos. The size and colour of the circles are proportional to the $\log_2$FC while the black outline indicates p<0.05. (D) pDockQ scores of the pairwise protein structural models based on Alphafold 2. Gray points indicate low confidence structural models (pDockQ<0.23), blue points indicate medium confidence structural models (0.23<pDockQ<0.5) and red points indicate high confidence structural models (pDockQ>0.5). (E) Representation of the D1-Piwi structural model predicted by Alphafold 2. A C-terminal linear peptide of D1 (311-355) threads through a cleft in the Piwi protein. All scale bars are 5μm.

Figure 4. D1 blocks localization of the piRNA biogenesis machinery to chromocenters. (A) A single nurse cell from an EGFP-Cuff ovary stained with D1 (magenta), Vasa (blue) and DAPI (gray). The arrowhead points to the DAPI-dense chromocenter. The insets highlight Cuff localization at the periphery of D1-containing chromocenters. (B) Localization of EGFP-Cuff in nurse cell nuclei from control ($D1^{LL03310}$/TM6C) and D1 mutant ($D1^{LL03310}/Df(3R)^{Bsc666}$) ovaries. Yellow arrowheads point to the location of the chromocenter. (C) Box-and-whisker plot of the size of nurse cell chromocenters relative to total nuclear size in control ($D1^{LL03310}$/TM6C) and D1 mutant ($D1^{LL03310}/Df(3R)^{Bsc666}$) ovaries. **** represents p<0.0001 from a student's t-test. (D) Box-and-whisker plot of EGFP-Cuff enrichment at chromocenters

relative to the rest of the nurse cell nucleus in control (*D1*^{LL03310}/TM6C) and D1 mutant (*D1*^{LL03310}/*Df(3R)*^{Bsc666}) ovaries. **** represents p<0.0001 from a student's t-test. All scale bars are 5μm.

Figure 5. Loss of D1 triggers transgenerational transposon de-repression and agametic gonads. (A) Crossing scheme used to generate F1 *D1* mutants and F2 *D1* mutants. (B) Brightfield images of ovaries from the indicated genotypes. Scale bars: 100μm. (C) Ovaries from the indicated genotypes were stained with Vasa (green) and DAPI (magenta). Scale bars: 5μm. (D) RNA *in situ* hybridization against Copia (green) and total polyadenylated mRNA (blue) was performed in ovaries of the indicated genotypes that were also stained for DAPI (magenta). Yellow and black arrowheads indicate undifferentiated germ cells in the *aubergine* mutant germarium. Scale bars: 5μm. (E) Percentage of germaria containing a normal population of germ cells, only germline stem cells/cystoblasts (GSC/CB only) or completely lacking germ cells in *yw* (n=48), *aubergine* mutant (n=66), F1 *D1* mutant (n=65) and F2 *D1* mutant (n=109). (F) Percentage of germaria with germ cells expressing Copia transcripts in *yw* (n=62), *aubergine* mutant (n=67), F1 *D1* mutant (n=71) and F2 *D1* mutant (n=50). (G) Box-and-whisker plot of copia expression normalized to total polyA mRNA in undifferentiated germ cells (GSCs/CBs) from *yw* (n=16), *aubergine* mutant (n=22), F1 *D1* mutant (n=18) and F2 *D1* mutant (n=29). **** represents p<0.0001 from one-way ANOVA.

Supplementary Figure 1. Schematic of the GFP-Prod allele. (A) A 400bp promoter sequence upstream of the Prod translation start site was placed upstream of a GFP-tagged and codon-optimized Prod sequence, inserted into the attP40 locus and recombined with the *prod*^{k08810} loss-of-function allele. (B) Rescue of prod mutant (*prod*$^{k08810/U}$) lethality was assessed using p400-GFP-Prod, *prod*^{k08810} (p400-GFP-Prod, 447 flies were counted from 5

crosses) and *prod*^{k08810} (Ctrl, 314 flies were counted from 5 crosses). Red line indicates median while dashed line indicates percent progeny for complete rescue.

Supplementary Figure 2: The D1-associated proteome in GSC-enriched testes. Volcano plot of the D1-associated proteome in GSC-enriched testes. The dashed lines mark $\log_2FC>1$ and $p<0.05$; blue points indicate proteins involved in transposon repression and red points indicate proteins involved in DNA. Number of biological replicates are indicated in the plot.

Supplementary Figure 3. Localization of D1 and Prod with respect to Piwi, Sov and Hmr in nurse cells and spermatogonia. Piwi-GFP (A, D), sov-GFP (B, E) and Hmr-HA (C, F) ovaries and testes were stained with D1/Prod (magenta), Vasa/Lamin (blue) and DAPI (gray). Arrowheads indicate sov or Hmr foci in proximity to chromocenter boundaries in individual nurse cells or spermatogonia. Scale bars: 5μm.

Supplementary Figure 4. Validation of D1 and Prod knockdowns in gonadal somatic cells and germline cells. (A, B) Individual egg chambers stained for DAPI (magenta) and Prod/D1 following *tj-gal4* mediated expression of *Prod* ^{RNAi} (KK104361) or *D1* ^{RNAi} (KK110307). *mCherry* ^{RNAi} was used as a control. (C, D) Anterior region of individual ovarioles were stained for DAPI (blue) and Prod (magenta) following *nos-gal4* mediated expression of *mCherry* ^{RNAi} (C) or *Prod* ^{RNAi} (D). Scale bars: 5μm.

Supplementary Figure 5. Loss of D1 or Prod does not trigger expression of transposon reporters in gonadal somatic cells or germ cells. (A) Individual egg chambers stained for beta-galactosidase (indicating expression of *gypsy-lacZ*) following *tj-gal4* mediated expression of *mCherry* ^{RNAi}, *sov* ^{RNAi} (HMC04875), *D1* ^{RNAi} (KK110307) or *Prod*

^{RNAi} (KK104361). Arrowhead indicates follicle cells expressing *gypsy-lacZ*. (B) Anterior region of individual ovarioles were stained for DAPI in the following strains carrying GFP-Burdock-lacZ: *nos* > *mCherry* ^{RNAi} (*mCherry* ^{GLKD}), *nos* > *Prod* ^{RNAi} (*Prod* ^{GLKD}), *nos* > *aub*^{RNAi} (HMS01945, *aub* ^{GLKD}), *D1*^{LL03310}*/TM6B* (Ctrl.) and *D1*^{LL03310}*/Df(3R)*^{Bsc666} (*D1* mutant). Arrowheads indicate germ cells expressing GFP-Burdock. Scale bars: 5μm.

Supplementary Figure 6. Loss of *aubergine* has a mild effect on Prod foci in female GSCs. (A) Female GSCs expressing *mCherry*^{RNAi} (*mCherry*^{GLKD}) or *aub*^{RNAi} (*aub*^{GLKD}) under the control of *nos-gal4* stained for Prod (magenta), Lamin (green) and Vasa (blue). Scale bars: 5μm. (B, C) Bar plot of the total number of Prod foci (B) or D1 foci (C) present per female GSC in *mCherry*^{GLKD}, *aub*^{GLKD}, *aubergine* heterozygous control (Ctrl.) and *aubergine* mutant (*aub*$^{QC42/HN2}$). n indicates the number of cells analyzed. ** indicates $p<0.01$ based on a one-way ANOVA.

Supplementary Figure 7. Localization of EGFP-Cuff relative to chromocenters. (A, B) EGFP-Cuff ovaries and testes were stained for D1/Prod (magenta), Vasa/Lamin (blue) and DAPI (gray). Individual nurse cells or spermatogonia are depicted. Scale bars: 5μm.

Supplementary Table 1. List of proteins detected by LC-MS/MS in embryonic lysates from GFP-D1, GFP-Prod and NLS-GFP strains. Fold change of average MS1 intensity and p-value for the GFP-D1 and GFP-Prod strains relative to the NLS-GFP control strain are shown for detected proteins. Proteins pulled down by GFP-D1 and GFP-Prod ($\log_2FC>1$, $p<0.05$) are highlighted in yellow.

Supplementary Table 2. List of proteins detected by LC-MS/MS in ovary lysates from GFP-D1, GFP-Prod and NLS-GFP strains. Fold change of average MS1 intensity and p-

value for the GFP-D1 and GFP-Prod strains relative to the NLS-GFP control strain are shown for detected proteins. Proteins pulled down by GFP-D1 and GFP-Prod ($log_2FC>1$, $p<0.05$) are highlighted in yellow.

Supplementary Table 3. List of proteins detected by LC-MS/MS in GSC-enriched testes lysates from GFP-D1, GFP-Prod and NLS-GFP strains. Fold change of average MS1 intensity and p-value for the GFP-D1 and GFP-Prod strains relative to the NLS-GFP control strain are shown for detected proteins. Proteins pulled down by GFP-D1 and GFP-Prod ($log_2FC>1$, $p<0.05$) are highlighted in yellow.

Supplementary Table 4. List of proteins associated with DNA repair in *Drosophila*.

Supplementary Table 5. List of proteins associated with transposon repression in *Drosophila*.

Supplementary Table 6. List of proteins detected by LC-MS/MS in embryonic lysates from GFP-Piwi and NLS-GFP strains. Fold change of average MS1 intensity and p-value for the GFP-Piwi strain relative to the NLS-GFP control strain is shown for detected proteins.

Supplementary Table 7. Predicted DockQ scores of protein pairs using Alphafold 2 structural models.

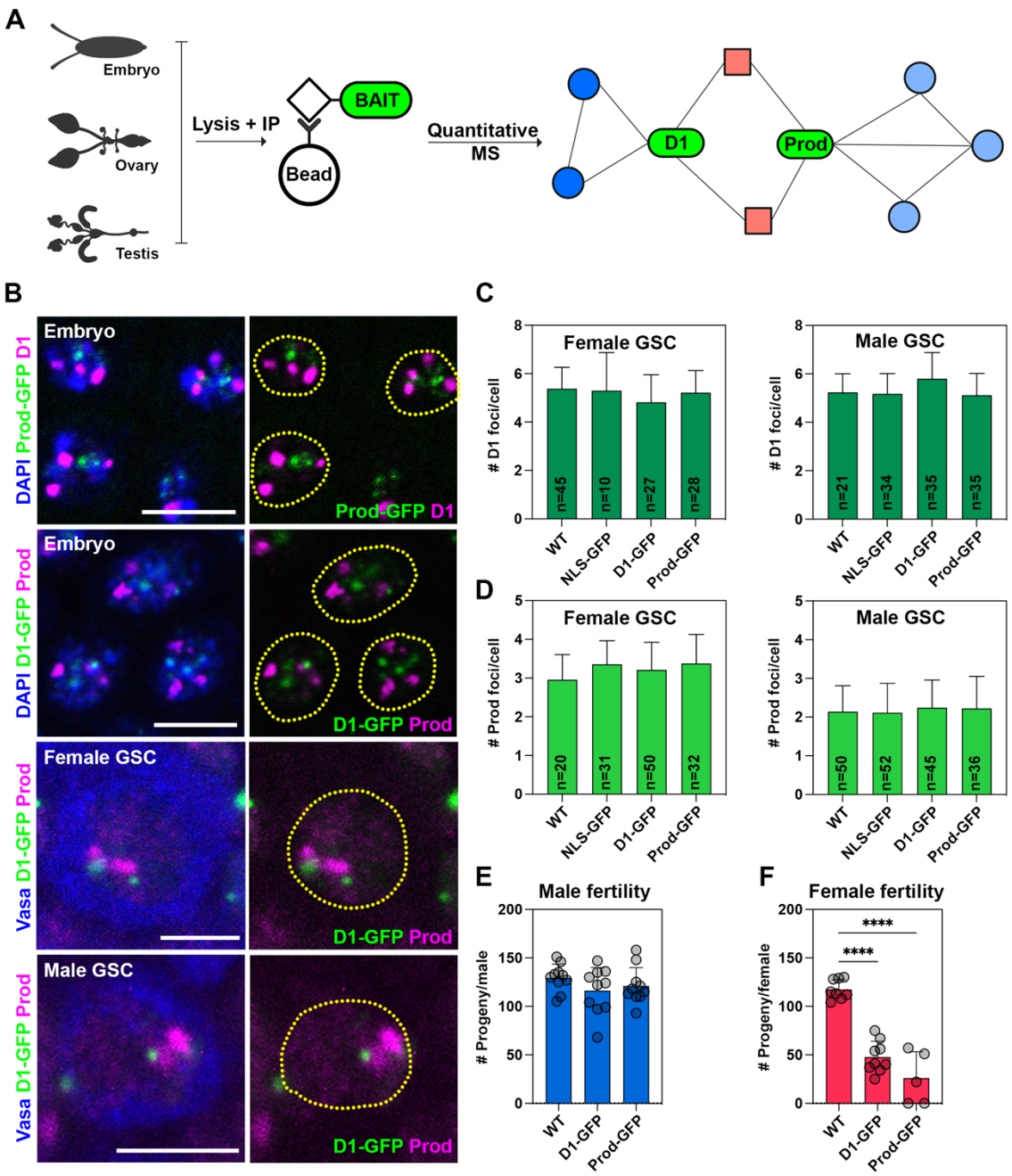

Figure 1: Experimental scheme for the characterization of the chromocenter-associated proteome. (A) Experimental scheme for the identification of D1- and Prod-associated proteins across multiple *Drosophila* tissues using quantitative mass spectrometry. (B) D1 and Prod are in proximity within the nucleus in multiple cell types including gastrulating embryos, female germ cells and male germ cells. (C) Bar plot of the total number of D1 foci present in female GSCs and male GSCs in the wildtype, *NLS-GFP*, *D1-GFP* and *Prod-GFP* strains. n indicates the number of cells analyzed. (D) Bar plot of the total number of Prod foci present in female GSCs and male GSCs in the wildtype, *NLS-GFP*, *D1-GFP* and *Prod-GFP* strains. n indicates the number of cells analyzed. (E, F) Bar plot of the number of progeny sired by a single tester male (E) or female (F) when mated with two wildtype flies of the opposite sex over the course of one week. **** represents $p<0.0001$ from an ordinary one-way ANOVA. All error bars are SD and all scale bars are 5μm.

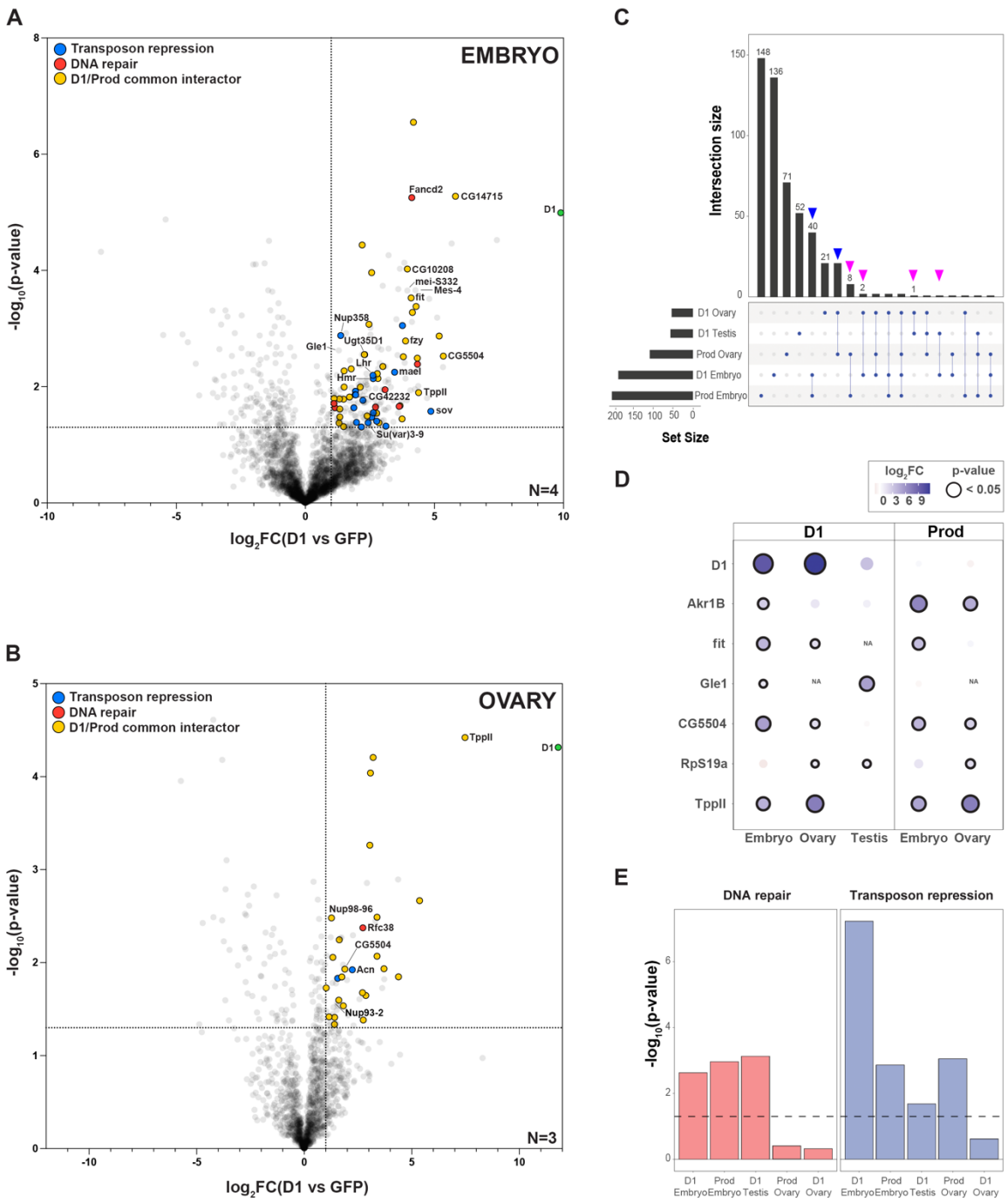

Figure 2: The chromocenter-associated proteome across *Drosophila* tissues. (A-B) Volcano plots of the D1-associated proteome in embryos (A) and ovaries (B). The dashed lines mark $\log_2FC>1$ and $p<0.05$; blue points indicate proteins involved in transposon repression, red points indicate proteins involved in DNA repair and yellow points indicate proteins pulled down by D1 and Prod. Number of biological replicates are indicated in the plot. (C) Intersection plot representing number of overlapping proteins in the D1 and Prod pulldowns from different tissues. Blue arrowheads indicate number of proteins common to D1 and Prod pulldown in the same tissue while magenta arrowheads indicate number of common proteins pulled down by D1 or Prod across different tissues (D) Enrichment of the selected proteins pulled down by D1 and Prod from more than one tissue. The size and colour of the circles are proportional to the $\log_2FC$ while the black outline indicates $p<0.05$.

(E) Hypergeometric test of DNA repair or transposon repression pathway proteins from the multi-tissue D1 and Prod interactome. Dashed line indicates p=0.05.

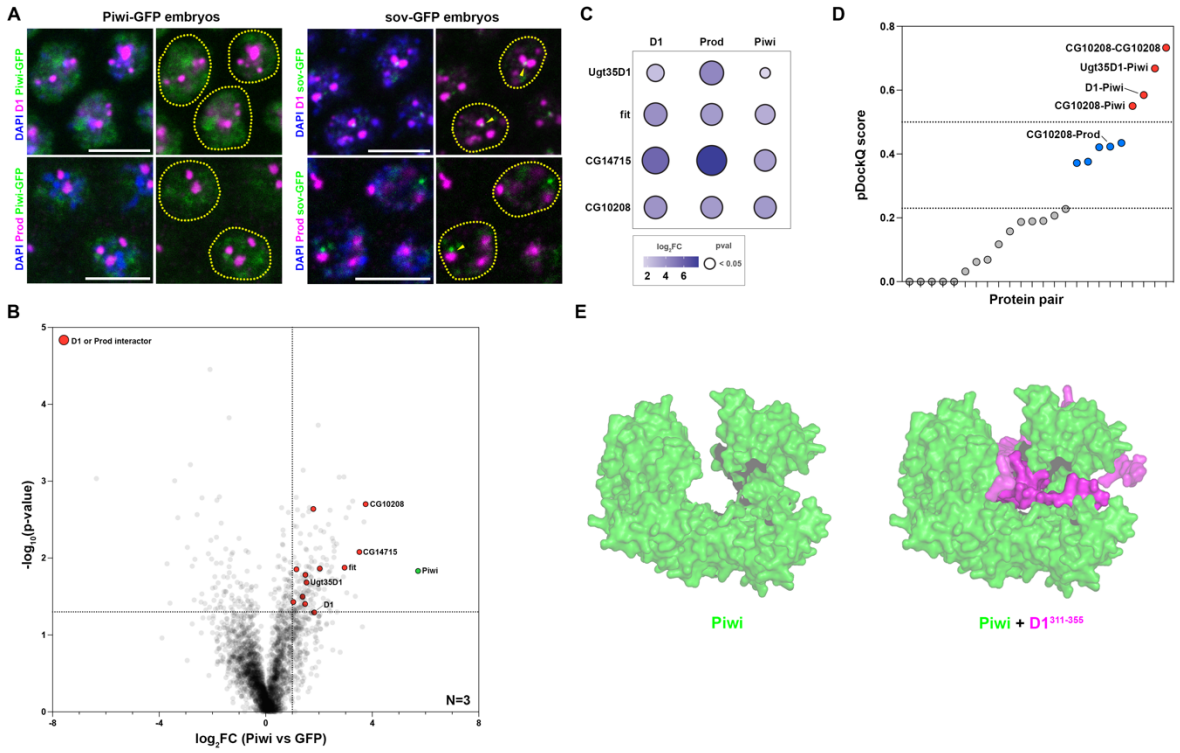

Figure 3. Molecular relationships between proteins involved in transposon repression and chromocenter formation (A) Piwi-GFP and Sov-GFP embryos were stained with D1/Prod (magenta) and DAPI (blue). Arrowheads indicate sov-GFP foci in proximity to chromocenter boundaries. (B) Volcano plot of the Piwi-associated proteome in embryos. The dashed lines mark $\log_2FC>1$ and $p<0.05$ while the red points indicate proteins identified in pulldowns of D1 or Prod from embryos. Number of biological replicates are indicated. (C) Enrichment of the selected proteins pulled down by D1, Prod and Piwi from embryos. The size and colour of the circles are proportional to the $\log_2FC$ while the black outline indicates $p<0.05$. (D) pDockQ scores of the pairwise protein structural models based on Alphafold 2. Gray points indicate low confidence structural models (pDockQ<0.23), blue points indicate medium confidence structural models (0.23<pDockQ<0.5) and red points indicate high confidence structural models (pDockQ>0.5). (E) Representation of the D1-Piwi structural model predicted by Alphafold 2. A C-terminal linear peptide of D1 (311-355) threads through a cleft in the Piwi protein. All scale bars are 5μm.

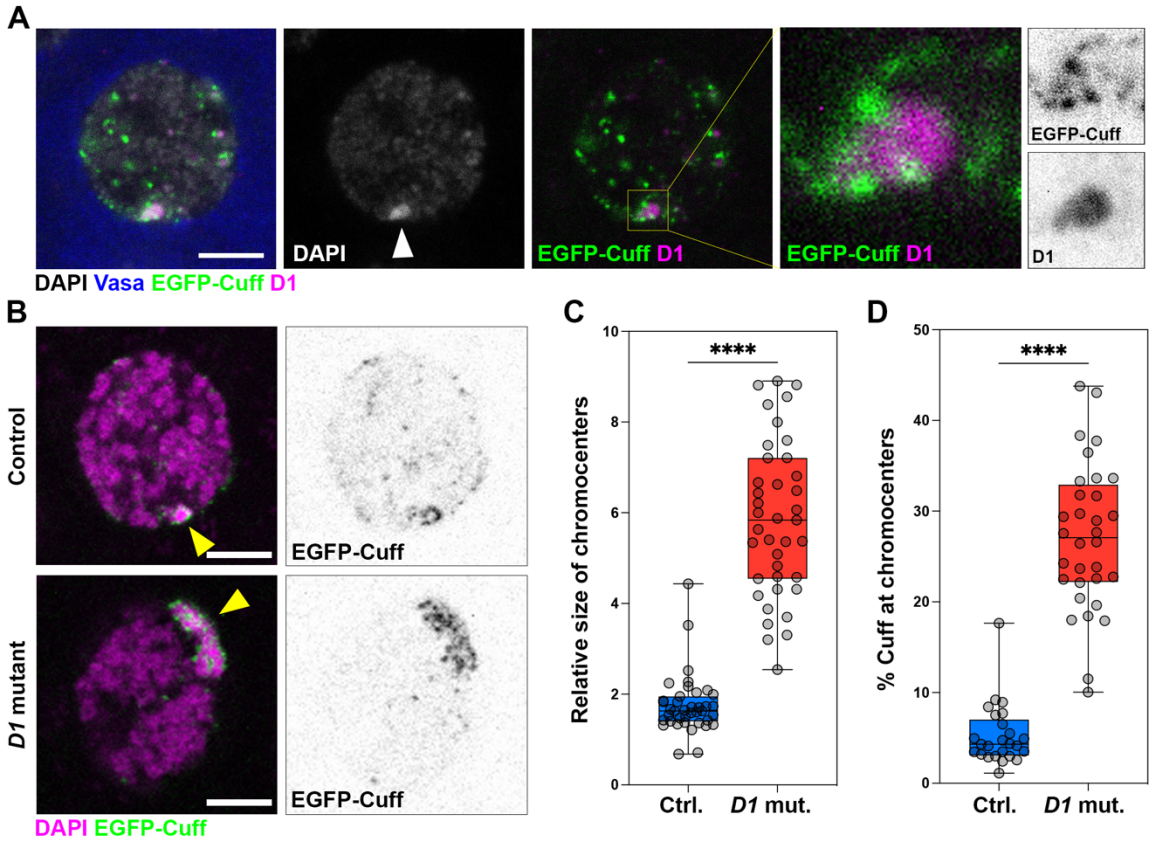

Figure 4. D1 blocks localization of the piRNA biogenesis machinery to chromocenters. (A) A single nurse cell from an EGFP-Cuff ovary stained with D1 (magenta), Vasa (blue) and DAPI (gray). The arrowhead points to the DAPI-dense chromocenter. The insets highlight Cuff localization at the periphery of D1-containing chromocenters. (B) Localization of EGFP-Cuff in nurse cell nuclei from control ($D1^{LL03310}$/TM6C) and D1 mutant ($D1^{LL03310}/Df(3R)^{Bsc666}$) ovaries. Yellow arrowheads point to the location of the chromocenter. (C) Box-and-whisker plot of the size of nurse cell chromocenters relative to total nuclear size in control ($D1^{LL03310}$/TM6C) and D1 mutant ($D1^{LL03310}/Df(3R)^{Bsc666}$) ovaries. **** represents $p<0.0001$ from a student's t-test. (D) Box-and-whisker plot of EGFP-Cuff enrichment at chromocenters relative to the rest of the nurse cell nucleus in control ($D1^{LL03310}$/TM6C) and D1 mutant ($D1^{LL03310}/Df(3R)^{Bsc666}$) ovaries. **** represents $p<0.0001$ from a student's t-test. All scale bars are 5μm.

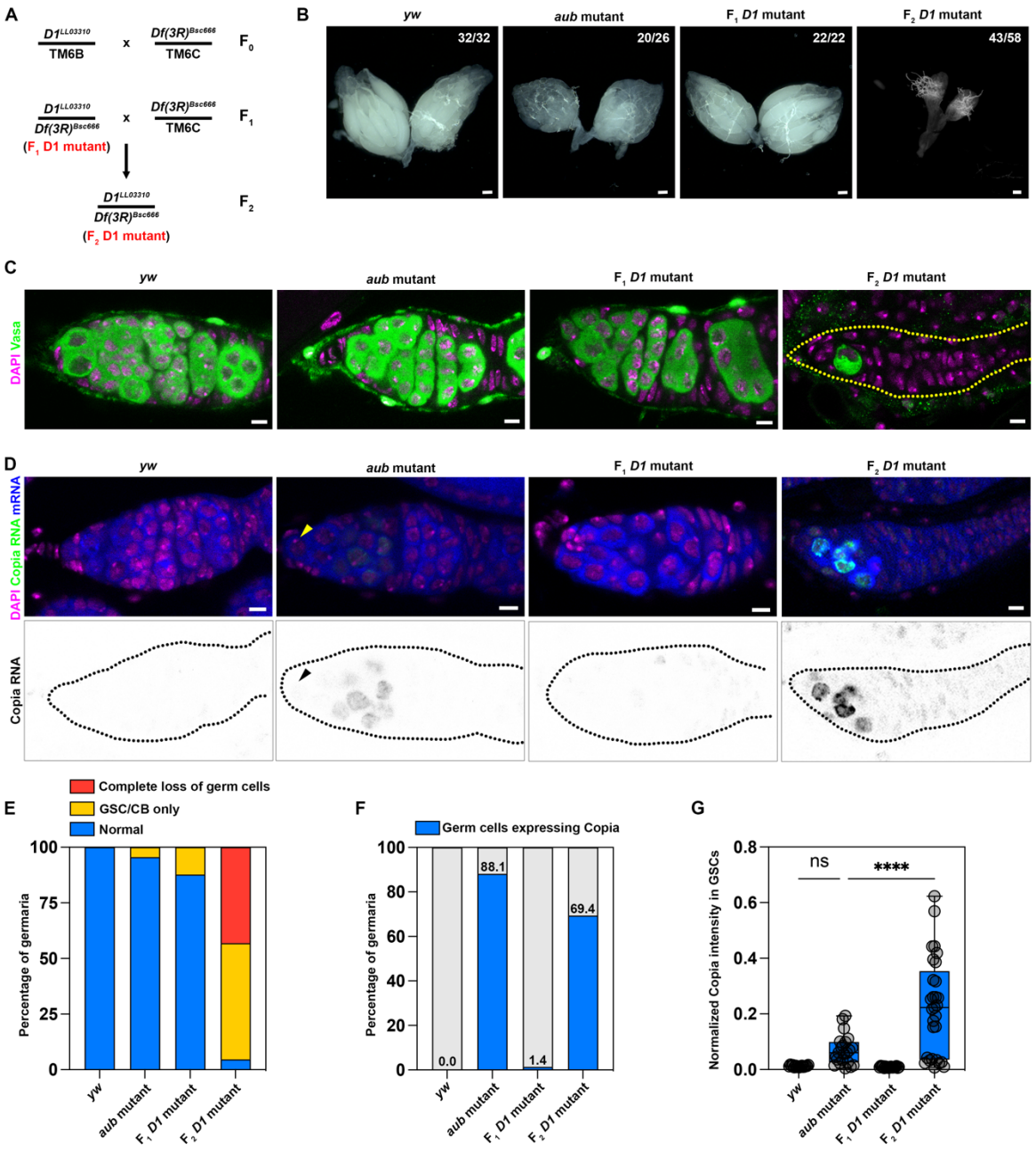

Figure 5. Loss of D1 triggers transgenerational transposon de-repression and agametic gonads. (A) Crossing scheme used to generate F1 *D1* mutants and F2 *D1* mutants. (B) Brightfield images of ovaries from the indicated genotypes. Scale bars: 100µm. (C) Ovaries from the indicated genotypes were stained with Vasa (green) and DAPI (magenta). Scale bars: 5µm. (D) RNA *in situ* hybridization against Copia (green) and total polyadenylated mRNA (blue) was performed in ovaries of the indicated genotypes that were also stained for DAPI (magenta). Yellow and black arrowheads indicate undifferentiated germ cells in the *aubergine* mutant germarium. Scale bars: 5µm. (E) Percentage of germaria containing a normal population of germ cells, only germline stem cells/cystoblasts (GSC/CB only) or completely lacking germ cells in *yw* (n=48), *aubergine* mutant (n=66), F1 *D1* mutant (n=65) and F2 *D1* mutant (n=109). (F) Percentage of germaria with germ cells expressing Copia transcripts in *yw* (n=62), *aubergine* mutant (n=67), F1 *D1* mutant (n=71) and F2 *D1* mutant (n=50). (G) Box-and-whisker plot of copia expression normalized to total polyA mRNA in undifferentiated germ cells (GSCs/CBs) from *yw* (n=16), *aubergine* mutant (n=22), F1 *D1* mutant (n=18) and F2 *D1* mutant (n=29). **** represents p<0.0001 from one-way ANOVA.