# Drosophila *small ovary* encodes a zinc-finger repressor required for ovarian differentiation

Leif Benner[1,2], Elias A. Castro[3], Cale Whitworth[1,4], Koen J.T. Venken[5], Haiwang Yang[1], Brian Oliver[1,6,7], Kevin R. Cook[4], and Dorothy A. Lerit[3,6]

**Author affiliations:**

[1]Section of Developmental Genomics, Laboratory of Cellular and Developmental Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda MD 20892. [2]Department of Biology, Johns Hopkins University, Baltimore MD 21218. [3]Department of Cell Biology, Emory University School of Medicine, Atlanta GA 30322. [4]Department of Biology, Indiana University, Bloomington, Indiana 47405. [5]Verna and Marrs McLean Department of Biochemistry and Molecular Biology, McNair Medical Institute at the Robert and Janice McNair Foundation, Dan L. Duncan Cancer Center, Center for Drug Discovery, Department of Pharmacology and Chemical Biology, Baylor College of Medicine, Houston, TX, 77030, USA

[6]**Corresponding authors:** dlerit@emory.edu briano@nih.gov

[7]**Lead contact**: briano@nih.gov

**Keywords**: chromatin, differentiation, gene expression, HP1, oogenesis, position-effect variegation

## Abstract

Repression is essential for coordinated cell type-specific gene regulation and for controlling the expression of transposons. In the *Drosophila* ovary, stem cells regeneration and differentiation require well-controlled expression, with de-repression leading to tissue degeneration and ovarian tumors. Additionally, the ovary is acutely sensitive to deleterious consequences of transposon de-repression. The *small ovary* (*sov*) locus was identified in a female sterile screen shows dramatic effects ovarian morphogenesis. We mapped the locus to the uncharacterized gene *CG14438* and reveal it encodes a zinc-finger protein that colocalizes with the essential Heterochromatin Protein 1 (HP1a). The distinct functions of Sov we report, including repression of inappropriate signaling, transposon silencing, and effects on position-effect variegation in the eye, suggest a central role in heterochromatin stabilization.

## Introduction

Multicellular organisms rely on stem cells. A combination of extrinsic and intrinsic signals function to maintain the balance between self-renewal of stem cells and the differentiation of progenitors needed for tissue homeostasis and development. The *Drosophila* adult ovary is a well-studied model for development (Fuller and Spradling 2007). It is organized into ~16 ovarioles, which are assembly lines of progressively maturing egg chambers. At the anterior tip of each ovariole is a structure called the germarium, harboring 2-3 germline stem cells (GSCs), which sustain egg development throughout the life of the animal. Surrounding the GSCs and their cystoblast daughters is a stem cell niche, a collection of nonproliferating somatic cells composed of terminal filament and cap cells. Like the GSCs, somatic stem cells divide to produce new stem cells and escort cells, which encase the germline. Each cystoblast is fated to differentiate and will undergo four rounds of cell division with incomplete cytokinesis to

produce a germline cyst interconnected by the fusome, a branched cytoskeletal structure. Germline cysts are surrounded by two escort cells. Toward the posterior of the germarium, a monolayer of somatic follicle cells replaces the escort cells to form an egg chamber.

Egg chamber development requires the careful coordination of distinct germline and somatic cell populations through signaling within a complex microenvironment. For example, extrinsic signals and transducers including Jak/Stat and the BMP homolog, Decapentapalegic (Dpp), are required to maintain the proliferative potential of GSCs (Bausek 2013; Gilboa 2015; S. Chen, Wang, and Xie 2011). Oriented divisions displace the primary cystoblast cell away from the stem cell preserving signals and permit the expression of differentiation factors, such as *bag-of-marbles* (*bam*). The importance of bam is revealed in bam loss-of-function mutants, resulting in tumorous germaria filled with undifferentiated germline cells containing many spectrosomes, the dot-fusome structure characteristic of GSCs and primary cystoblasts, as a result of complete cytokinesis (Lin and Spradling 1995; D. McKearin and Ohlstein 1995). Controlling this signaling requires repression in addition to activation.

Dynamic changes in chromatin landscapes within the germline and somatic cells are critical for oogenesis (Barton et al. 2016; X. Li et al. 2017; Börner et al. 2016; Peng et al. 2016; Soshnev et al. 2013; McConnell, Dixon, and Calvi 2012). Nucleosomes generally repress transcription by competing for DNA binding with transcription factors (Lorch and Kornberg 2017; Kouzarides 2007; Jenuwein and Allis 2001). Further levels of repression depend largely on histone modifications. For example, the activities of both the H3K9 methyltransferase SETDB1, encoded by *eggless* (*egg*) (X. Wang et al. 2011; Clough et al. 2007; Clough, Tedeschi, and Hazelrigg 2014), and the H3K4 demethylase encoded by *lysine-specific demethylase 1* (Di Stefano et al. 2007; Rudolph et al. 2007; Eliazer, Shalaby, and Buszczak

2011; Eliazer et al. 2014) are required in the somatic cells of the germarium for GSC maintenance, normal patterns of differentiation, and germline development. Genome-wide profiling hints at a progression from open chromatin in germline stem cells to a more closed state during egg chamber differentiation (T. Chen and Dent 2014). Disrupting this progressive repression in the ovary could result in stem cell over-proliferation and defective development.

There are other roles for regulated chromatin states in oogenesis. The germline has been repeatedly hijacked to promote the vertical transmission of transposons resulting in a loss of fitness in the *Drosophila* host (Charlesworth and Langley 1989; Pelisson et al. 2002). This increased mobilization of transposons in gonads is countered by host responses, such as induced heterochromatin formation, to repress transposon activity. The proximity of condensed heterochromatin can dampen the activity of genes and transgenes (Elgin and Reuter 2013). The conserved Heterochromatin-Protein 1a (HP1a), encoded by *Su(var)205*, is critical for heterochromatin formation (James and Elgin 1986; Eissenberg et al. 1990; Clark and Elgin 1992).

Originally described over 40 years ago (J. D. Mohler 1977), the *small ovary* (*sov*) locus mutants have range of phenotypes including disorganized ovariole structure, egg chamber fusions, and undifferentiated tumors, culminating in complete ovarian degeneration (Wayne et al. 1995). In the present work, we demonstrate *sov* encodes an unusually long 21 $C_2H_2$ Zinc finger (ZnFs) nuclear protein that colocalizes with HP1. Previous work showed this ZnF protein in complex with HP1a (Alekseyenko et al. 2014). RNA-seq analysis shows that *sov* activity is required to repress the expression of a large number of genes in the ovary as well as transposons. Additionally, *sov* mutations are strong dominant suppressors of Position Effect Varigation (PEV) in the eye, a key characteristic of HP1 (Elgin and Reuter 2013), which is

consistent with Sov and HP1a colocalization in the nucleus. These data indicate that Sov is a potent repressor of gene expression.

## Results

### *small ovary is CG14438*

To localize *sov*, we used complementation of female-sterile and lethal *sov* mutations with preexisting and custom-generated deficiencies, duplications, and transposon insertions to refine previous mapping (J. D. Mohler 1977; D. Mohler and Carroll 1984). Our results (Fig 1A,B) suggested that either the protein-coding *CG14438* gene or the intronic, noncoding *CR43496* gene was *sov*. The *sov* locus mapped to the non-complementing deletion *Df(1)sov*. Rescue of female sterility and/or lethality of $sov^{2}$, $sov^{ML150}$, $sov^{EA42}$ and *Df(1)sov* with $Dp(1;3)sn^{13a1}$ (molecularly undefined, not shown), *Dp(1;3)DC486*, and *PBac{GFP-sov}* confirmed our mapping. We could not replicate separability of $sov^{EA42}$ sterility and lethality with $Dp(1;3)sn^{13a1}$ (Wayne et al. 1995). Female sterile and lethal alleles of *sov* map identically.

We performed genome sequencing to determine if the lesions in *sov* alleles occur in *CG14439* or *CR43496* (Table S1). While *CR43496* contained three polymorphisms in *sov* alleles relative to the reference, we identified the same polymorphisms in all mutant *sov* alleles and in $sov^{+}$ alleles in $w^{1118}$ and *Oregon-R* lines. Thus, *CR43496* is highly unlikely to be *sov.* In contrast, we found disruptive mutations in *CG14438* (Fig 1C), which encodes an unusually long 3,313 residue protein with 21 $C_2H_2$ ZnFs, nuclear localization motifs, and coiled-coil regions (Fig 1C). We identified a nonsense mutation (G to A at position 6,764,462) in the *CG14438* open reading frame of $sov^{EA42}$ which is predicted to truncate the *sov* protein before the ZnF domains. We found a frameshift insertion (T at position 6,769,742) located towards the

118 end of *CG14438* in *sov*$^{2}$ that encodes 30 novel residues followed by a stop codon within the C-
119 terminal ZnF, as well as removing the terminal predicted NLS motif (Fig 1C). We found a
120 missense mutation (C to G at position 6,763,888) in *sov*$^{ML150}$ that results in a glutamine to
121 glutamate substitution within a predicted coiled-coil domain. While this is a conservative
122 substitution, and while glutamate residues are common in coiled-coil domains, glutamine to
123 glutamic acid substitutions are especially disruptive in the coiled-coil region of the sigma
124 transcription factor (Hsieh, Tintut, and Gralla 1994). We conclude that *CG14438* is *sov*.
125
126 *sov expression*
127 To determine where *sov* is expressed and if it encodes multiple isoforms, we analyzed
128 expression in adult tissues by RNA-Seq. We noted that *sov* is broadly expressed as a single
129 mRNA isoform in adult tissues, with highest expression in ovaries (Fig 1D). The modENCODE
130 (Graveley et al. 2011; J. B. Brown et al. 2014) and FlyAtlas (Robinson et al. 2013; Leader et al.
131 2018) reference sets show a similar enrichment in the ovary, as well as in early embryos. The
132 enrichment of *sov* in the ovary is consistent with its role in oogenesis.
133 To determine if and where Sov was expressed in the ovary, we examined the
134 distribution of GFP-Sov in developing egg chambers (Fig 2A). We observed GFP-Sov signal in
135 all cell types. By pairing localization analysis of Sov with antibodies recognizing Vasa (Vas) to
136 label the germline (Fig 2B), we observed particularly striking nuclear localization of Sov
137 surrounded by perinuclear Vas in the germline cells within region 1 of the germarium, with
138 highest levels evident within the GSCs (dashed circles). We also used Traffic jam (Tj) to label
139 the soma (Fig 2C). We observed Sov nuclear staining in the Tj positive somatic and follicle

stem cells (dashed ovals) (Fig 2C). These data indicate that Sov is expressed in the germline and soma in the ovary.

*Sov function*

To examine *sov* function in the germarium, we compared a strong allelic combination, $sov^{2}/sov^{EA42}$ to cell-type specific knockdown of *sov* using a UAS short hairpin *sov* RNAi construct (P{TRiP.HMC04875}; hereafter, $sov^{RNAi}$). The germaria of heterozygous *sov* females are wild type (Fig 3A), but the *sov* mutants often show an ovarian tumor phenotype, with greater than the expected number of germ cells with dot spectrosomes (Fig 3B), indicating that germ cells undergo complete cytokinesis. Ovarian tumors were observed in ~40% of *sov* mutants (N=12/31 $sov^{EA42}/sov^{2}$ egg chambers versus N=0/47 controls). These findings are consistent with GSC hyperproliferation and/or failed differentiation. We also observed tumors and nurse cell nuclei residing within common follicles (Fig 3B) in about one-third of *sov* mutants (N=11/31 $sov^{EA42}/sov^{2}$ versus N=0/24 in controls). This phenotype occurs when follicle cells either fail to separate egg chambers, or where those chambers fuse (Goode, Wright, and Mahowald 1992). In older *sov* ovaries, we observed extensive cell death (Fig S1). Additionally, we found that rare $sov^{EA42}/Y$ males that escaped lethality had no germ cells (Fig S2). Thus, *sov* functions widely.

To examine cell-type specific functions of *sov*, we used *tj-GAL4* to express $sov^{RNAi}$ in somatic escort and follicle cells and *nanos-GAL4* (*nos-GAL4*) to express $sov^{RNAi}$ (or $mCherry^{RNAi}$ controls) in the germline. Relative to the controls, the $tj>sov^{RNAi}$ germaria (Fig 3C,D) were abnormal, showing ovarian tumor phenotypes similar to those seen in *sov* mutants. Germariums were often filled with germline cells with dot spectrosomes and the

follicle cells encroached anteriorly. Similar results were also observed using other somatic drivers (*c587-GAL4* and *da-GAL4*). Interestingly, *nos>sov$^{RNAi}$* did not impaired germaria morphology (Fig 3E,F) and egg chambers representing all 14 morphological stages of oogenesis appeared phenotypically normal. However, eggs produced from *nos>sov$^{RNAi}$* mothers arrested during early embryogenesis, indicating that maternal *sov* is required for embryonic development. While the expression patterns and germline RNAi suggest that *sov* is deployed in both the soma and germline, Wayne et al. (1995) reported *sov* to be somatic line dependent. We conducted a germline clonal analysis and found instead that *sov*$^{+}$ is required in the germline (Fig S3). Additionally, *sov$^{RNAi}$* resulted in embryonic germline defects in a high-throughput study (Jankovics et al. 2014). Taken together, these data indicate that there are both somatic and germline requirements for *sov* in oogenesis.

To explore the differentiation of the germline further, we examined the distribution of Bam, which is expressed in germarium region 1 germ cells (D. M. McKearin and Spradling 1990; D. McKearin and Ohlstein 1995; Ohlstein and McKearin 1997). In the absence of Bam, germ cells hyper-proliferate, resulting in tumors composed of 2-cell cysts. Given the prevalence of undifferentiated tumorigenic germ cells upon somatic *sov$^{RNAi}$* knockdown, we asked if *sov* has a non-autonomous role in promoting Bam expression in germ cells. Indeed, somatic knockdown of *sov* results in significantly less Bam expressed in germ cells relative to controls (Fig 3G,H; Fig S4). Both *tj-GAL4>sov$^{RNAi}$* and *c587-GAL4>sov$^{RNAi}$* resulted in strong repression of Bam expression. Germline depletion of *sov* did not alter Bam localization (Fig S4). The reduction in Bam expression in the germline is consistent with a non-autonomous role of Sov function in the soma for differentiation signals directed to the germline.

185 To examine the defective follicle encapsulation of the germline cysts following *sov*
186 depletion (resulting in fused follicles), we used the oocyte-specific expression of Oo18 RNA
187 binding protein (Orb) (Lantz et al. 1994) to count the number of oocytes per cyst. Orb specifies
188 the future single oocyte at the posterior of the 16-germ cell egg chamber in control egg
189 chambers (Fig 3I, Fig S2D). When we examined ovaries where *sov* expression had been
190 knocked down in somatic cells, we found examples of both egg chambers with either too many
191 oocytes (Fig 3J) or no oocytes (Fig 3K, Fig S4). In a wild type 16-cell germline cyst, one of the
192 two cells that has four ring canals becomes the oocyte, and this feature can be used to
193 determine if extra germline divisions had occurred in cysts, or if multiple cysts were enveloped
194 by the follicle cells. We saw that the egg chambers with multiple Orb+ cells always had more
195 than 16 germ cells, and in the example shown, all three Orb+ cells had four ring canals,
196 indicating that egg chamber fusion had occurred (Fig 3J). Taken together, we conclude that
197 Sov functions in the soma to ensure the proper differentiation of the germline.
198
199 *Sov represses gene expression in the ovary*
200 To provide mechanistic insights into the function of Sov, we performed transcriptome profiling
201 using triplicated Poly-A$^+$ RNA-seq analyses of ovaries from *sov*, *sov*$^{RNAi}$, and control females.
202 The gene expression profiles of ovaries from sterile females were markedly different from
203 controls, primarily due to de-repression in the mutants (Fig 4A; Table S2). Among genes
204 showing differences in expression (FDR *p*adj < 0.05), we found 1,752 genes with >4-fold
205 increased expression in mutants, while there were only 172 genes with >4-fold decreased
206 expression (Table S3). To explore where these de-repressed genes are normally expressed,
207 we examined their expression in other female tissues and in testes in a set of quadruplicated

RNA-seq experiments (a lab resource for this and other studies; Fig 4B). Many of the genes suppressed by Sov in ovaries were highly expressed in other female tissues, or in the testis, indicating that wildtype Sov prevents ectopic gene expression.

To determine what types of genes are de-repressed in *sov* mutants, we performed Gene Ontology (GO)(Gene Ontology Consortium 2015) term analysis (Fig S5; Table S4). The genes with lower expression in *sov* mutants had only a few significant GO terms that were oogenic in nature (as expected given the general lack of mature eggs in *sov* mutants). For example, there was poor expression of the chorion genes that are required to build the eggshell (Orr-Weaver 1991), as well as genes required for follicle cell development. In contrast, the de-repressed genes in *sov* mutants showed that there was a significant enrichment of genes with cell signaling GO terms, including neuronal communication. We found that many genes repressed by Sov in the ovary are indeed expressed in the head. For example, the *heartless* (*htl*) locus, which encodes a FGFR tyrosine kinase receptor important for neuron/glia communication (Stork et al. 2014), was de-repressed in *sov* ovaries and highly expressed in the head (Fig 4C). This intriguing result suggests that *sov* normally functions to repress a host of signaling pathways. This de-repressed signaling in *sov* mutants is likely catastrophic for communication between various somatic cells and the germline during egg chamber development. This could well explain the variety of mutant phenotypes observed in *sov* mutants.

The general repressive role of *sov* was also revealed by a dramatic and coherent elevation of transposon expression in the mutants (Fig 4D; Table S3). 138 transposable element classes were detected in our gene expression profiling of *sov* mutants and controls. Of these, 91 had increased expression (FDR *p*adj < 0.05) >4-fold in *sov* mutants, while none

had >4-fold decreased expression. The DNA, LINE, and LTR families of transposable elements were all de-repressed in *sov* mutants. Interestingly, the somatic knockdown of *sov* resulted in the most dramatic de-repression of the *gypsy* and *copia* classes of transposons, which are *Drosophila* retroviruses that develop in somatic cells and are exported to the developing germline (Yoshioka et al. 1990). Thus, wild type *sov* may be important to protect the germline from infection. Equally interesting, germline knockdown resulted in greatest de-repression of the *HeT-A*, *TAHRE*, and *TART* transposons that compose the telomere (Mason, Frydrychova, and Biessmann 2008). Poorly formed telomeres result in loss of material from chromosome ends and telomere fusion, which results in poor anaphase separation and aneuploidy (Cenci, Ciapponi, and Gatti 2005). This finding suggests that the general role of *sov* in silencing transposon expression includes control of both transposon functions necessary for normal cellular metabolism, exemplified by the telomere transposons, as well as detrimental activities of retrovirus-like elements. In addition to widespread dis-regulation of signaling and development pathways in *sov* mutants, transposon dysgenesis may contribute to the severe defects in *sov* ovarian development.

*Sov is a Suppressor of PEV and co-localizes with HP1*

The repressive function of Sov is reminiscent of HP1a function as general repressor of gene expression. HP1a was first characterized in *Drosophila* as a suppressor of PEV, a process that reduces gene expression due to spreading of heterochromatin into a gene region (Clark and Elgin 1992). To test the hypothesis that *sov* negatively regulates gene expression by promoting heterochromatin formation, we examined the role of *sov* in PEV in the eye, where patches of white$^{+}$ (red pigmented) and white$^{-}$ (not red pigmented) eye facets are easily

observed (Fig 5A). If, like HP1a, Sov represses gene expression by promoting heterochromatin formation (Eissenberg et al. 1990), then it should suppress PEV, which would be seen as increased eye pigmentation. We obtained five different variegating $w^+$ transgene insertions associated with either the heterochromatic pericentric region of chromosome arm 2L or spread along the length of the heterochromatin-rich chromosome 4. In control animals, these insertions show a characteristic eye variegation pattern (Fig 5B). Consistent with the allelic strengths seen in previous experiments, the weak $sov^2$ allele did not suppress PEV (Fig 5C) but the stronger $sov^{ML150}$, $sov^{EA42}$, and *Df(1)sov* mutations dominantly suppressed PEV (Fig 5D–F). These data demonstrate a role for Sov in heterochromatin formation.

The strong repressive function of Sov is reminiscent of HP1. If Sov is in complex with HP1a (Alekseyenko et al. 2014), then Sov and HP1 should colocalize in cells. To see if this was true, we followed HP1a-RFP and GFP-Sov localization in 1–2 hr live embryos where both are abundantly expressed (Rudolph et al. 2007). During *Drosophila* cleavage, nuclei undergo rapid synchronous nuclear divisions prior to cellularization at cleavage division/nuclear cycle (NC) 14 (Foe and Alberts 1983). HP1a and Sov colocalized in all NC 10–12 embryos we examined (Fig 6A; N=7 embryos). During prophase, both HP1a and Sov were enriched in regions of condensed DNA (Fig 6A, 4:30). Whereas low levels of HP1a decorated DNA throughout division, Sov was depleted during mitosis (Fig 6A, 10:00 and 11:10). Upon re-entry into interphase, Sov localization to nuclei resumed and was coincident with HP1a. Formation of heterochromatin is contemporaneous with or slightly proceeds HP1a apical subnuclear localization in NC 14 (Rudolph et al. 2007; Yuan and O'Farrell 2016). At this stage, we observed a strong colocalization of HP1a and Sov. Measuring the distribution of HP1a and Sov (Fig 6B,B') confirmed high levels of colocalization within HP1a subnuclear domains (Fig 6C,

277 shaded region). These data support the idea that HP1a and Sov are deposited maternally
278 where they assemble into a complex.
279
280 **Discussion**
281 Sov is a novel heterochromatin-associated protein
282 Genomes contain large blocks of DNA of potentially mobile transposons and immobile mutated
283 derivatives (Vermaak and Malik 2009). The cell keeps these transposons from wreaking havoc
284 on the genome by actively suppressing their expression through condensation into
285 heterochromatin. However, some tightly regulated expression from heterochromatin is required
286 for normal cellular function. For example, the telomeres of *Drosophila* are maintained by
287 transposition of mobile elements from heterochromatic sites (Mason, Frydrychova, and
288 Biessmann 2008), and histone and rRNA genes are located within heterochromatic regions
289 (Yasuhara and Wakimoto 2006). Weakening of heterochromatin by suppressors of variegation
290 results in de-repression of gene expression at the edges of heterochromatin blocks,
291 suggesting that the boundaries between repressed and active chromatin can expand and
292 contract (Weiler and Wakimoto 1995; Reuter and Spierer 1992). HP1a, encoded by
293 *Su(var)*205*,* is a central component of heterochromatin (Ebert et al. 2006) that shows the same
294 strong suppression of variegation that we observed in *sov*/+ flies. Similarly, HP1a is also
295 required to repress the expression of transposons (Vermaak and Malik 2009).
296 Repression is often stable and emerging themes suggest a robust set of activities,
297 rather than a single component, maintain a chromatin state. For example, long-term repression
298 is stabilized with complexes, such as the repressive Polycomb Group (PcG) which provides an
299 epigenetic memory function (Kassis, Kennison, and Tamkun 2017). In order to create a stable

epigenetic state, the PcG complexes have subunits that modify histones and subunits that bind these modifications. In addition to being tethered to the chromatin via histones, the PcG complexes also include DNA-binding proteins, such as the YY1-like ZnF protein Pleiohomeotic (Pho) that further reinforce localization. The DNA anchor proteins in PcG complexes are at least partially redundant with the histone-binding components (J. L. Brown et al. 2003), suggesting that localization is robust due to multiple independent localization mechanisms.

In mammals, HP1a also has both histone and DNA anchoring. DNA anchors are provided by a family of KRAB ZnF proteins that have undergone a massive radiation during evolution (Yang, Wang, and Macfarlan 2017; Ecco, Imbeault, and Trono 2017). KRAB proteins are essential for repression of transposons and more generally for a properly regulated genome. The KRAB family members use a KRAB-Associated Protein (KAP1) adapter protein to associate with HP1a and the SETDB1 methylase that modifies histones to enable HP1a binding (Fig 7). Although KRAB proteins have not been identified in *Drosophila*, the *Drosophila bonus* (*bon*) gene encodes a KAP1 homolog (Beckstead et al. 2005). *Drosophila* also has a SETDB1 encoded by *egg*, and loss of *egg* results in ovarian phenotypes reminiscent of those observed in *sov* loss-of-function females (Clough et al. 2007; Clough, Tedeschi, and Hazelrigg 2014; X. Wang et al. 2011). Like *sov* and HP1a, *bon* is a modifier of variegation, raising the possibility that they collectively coordinate gene regulation. In support of this hypothesis, Bon, Sov, and Egg were all identified in the same biochemical complex as HP1a (Alekseyenko et al. 2014). It is tempting to speculate that the single very long and ZnF-rich Sov protein plays the same direct tethering role as the large family of KRAB ZnF proteins in mammals (Fig 7). If Sov uses subsets of fingers to bind DNA, it could localize to many different sequences in a combinatorial fashion. Additionally, the complex containing Sov and HP1a also contains RNA

(Alekseyenko et al. 2014) and could act as a tether via a series of RNA intermediates as occurs in sex chromosome inactivation, for example (J. T. Lee 2009). As in the case of Pho in the PcG complexes, Sov might be a robustness factor rather than an absolute requirement, as we did not observe gross delocalization of HP1a following $sov^{RNAi}$ (not shown). Further work will be required to fully understand the relationship between Sov and HP1a.

Repression by Sov promotes germline differentiation

The fact that there have been many female sterile alleles of *sov* isolated suggests that the ovary is particularly sensitive to reduced *sov* activity. The female sterility phenotype is complex and somewhat variable, but partial *sov* loss of function is characterized by somatic-dependent differentiation defects. Phenotypes include defective follicle cell encasement of egg chambers, resulting in follicles with multiple oocytes, and accumulation of germline tumor cells that undergo complete cytokinesis instead of forming mature and differentiated 16-cell egg chambers. Stronger alleles also show a germline-dependent block in development and lethality. Diverse functions sometimes result from isoform and/or localization diversity. While *sov* functions seem diverse, the locus expresses a single major isoform and Sov seems to be consistently located in the nucleus. It is more likely that one mechanism account for all *sov* function. We suggest that Sov in regulates the chromatin landscape that helps repress stem cell identity and promote differentiation. This same mechanism represses ectopic signaling and tissue-inappropriate responses in the ovary, controls heterochromatin formation in the eye and represses transposons.

We observed dramatic de-repression of transposons in *sov* mutants. In addition to our observations, Czech et al (2013) found evidence for a role of *sov* in transposon repression in a

346 genome-wide screen. Transposon repression raises the possibility that Sov interacts with the
347 Piwi pathway (Brennecke et al. 2007; Yin and Lin 2007; Teixeira et al. 2017) which directly
348 interacts with HP1a to facilitate heterochromatin formation and promote transposon silencing
349 (Brower-Toland et al. 2007; S. H. Wang and Elgin 2011). However, genes involved more
350 strictly with transposable element regulation, such as Piwi, do not have strong dosage effects
351 on PEV (Gu and Elgin 2013), while *sov* and *HP1a*/*Su(var)205* do. This distinction suggests
352 that Sov has a more general effect on heterochromatin, rather than specificity for transposon
353 repression. Nevertheless, interplay between Piwi, Sov, and HP1a remains possible.
354
355 Conclusions
356 Our data show that *sov* encodes a repressor of gene expression and transposons in the
357 ovary. In the absence of *sov*, genes that are normally expressed at very low levels in ovary,
358 are activated. These same genes show dynamic patterns of expression in other adult tissues,
359 suggesting that they are de-repressed in the absence of Sov. The failure to restrict gene
360 expression results in a range of phenotypes including ovarian tumors, defective oogenesis,
361 and tissue degeneration. These data support the idea that Sov encodes a novel and essential
362 protein that is generally repressive, possibly via interactions with HP1a.
363
364 **Materials and Methods**
365 We have adopted the FlyBase-recommended resources table which includes all genetic,
366 biological, cell biology, genomics, manufactured reagents and algorithmic resources used in
367 this study (Table S5).
368

## Flies and genetics

The *small ovaries* (*sov*) locus was defined by three non-complementing female sterile mutations including $sov^{2}$, which mapped to ~19 cM on the X chromosome in the 5D5-6;6E4-5 interval (D. Mohler and Carroll 1984; J. D. Mohler 1977), refined to 6BD (Wayne et al. 1995). Cook *et al.* (2012) placed *sov* in the 4-gene *CG14438–shf* interval (Fig. 1; Table S5). *sov* mutations complemented $shf^{2}$, but not $P\{SUPor\text{-}P\}sov^{KG00226}$ and $P\{GawB\}sov^{NP6070}$, suggesting that *CG14438* or *CR43496* were *sov*, which we confirmed by generating *Df(1)sov* (X:6756569..6756668; 6770708, 6C12;6C13) as a FLP-induced recombination event between $P\{XP\}sov^{d07849}$ and *PBac{RB}e03842* (Parks et al. 2004; Cook et al. 2012) to remove only *CG14438* and *CR43496.*

We used the dominant female sterile technique for germline clones (Chou and Perrimon 1996). Test chromosomes were free of linked lethal mutations by male viability ($sov^{2}$) or rescue by *Dp(1;3)DC486* ($sov^{EA42}$ and *Df(1)sov*), were recombined with $P\{ry^{+t7.2}=neoFRT\}19A$, and verified by complementation tests and PCR. We confirmed $P\{ry^{+t7.2}=neoFRT\}19A$ functionality by crossing to $P\{w^{+mC}=GMR\text{-}hid\}SS1$, $y^{1}$ $w^{*}$ $P\{ry^{+t7.2}=neoFRT\}19A$; $P\{w^{+m*}=GAL4\text{-}ey.H\}SS5$, $P\{w^{+mC}=UAS\text{-}FLP.D\}JD2$ and scoring for large eye size. We crossed females with FRT chromosomes to $P\{w^{+mC}=ovoD1\text{-}18\}P4.1$, $P\{ry^{+t7.2}=hsFLP\}12$, $y^{1}$ $w^{1118}$ $sn^{3}$ $P\{ry^{+t7.2}=neoFRT\}19A/Y$ males for 24 hours of egg laying at 25°C. We heat-shocked for 1hr at 37°C on days 2 and 3. We dissected females (5d post-eclosion) to score for $ovo^{D1}$ or wildtype morphology.

We generated *PBac{GFP-sov}* from P[acman] BAC clone CH322-191E24 (X:6753282--6773405) (Venken et al. 2009) grown in the SW102 strain (Warming et al. 2005). In step one, we integrated the positive/negative marker CP6-RpsL/Kan (CP6 promoter with a bi-cistronic

392 cassette encoding the RpsL followed by the Kan) between the first two codons of *sov*, and
393 selected (15 µg/ml Kanamycin). We integrated at the galK operon in DH10B bacteria using
394 mini-lambda-mediated recombineering (Court et al. 2003). We amplified DH10B::CP6-
395 RpsL/Kan DNA using primers N-CG14438-CP6-RN-F and -R. Correct events were identified
396 by PCR, as well as resistance (15 µg/ml Kanamycin) and sensitivity (250 µg/ml Kan). In step
397 two, we replaced the selection markers with a multi-tag sequence (Venken et al. 2011),
398 tailored for N-terminal tagging (N-tag) and counter-selected (250 µg/ml Kanamycin). The N-tag
399 (3xFlag tag, TEV protease site, StrepII tag, superfolder GFP, FlAsH tetracysteine tag, and
400 flexible 4xGlyGlySer (GGS) linker) was *Drosophila* codon optimized in a R6Kγ plasmid. We
401 transformed plasmid into EPI300 for copy number amplification. We confirmed correct events
402 by PCR and Sanger DNA sequencing. Tagged P[acman] BAC clone DNA was injected into $y^1$
403 *M{vas-int.Dm}ZH-2A w*; PBac{y+-attP-3B}*$^{VK00033}$ embryos, resulting in $w^{1118}$*; PBac{y[+mDint2]*
404 $w^{+mC}$*=FTSF.GFP-sov}*$^{VK00033}$.
405
406 Microscopy
407 We fixed ovaries in 4 or 5% EM-grade paraformaldehyde in PBS containing 0.1 or 0.3%
408 Triton X-100 (PBTX) for 10-15min, washed 3x 15 min in PBTX, and blocked >30min in 2%
409 normal goat serum, 0.5-1% bovine serum albumin (BSA) in PBS with 0.1% Tween-20 or 0.1%
410 Triton X-100. Antibodies and DAPI were diluted into blocking buffer. We incubated in primary
411 antibodies overnight at 4 °C and secondaries 2–3hr at room temperature. Embryos (1–2 hr)
412 were prepared for live imaging in halocarbon oil (Lerit et al. 2015). We imaged ovaries and
413 embryos using a Nikon Ti-E system or Zeiss LSM 780 microscope and eyes with a Nikon SMZ.
414

DNA-Seq

Genomic DNA was extracted from 30 whole flies per genotype (Huang, Rehm, and Rubin 2009)(Sambrook and Russell 2006) to prepare DNA-seq libraries (Nextera DNA Library Preparation Kit). We used 50 bp, single-end sequencing (Illumina HiSeq 2500, CASAVA base calling). Sequence data are available at the SRA (SRP14438). We mapped DNAseq reads to FlyBase r6.16 genome with Hisat2 (-k 1 --no-spliced-alignment)(Kim, Langmead, and Salzberg 2015). We used mpileup and bcftools commands from SAMtools within the genomic region X:6756000–6771000 (H. Li et al. 2009; H. Li 2011) for variant calling and snpEFF to determine the nature of variants in *sov* mutants (Cingolani et al. 2012).

RNA-Seq

Stranded PolyA+ RNA-seq libraries from *sov* and control ovaries (Table S5) were created (H. Lee, Cho, et al. 2016) and are available at GEO (GSE113977). We extracted total RNA (Qiagen RNeasy Mini Kit) in biological triplicate from 15 ovaries (4-5d post-eclosion) and used 200 ng with 10 pg of ERCC spike-in control RNAs (pools 78A or 78B) for libraries (Jiang et al. 2011; Zook et al. 2012; Pine et al. 2016; H. Lee, Pine, et al. 2016). We used 50 bp, single-end sequencing as above. Tissue expression analysis are from GEO accession GSE99574, a resource for comparing gene expression patterns.

We mapped RNA-seq reads to FlyBase r6.21 with Hisat2 (-k 1 --rna-strandness R --dta)(Kim, Langmead, and Salzberg 2015). We determined read counts for each attribute of the FlyBase r6.21 GTF file (with ERCC and transposable element sequences), with HTSeq-count (Anders, Pyl, and Huber 2015). Transposon sequences were from the UCSC Genome Browser RepeatMasker track (Casper et al. 2018; Smit, Hubley, and Green 2013-2015).

438 We conducted differential expression analysis with DESeq2 (pAdjustMethod = "fdr")
439 (Love, Huber, and Anders 2014). We removed genes with read counts ≤1 and read counts for
440 transposable elements with >1 location were summed for the DESeq2 analysis. *Df(1)sov/sov*$^{2}$
441 replicate 3 and *sov*$^{2}$*/w*$^{1118}$ replicate 1 were failed. For *sov* mutant *vs.* control DESeq2 analysis,
442 all *sov* mutants were compared to all wild type controls. For tissue types, we compared each
443 sexed tissue to sexed whole organism. We used reads per kb per million reads (RPKM) for
444 gene-level expression.
445 For heatmaps, we calculated Euclidean distance, performed hierarchical cluster
446 analysis (agglomeration method = Ward), and mean-subtracted scaled across genotypes.
447 Mean sample RPKM correlation values (Table S1) were calculated by cor.test() function with
448 Pearson correlation coefficient in R (R Core Team 2017). We represented de-repressed genes
449 as mean-subtracted scaled values across tissues in the heatmap (Table S6).
450 For read density tracks, replicate raw read files were combined. Bedgraph files were
451 created with bedtools genomecov (Quinlan and Hall 2010) visualized on the UCSC genome
452 browser (Kent et al. 2002). Tracks were scaled by the number of reads divided by total reads
453 per million.
454 We used ClueGO (Bindea et al. 2009), with Cytoscape (Shannon et al. 2003) for one-
455 sided enrichment analysis (see Table S4).
456 Images were assembled using ImageJ and Photoshop. We used ROI tool in ImageJ,
457 plotted/analyzed image data (Microsoft Excel and GraphPad Prism), and calculated
458 significance by D'Agnostino and Pearson normality tests, followed Student's two-tailed t-test or
459 Mann-Whitney tests.
460

461 **Acknowledgements**

462 We thank Astrid Haase, Nasser Rusan, the Oliver and Lerit labs, and Dawson Mohler
463 (deceased) for stimulating discussions. We used resources from FlyBase, The Hybridoma
464 Bank, and the NIH biowulf (Table S5). This research was supported in part by the Intramural
465 Research Program of the NIH, NIDDK to BO; NIH Grant 5K22HL126922-02 to DAL; NIH
466 Graduate Partners Program to LB; Baylor College of Medicine, the Albert and Margaret Alkek
467 Foundation, the McNair Medical Institute at The Robert and Janice McNair Foundation, the
468 Cancer Prevention and Research Institute of Texas (R1313), and NIH grant 1R21HG006726 to
469 KJTV; and NIH grants P40OD018537 and R24RR014106, and the National Science
470 Foundation DBI-9816125 to KRC.

471

472 **Author contributions**

473 LB, KC, BO, and DAL designed the project and analyzed data. LB, KC, CW, and KJTV
474 performed genetics. LB and HY performed genomics. LB, EAC, CW and DAL performed
475 imaging. LB, BO, and DAL wrote the manuscript. All authors reviewed data and provided
476 feedback on the manuscript.

477

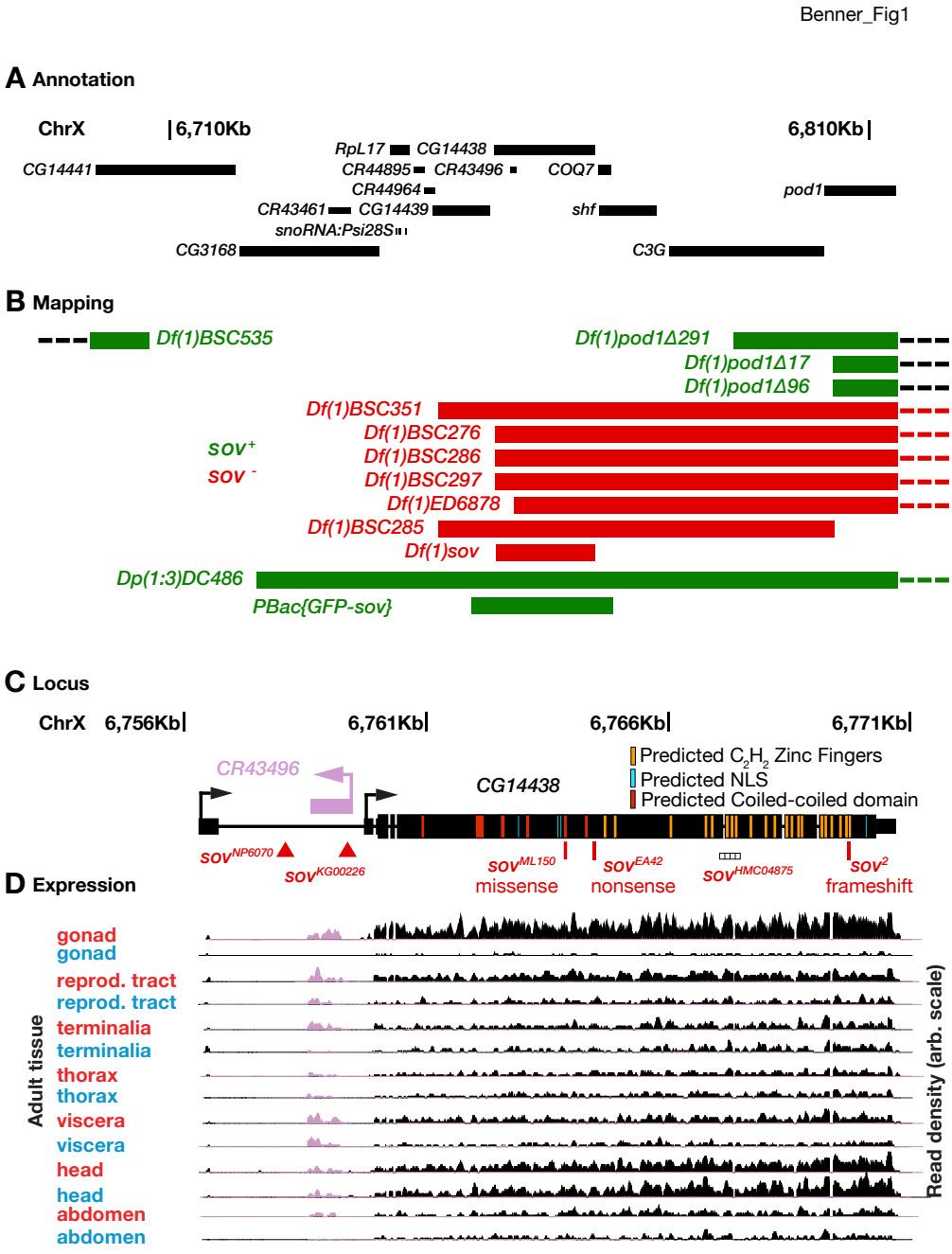



**Figure 1. *sov* is *CG14438*.** (A) Genes of the genomic interval X:6710000–6810000 (Gramates et al. 2017). (B) Deficiency (*Df*) and Duplication (*Dp*) mapping. (C) Schematic of the *CG14438* (black) and CR43496 (purple) genes. Transcription start (bent arrows), introns (thin lines) non-coding regions (medium lines) and coding regions (thick lines) are shown. Transposon insertions (triangles), point mutations (red lines), the region targeted by the shRNAi transgene (base-paired), and Sov protein features are shown. Non-complementing (*sov*$^{-}$, red) and complementing (*sov*$^{+}$, green) alleles and rearrangements are shown. (D) RNA expression tracks by tissue type from female (red) or male (blue) adults.

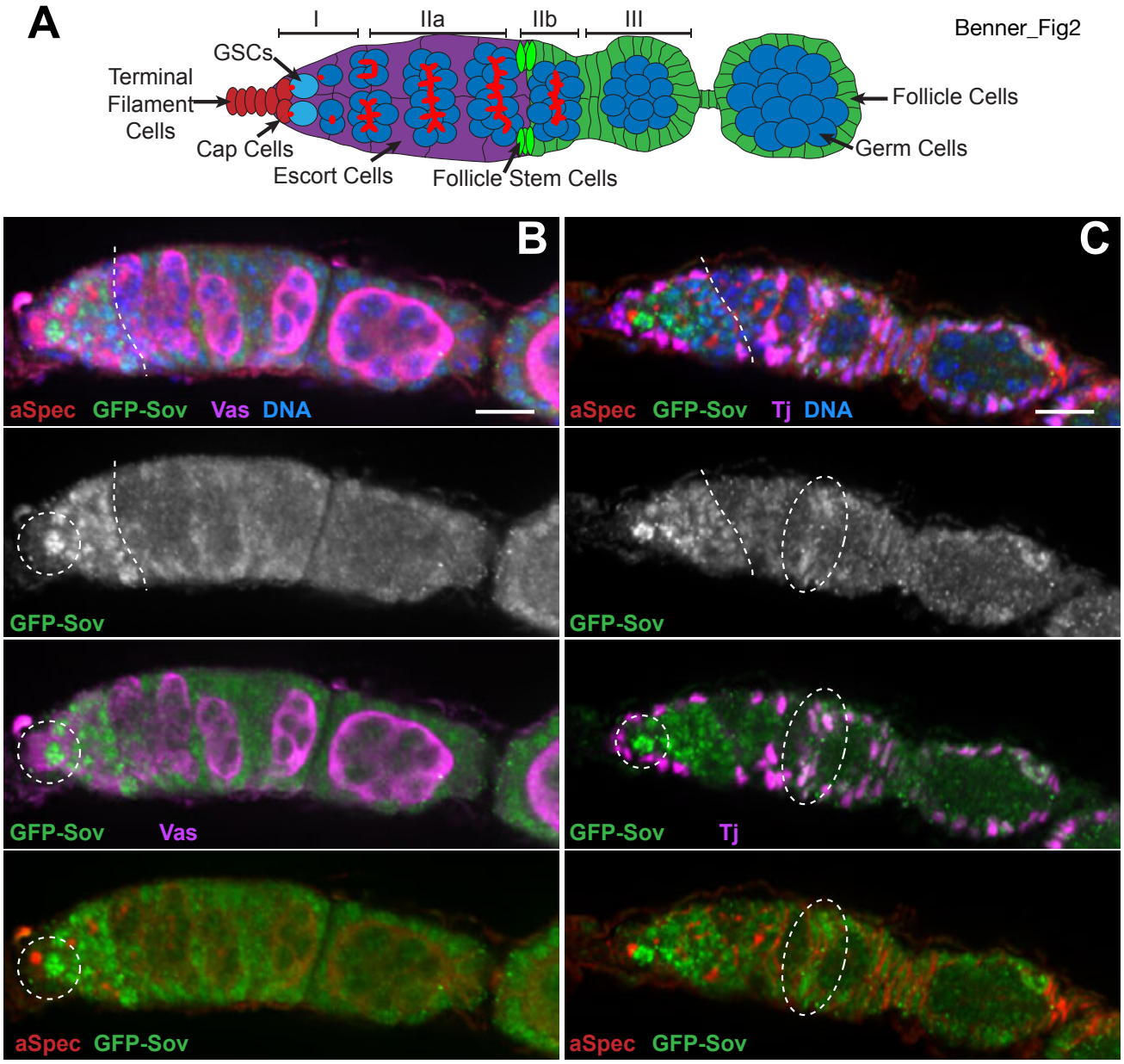



**Figure 2. Sov germarium expression.** (A) Cartoon showing germarium regions I-III and young egg chamber with cell types. (B,C) 1d post-eclosion, visualized for GFP-Sov (green), anti-αSpectrin (red). Images show single optical sections. (B) Sov contrasted with anti-Vas germline (magenta), or (C) -Tj somatic staining (magenta). GFP-Sov expression regions of interest (see text) are shown (dashed lines). Bars: 10μm.

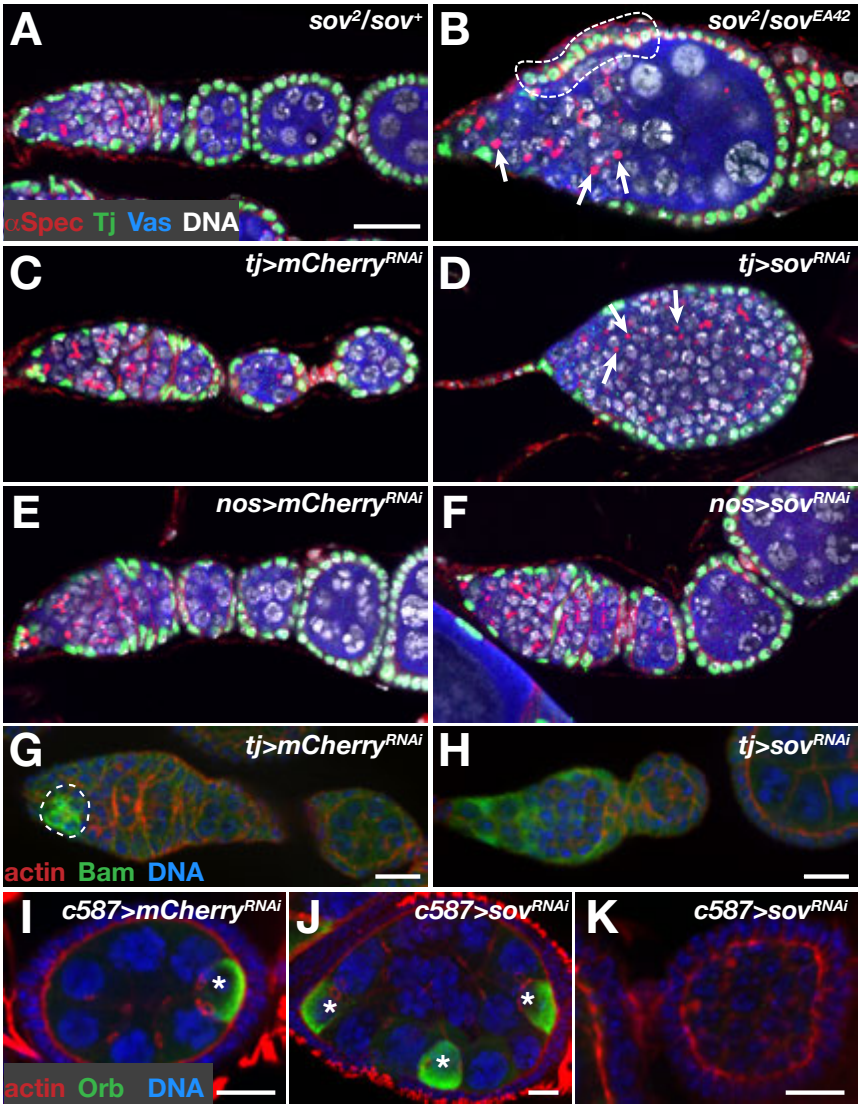
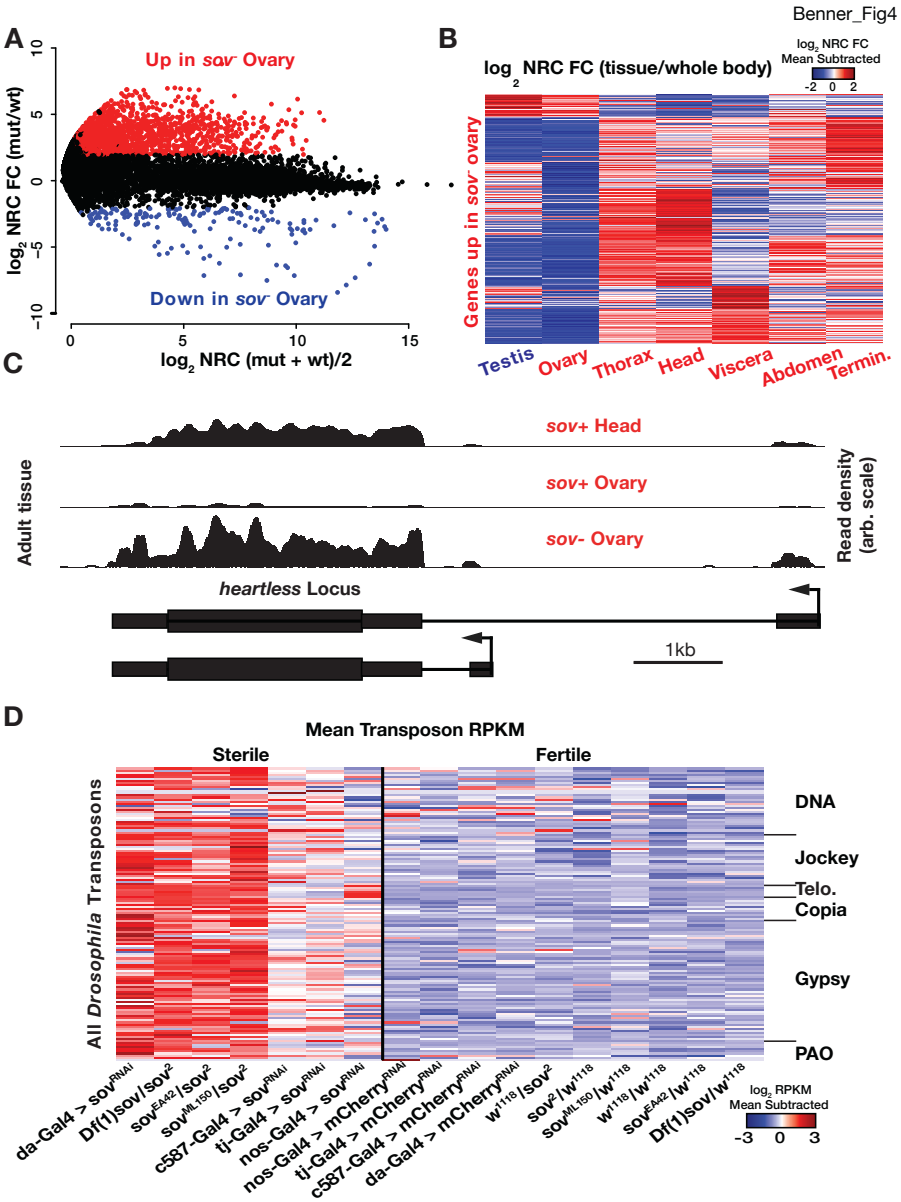



**Figure 3. *sov* mutant defects.** (A-K) images of indicated genotype (white). (A-F) Germaria (4-5d post-eclosion) of stained with anti-Vas (blue), -Tj (green), -αSpectrin (red), and DAPI (white). Dot spectrosomes (arrows) and displaced follicle cells (dashed lines) are shown. Images are single optical sections. (G, H) Germaria (1d post-eclosion) stained with anti-Bam (green), -actin (red), and DAPI (blue). Bam region of interest is shown (dashed lines). (I–K) Germaria (1d post-eclosion) anti-Orb (green), -actin (red), and DAPI (blue). Orb+ cells shown (*). Single optical sections (A-H) and maximum intensity projections (I-K). Bars: (A–F) 20μm, (G–K) 10μm.




505 **Figure 4. *sov* mutant transcriptome.** (A) Relative expression in *sov* mutant vs *sov*+ ovaries
506 plotted against the mean expression in both sample types. Units are $Log_2$ normalized read
507 counts (NRC). Data points are genes, those with >4-fold change ($log_2$ 2, red) or <–4 ($log_2$ -2,
508 blue) and FDR *p*adj value <0.05 are highlighted. (B) Tissue-biased expression in wildtype for
509 genes de-repressed in *sov* mutant ovaries. Heatmap from mean-subtracted ratios scaled
510 across tissues (red=higher; blue=lower). (C) The *heartless* locus (see Fig1 for format) showing
511 RNA-seq normalized read densities of *sov*+ from head and *sov*$^{+}$ or *sov*$^{-}$ ovarian tissues. (D)
512 Transposable element expression in *sov* mutants and control alleles (as indicated). Heatmap
513 from mean-subtracted reads (in RPKM. red=higher; blue=lower) scaled for each transposable
514 element (rows) across genotypes (columns). Some element subtypes are shown for DNA, non-
515 LTR (Jockey), telomeric repeat (Telo), and LTR (Gypsy, Copia, and PAO) are indicated.
516

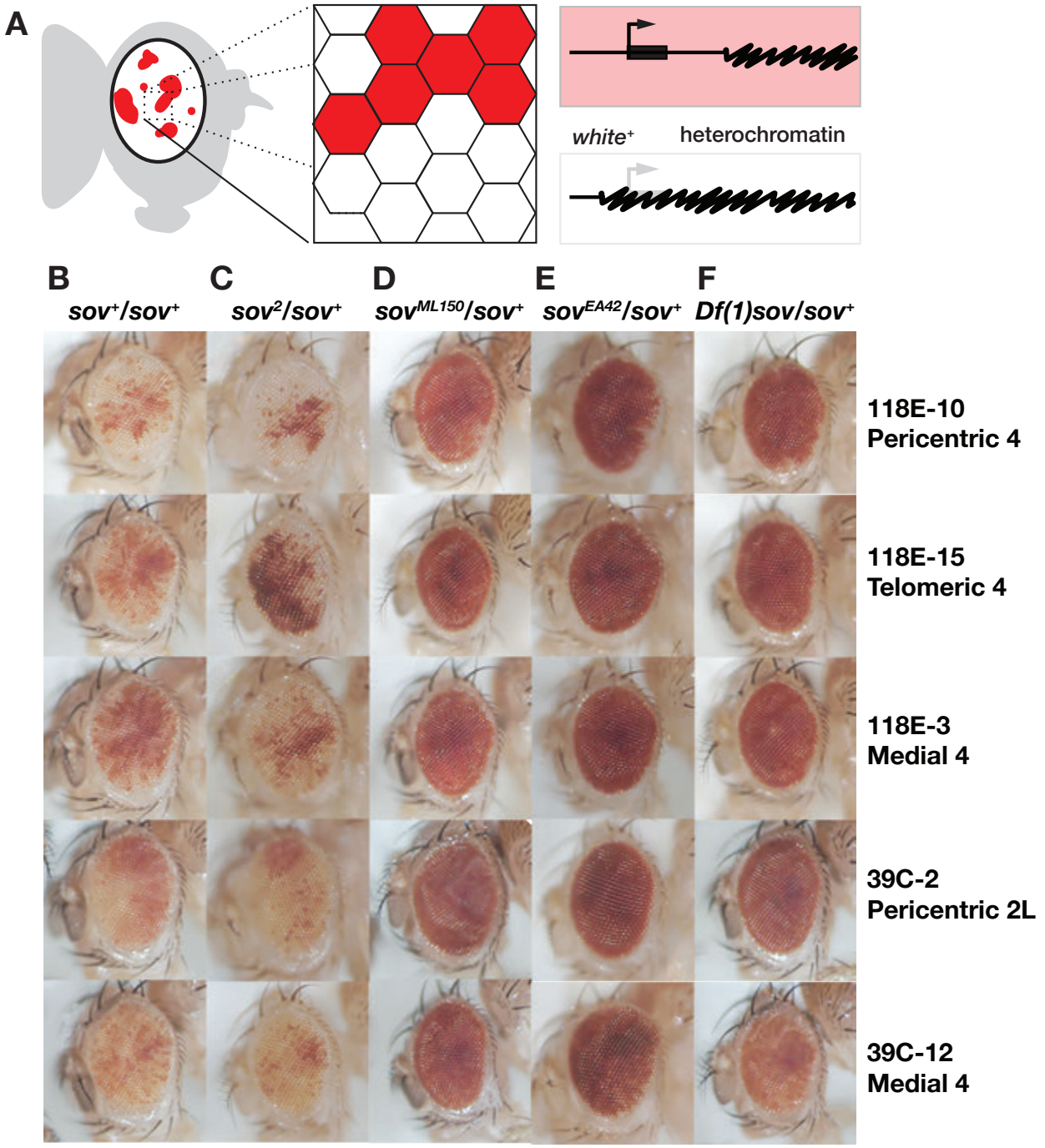



**Figure 5. Sov is a dominant suppressor of position effect variegation.** (A) Cartoon of position effect variegation (PEV) in the eye. Expression of the *white* gene (bent arrow, thick bar, red) near heterochromatin (squiggled) region can be silenced (white) by spreading. (B–F) Eyes with variegated expression of *P{hsp26-pt-T}* transgene inserted into different chromosomal positions (rows) in different *sov* backgrounds (columns).

Benner_Fig6

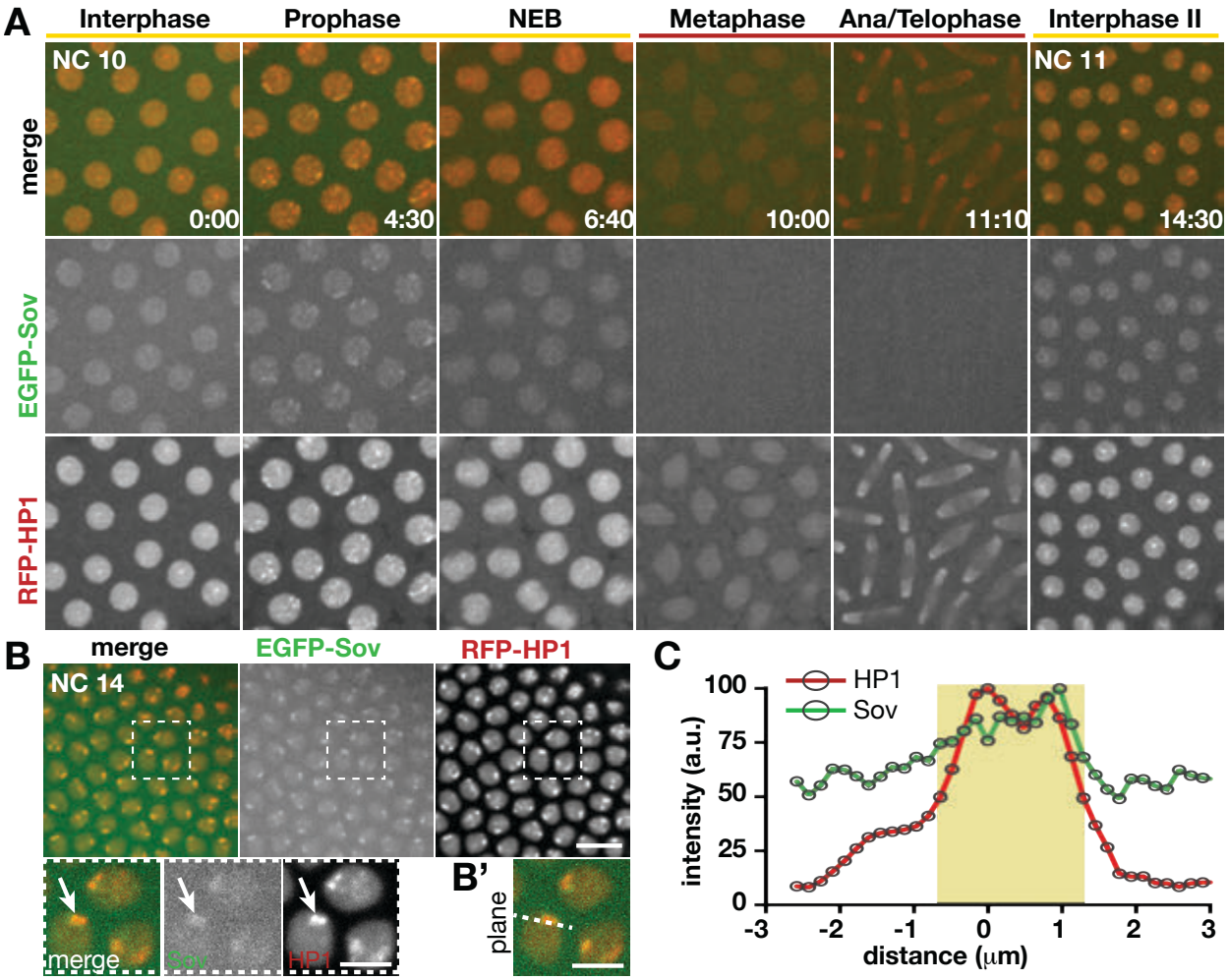



**Figure 6. Sov colocalizes with HP1a.** Stills from live imaging of embryos expressing GFP-Sov and RFP-HP1a. (A) Localization of GFP-Sov and RFP-HP1a (rows) in stills from live imaging of an embryo expressing progressing from NC 10 (left) to NC 11 (right) with stages (columns, nuclear envelope breakdown=NEB) and time (min:s in merge row) shown. (B) expressing GFP-Sov and RFP-HP1a (columns) at NC 14. Boxed regions are magnified in insets below. An HP1a subnuclear domain is shown (arrows). (B') Single optical section containing peak HP1a fluorescence of the same region shown in (B). Dashed line indicates region used for linescan analysis. (A,B) Maximum projections through 1.5 μm volume at 1F/30s. Bars: 10 μm; insets, 5 μm. (C) Histogram of HP1a and Sov fluorescence intensity measured in (B'). Fluorescence levels (arbitrary units) normalized to the peak fluorescence intensity for each channel and the distance (μm) to peak HP1a signal. Line scans were scaled –3 to 3μm, where 0.0 is peak HP1a fluorescence. Half maximum HP1a fluorescence is shown (yellow shade).

Benner_Fig7

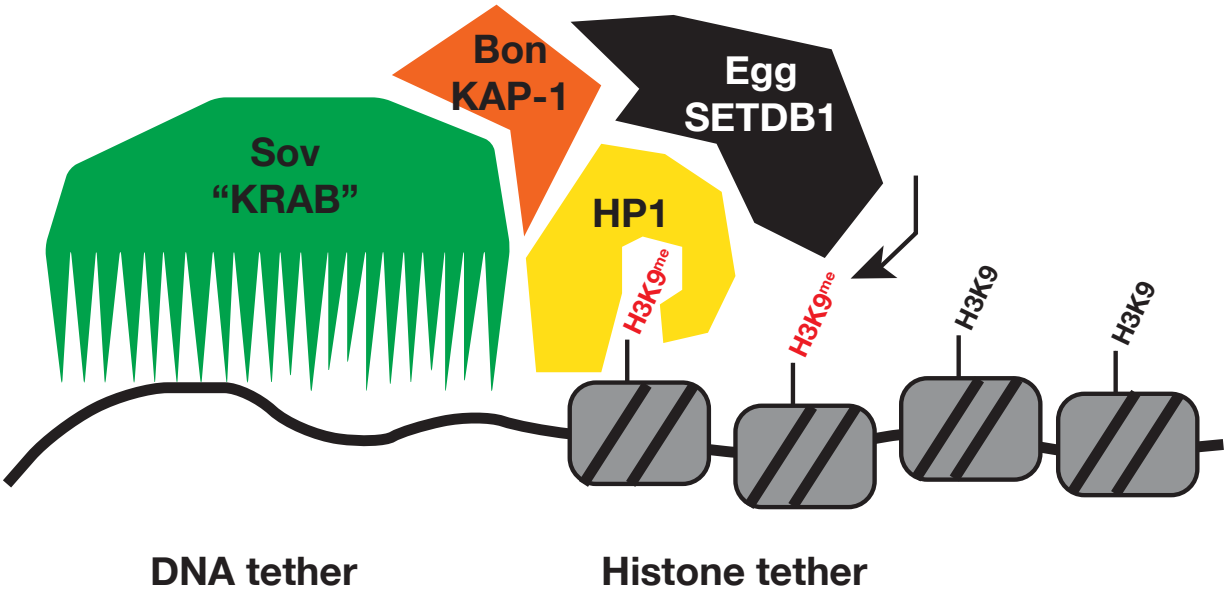



**Figure 7. Working model of *sov* function.** Sov protein function and attributes are analogous to the mammalian KRAB proteins (green). Like KRAB proteins, Sov is in complex with HP1a (yellow), Bon (mammalian KAP-1, orange) and Egg (mammalian SETDB1, black) providing a DNA tether that functions along with the H3K9me tether written by Egg (arrow, red lettering) and bound by HP1, to stabilize a repressed state.

## References

Alekseyenko, Artyom A., Andrey A. Gorchakov, Barry M. Zee, Stephen M. Fuchs, Peter V. Kharchenko, and Mitzi I. Kuroda. 2014. "Heterochromatin-Associated Interactions of Drosophila HP1a with dADD1, HIPP1, and Repetitive RNAs." *Genes & Development* 28 (13): 1445–60.

Anders, Simon, Paul Theodor Pyl, and Wolfgang Huber. 2015. "HTSeq--a Python Framework to Work with High-Throughput Sequencing Data." *Bioinformatics* 31 (2): 166–69.

Barton, Lacy J., Kaylee E. Lovander, Belinda S. Pinto, and Pamela K. Geyer. 2016. "Drosophila Male and Female Germline Stem Cell Niches Require the Nuclear Lamina Protein Otefin." *Developmental Biology* 415 (1): 75–86.

Bausek, Nina. 2013. "JAK-STAT Signaling in Stem Cells and Their Niches in Drosophila." *JAK-STAT* 2 (3): e25686.

Beckstead, R. B., S. S. Ner, K. G. Hales, T. A. Grigliatti, B. S. Baker, and H. J. Bellen. 2005. "Bonus, a Drosophila TIF1 Homolog, Is a Chromatin-Associated Protein That Acts as a Modifier of Position-Effect Variegation." *Genetics* 169 (2): 783–94.

Bindea, Gabriela, Bernhard Mlecnik, Hubert Hackl, Pornpimol Charoentong, Marie Tosolini, Amos Kirilovsky, Wolf-Herman Fridman, Franck Pagès, Zlatko Trajanoski, and Jérôme Galon. 2009. "ClueGO: A Cytoscape Plug-in to Decipher Functionally Grouped Gene Ontology and Pathway Annotation Networks." *Bioinformatics* 25 (8): 1091–93.

Börner, Kenneth, Dhawal Jain, Paula Vazquez-Pianzola, Sandra Vengadasalam, Natascha Steffen, Dmitry V. Fyodorov, Pavel Tomancak, Alexander Konev, Beat Suter, and Peter B. Becker. 2016. "A Role for Tuned Levels of Nucleosome Remodeler Subunit ACF1 during Drosophila Oogenesis." *Developmental Biology* 411 (2): 217–30.

Branco, Joana, Ismael Al-Ramahi, Lubna Ukani, Alma M. Pérez, Pedro Fernandez-Funez, Diego Rincón-Limas, and Juan Botas. 2008. "Comparative Analysis of Genetic Modifiers in Drosophila Points to Common and Distinct Mechanisms of Pathogenesis among Polyglutamine Diseases."

*Human Molecular Genetics* 17 (3): 376–90.

Brennecke, Julius, Alexei A. Aravin, Alexander Stark, Monica Dus, Manolis Kellis, Ravi Sachidanandam, and Gregory J. Hannon. 2007. "Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila." *Cell* 128 (6): 1089–1103.

Brower-Toland, Brent, Seth D. Findley, Ling Jiang, Li Liu, Hang Yin, Monica Dus, Pei Zhou, Sarah C. R. Elgin, and Haifan Lin. 2007. "Drosophila PIWI Associates with Chromatin and Interacts Directly with HP1a." *Genes & Development* 21 (18): 2300–2311.

Brown, James B., Nathan Boley, Robert Eisman, Gemma E. May, Marcus H. Stoiber, Michael O. Duff, Ben W. Booth, et al. 2014. "Diversity and Dynamics of the Drosophila Transcriptome." *Nature* 512 (7515): 393–99.

Brown, J. Lesley, Cornelia Fritsch, Jürg Mueller, and Judith A. Kassis. 2003. "The Drosophila Pho-like Gene Encodes a YY1-Related DNA Binding Protein That Is Redundant with Pleiohomeotic in Homeotic Gene Silencing." *Development* 130 (2): 285–94.

Casper, Jonathan, Ann S. Zweig, Chris Villarreal, Cath Tyner, Matthew L. Speir, Kate R. Rosenbloom, Brian J. Raney, et al. 2018. "The UCSC Genome Browser Database: 2018 Update." *Nucleic Acids Research* 46 (D1): D762–69.

Cenci, Giovanni, Laura Ciapponi, and Maurizio Gatti. 2005. "The Mechanism of Telomere Protection: A Comparison between Drosophila and Humans." *Chromosoma* 114 (3): 135–45.

Charlesworth, B., and C. H. Langley. 1989. "The Population Genetics of Drosophila Transposable Elements." *Annual Review of Genetics* 23: 251–87.

Chen, Shuyi, Su Wang, and Ting Xie. 2011. "Restricting Self-Renewal Signals within the Stem Cell Niche: Multiple Levels of Control." *Current Opinion in Genetics & Development* 21 (6): 684–89.

Chen, Taiping, and Sharon Y. R. Dent. 2014. "Chromatin Modifiers and Remodellers: Regulators of Cellular Differentiation." *Nature Reviews. Genetics* 15 (2): 93–106.

Chou, T. B., and N. Perrimon. 1996. "The Autosomal FLP-DFS Technique for Generating Germline Mosaics in Drosophila Melanogaster." *Genetics* 144 (4): 1673–79.

598 Cingolani, Pablo, Adrian Platts, Le Lily Wang, Melissa Coon, Tung Nguyen, Luan Wang, Susan J.
599 Land, Xiangyi Lu, and Douglas M. Ruden. 2012. “A Program for Annotating and Predicting the
600 Effects of Single Nucleotide Polymorphisms, SnpEff: SNPs in the Genome of Drosophila
601 Melanogaster Strain w1118; Iso-2; Iso-3.” *Fly* 6 (2): 80–92.
602 Clark, R. F., and S. C. Elgin. 1992. “Heterochromatin Protein 1, a Known Suppressor of Position-Effect
603 Variegation, Is Highly Conserved in Drosophila.” *Nucleic Acids Research* 20 (22): 6067–74.
604 Clough, Emily, Woongjoon Moon, Shengxian Wang, Kathleen Smith, and Tulle Hazelrigg. 2007.
605 “Histone Methylation Is Required for Oogenesis in Drosophila.” *Development* 134 (1): 157–65.
606 Clough, Emily, Thomas Tedeschi, and Tulle Hazelrigg. 2014. “Epigenetic Regulation of Oogenesis and
607 Germ Stem Cell Maintenance by the Drosophila Histone Methyltransferase Eggless/dSetDB1.”
608 *Developmental Biology* 388 (2): 181–91.
609 Cook, R. Kimberley, Stacey J. Christensen, Jennifer A. Deal, Rachel A. Coburn, Megan E. Deal, Jill M.
610 Gresens, Thomas C. Kaufman, and Kevin R. Cook. 2012. “The Generation of Chromosomal
611 Deletions to Provide Extensive Coverage and Subdivision of the Drosophila Melanogaster
612 Genome.” *Genome Biology* 13 (3): R21.
613 Court, Donald L., Srividya Swaminathan, Daiguan Yu, Helen Wilson, Teresa Baker, Mikail Bubunenko,
614 James Sawitzke, and Shyam K. Sharan. 2003. “Mini-Lambda: A Tractable System for
615 Chromosome and BAC Engineering.” *Gene* 315 (October): 63–69.
616 Di Stefano, Luisa, Jun-Yuan Ji, Nam-Sung Moon, Anabel Herr, and Nicholas Dyson. 2007. “Mutation of
617 Drosophila Lsd1 Disrupts H3-K4 Methylation, Resulting in Tissue-Specific Defects during
618 Development.” *Current Biology: CB* 17 (9): 808–12.
619 Ebert, Anja, Sandro Lein, Gunnar Schotta, and Gunter Reuter. 2006. “Histone Modification and the
620 Control of Heterochromatic Gene Silencing in Drosophila.” *Chromosome Research: An*
621 *International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome*
622 *Biology* 14 (4): 377–92.
623 Ecco, Gabriela, Michael Imbeault, and Didier Trono. 2017. “KRAB Zinc Finger Proteins.” *Development*

144 (15): 2719–29.

Eissenberg, J. C., T. C. James, D. M. Foster-Hartnett, T. Hartnett, V. Ngan, and S. C. Elgin. 1990. "Mutation in a Heterochromatin-Specific Chromosomal Protein Is Associated with Suppression of Position-Effect Variegation in Drosophila Melanogaster." *Proceedings of the National Academy of Sciences of the United States of America* 87 (24): 9923–27.

Elgin, Sarah C. R., and Gunter Reuter. 2013. "Position-Effect Variegation, Heterochromatin Formation, and Gene Silencing in Drosophila." *Cold Spring Harbor Perspectives in Biology* 5 (8): a017780.

Eliazer, Susan, Victor Palacios, Zhaohui Wang, Rahul K. Kollipara, Ralf Kittler, and Michael Buszczak. 2014. "Lsd1 Restricts the Number of Germline Stem Cells by Regulating Multiple Targets in Escort Cells." *PLoS Genetics* 10 (3): e1004200.

Eliazer, Susan, Nevine A. Shalaby, and Michael Buszczak. 2011. "Loss of Lysine-Specific Demethylase 1 Nonautonomously Causes Stem Cell Tumors in the Drosophila Ovary." *Proceedings of the National Academy of Sciences of the United States of America* 108 (17): 7064–69.

Foe, V. E., and B. M. Alberts. 1983. "Studies of Nuclear and Cytoplasmic Behaviour during the Five Mitotic Cycles That Precede Gastrulation in Drosophila Embryogenesis." *Journal of Cell Science* 61 (May): 31–70.

Fuller, Margaret T., and Allan C. Spradling. 2007. "Male and Female Drosophila Germline Stem Cells: Two Versions of Immortality." *Science* 316 (5823): 402–4.

Gene Ontology Consortium. 2015. "Gene Ontology Consortium: Going Forward." *Nucleic Acids Research* 43 (Database issue): D1049–56.

Gilboa, Lilach. 2015. "Organizing Stem Cell Units in the Drosophila Ovary." *Current Opinion in Genetics & Development* 32 (June): 31–36.

Goode, S., D. Wright, and A. P. Mahowald. 1992. "The Neurogenic Locus Brainiac Cooperates with the Drosophila EGF Receptor to Establish the Ovarian Follicle and to Determine Its Dorsal-Ventral Polarity." *Development* 116 (1): 177–92.

Gramates, L. Sian, Steven J. Marygold, Gilberto Dos Santos, Jose-Maria Urbano, Giulia Antonazzo,

Beverley B. Matthews, Alix J. Rey, et al. 2017. “FlyBase at 25: Looking to the Future.” *Nucleic Acids Research* 45 (D1): D663–71.

Graveley, Brenton R., Angela N. Brooks, Joseph W. Carlson, Michael O. Duff, Jane M. Landolin, Li Yang, Carlo G. Artieri, et al. 2011. “The Developmental Transcriptome of Drosophila Melanogaster.” *Nature* 471 (7339): 473–79.

Gu, Tingting, and Sarah C. R. Elgin. 2013. “Maternal Depletion of Piwi, a Component of the RNAi System, Impacts Heterochromatin Formation in Drosophila.” *PLoS Genetics* 9 (9): e1003780.

Hsieh, M., Y. Tintut, and J. D. Gralla. 1994. “Functional Roles for the Glutamines within the Glutamine-Rich Region of the Transcription Factor Sigma 54.” *The Journal of Biological Chemistry* 269 (1): 373–78.

Huang, Audrey M., E. Jay Rehm, and Gerald M. Rubin. 2009. “Quick Preparation of Genomic DNA from Drosophila.” *Cold Spring Harbor Protocols* 2009 (4): db.prot5198.

James, T. C., and S. C. Elgin. 1986. “Identification of a Nonhistone Chromosomal Protein Associated with Heterochromatin in Drosophila Melanogaster and Its Gene.” *Molecular and Cellular Biology* 6 (11): 3862–72.

Jankovics, Ferenc, László Henn, Ágnes Bujna, Péter Vilmos, Kerstin Spirohn, Michael Boutros, and Miklós Erdélyi. 2014. “Functional Analysis of the Drosophila Embryonic Germ Cell Transcriptome by RNA Interference.” *PloS One* 9 (6): e98579.

Jenuwein, T., and C. D. Allis. 2001. “Translating the Histone Code.” *Science* 293 (5532): 1074–80.

Jiang, Lichun, Felix Schlesinger, Carrie A. Davis, Yu Zhang, Renhua Li, Marc Salit, Thomas R. Gingeras, and Brian Oliver. 2011. “Synthetic Spike-in Standards for RNA-Seq Experiments.” *Genome Research* 21 (9): 1543–51.

Kassis, Judith A., James A. Kennison, and John W. Tamkun. 2017. “Polycomb and Trithorax Group Genes in Drosophila.” *Genetics* 206 (4): 1699–1725.

Kent, W. James, Charles W. Sugnet, Terrence S. Furey, Krishna M. Roskin, Tom H. Pringle, Alan M. Zahler, and David Haussler. 2002. “The Human Genome Browser at UCSC.” *Genome Research*

12 (6): 996–1006.

Kim, Daehwan, Ben Langmead, and Steven L. Salzberg. 2015. “HISAT: A Fast Spliced Aligner with Low Memory Requirements.” *Nature Methods* 12 (4): 357–60.

Kouzarides, Tony. 2007. “Chromatin Modifications and Their Function.” *Cell* 128 (4): 693–705.

Lantz, V., J. S. Chang, J. I. Horabin, D. Bopp, and P. Schedl. 1994. “The Drosophila Orb RNA-Binding Protein Is Required for the Formation of the Egg Chamber and Establishment of Polarity.” *Genes & Development* 8 (5): 598–613.

Leader, David P., Sue A. Krause, Aniruddha Pandit, Shireen A. Davies, and Julian A. T. Dow. 2018. “FlyAtlas 2: A New Version of the Drosophila Melanogaster Expression Atlas with RNA-Seq, miRNA-Seq and Sex-Specific Data.” *Nucleic Acids Research* 46 (D1): D809–15.

Lee, Hangnoh, Dong-Yeon Cho, Cale Whitworth, Robert Eisman, Melissa Phelps, John Roote, Thomas Kaufman, et al. 2016. “Effects of Gene Dose, Chromatin, and Network Topology on Expression in Drosophila Melanogaster.” *PLoS Genetics* 12 (9): e1006295.

Lee, Hangnoh, P. Scott Pine, Jennifer McDaniel, Marc Salit, and Brian Oliver. 2016. “External RNA Controls Consortium Beta Version Update.” *Journal of Genomics* 4 (July): 19–22.

Lee, Jeannie T. 2009. “Lessons from X-Chromosome Inactivation: Long ncRNA as Guides and Tethers to the Epigenome.” *Genes & Development* 23 (16): 1831–42.

Lerit, Dorothy A., Holly A. Jordan, John S. Poulton, Carey J. Fagerstrom, Brian J. Galletta, Mark Peifer, and Nasser M. Rusan. 2015. “Interphase Centrosome Organization by the PLP-Cnn Scaffold Is Required for Centrosome Function.” *The Journal of Cell Biology* 210 (1): 79–97.

Li, Heng. 2011. “A Statistical Framework for SNP Calling, Mutation Discovery, Association Mapping and Population Genetical Parameter Estimation from Sequencing Data.” *Bioinformatics* 27 (21): 2987–93.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. “The Sequence Alignment/Map Format and SAMtools.” *Bioinformatics* 25 (16): 2078–79.

Lin, H., and A. C. Spradling. 1995. “Fusome Asymmetry and Oocyte Determination in Drosophila.” *Developmental Genetics* 16 (1): 6–12.

Li, Xuanying, Christopher W. Seidel, Leanne T. Szerszen, Jeffrey J. Lange, Jerry L. Workman, and Susan M. Abmayr. 2017. “Enzymatic Modules of the SAGA Chromatin-Modifying Complex Play Distinct Roles in Drosophila Gene Expression and Development.” *Genes & Development* 31 (15): 1588–1600.

Lorch, Yahli, and Roger D. Kornberg. 2017. “Chromatin-Remodeling for Transcription.” *Quarterly Reviews of Biophysics* 50 (January): e5.

Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12): 550.

Mason, James M., Radmila Capkova Frydrychova, and Harald Biessmann. 2008. “Drosophila Telomeres: An Exception Providing New Insights.” *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 30 (1): 25–37.

McConnell, Kristopher H., Michael Dixon, and Brian R. Calvi. 2012. “The Histone Acetyltransferases CBP and Chameau Integrate Developmental and DNA Replication Programs in Drosophila Ovarian Follicle Cells.” *Development* 139 (20): 3880–90.

McKearin, D. M., and A. C. Spradling. 1990. “Bag-of-Marbles: A Drosophila Gene Required to Initiate Both Male and Female Gametogenesis.” *Genes & Development* 4 (12B): 2242–51.

McKearin, D., and B. Ohlstein. 1995. “A Role for the Drosophila Bag-of-Marbles Protein in the Differentiation of Cystoblasts from Germline Stem Cells.” *Development* 121 (9): 2937–47.

Mohler, Dawson, and Andrea Carroll. 1984. “Report of Dawson Mohler and Andrea Carroll.” *Drosophila Information Service* 60 (June): 236–41.

Mohler, J. D. 1977. “Developmental Genetics of the Drosophila Egg. I. Identification of 59 Sex-Linked Cistrons with Maternal Effects on Embryonic Development.” *Genetics* 85 (2): 259–72.

Ohlstein, B., and D. McKearin. 1997. “Ectopic Expression of the Drosophila Bam Protein Eliminates Oogenic Germline Stem Cells.” *Development* 124 (18): 3651–62.

728 Orr-Weaver, T. L. 1991. “Drosophila Chorion Genes: Cracking the Eggshell’s Secrets.” *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 13 (3): 97–105.

730 Parks, Annette L., Kevin R. Cook, Marcia Belvin, Nicholas A. Dompe, Robert Fawcett, Kari Huppert, Lory R. Tan, et al. 2004. “Systematic Generation of High-Resolution Deletion Coverage of the Drosophila Melanogaster Genome.” *Nature Genetics* 36 (3): 288–92.

733 Pelisson, A., L. Mejlumian, V. Robert, C. Terzian, and A. Bucheton. 2002. “Drosophila Germline Invasion by the Endogenous Retrovirus Gypsy: Involvement of the Viral Env Gene.” *Insect Biochemistry and Molecular Biology* 32 (10): 1249–56.

736 Peng, Jamy C., Anton Valouev, Na Liu, and Haifan Lin. 2016. “Piwi Maintains Germline Stem Cells and Oogenesis in Drosophila through Negative Regulation of Polycomb Group Proteins.” *Nature Genetics* 48 (3): 283–91.

739 Pine, P. Scott, Sarah A. Munro, Jerod R. Parsons, Jennifer McDaniel, Anne Bergstrom Lucas, Jean Lozach, Timothy G. Myers, Qin Su, Sarah M. Jacobs-Helber, and Marc Salit. 2016. “Evaluation of the External RNA Controls Consortium (ERCC) Reference Material Using a Modified Latin Square Design.” *BMC Biotechnology* 16 (1): 54.

743 Quinlan, Aaron R., and Ira M. Hall. 2010. “BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features.” *Bioinformatics* 26 (6): 841–42.

745 R Core Team. 2017. “A Language and Environment for Statistical Computing.” Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

747 Reuter, G., and P. Spierer. 1992. “Position Effect Variegation and Chromatin Proteins.” *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 14 (9): 605–12.

749 Robinson, Scott W., Pawel Herzyk, Julian A. T. Dow, and David P. Leader. 2013. “FlyAtlas: Database of Gene Expression in the Tissues of Drosophila Melanogaster.” *Nucleic Acids Research* 41 (Database issue): D744–50.

752 Rudolph, Thomas, Masato Yonezawa, Sandro Lein, Kathleen Heidrich, Stefan Kubicek, Christiane Schäfer, Sameer Phalke, et al. 2007. “Heterochromatin Formation in Drosophila Is Initiated through

754 Active Removal of H3K4 Methylation by the LSD1 Homolog SU(VAR)3-3." *Molecular Cell* 26 (1):
755 103–15.
756 Sambrook, Joseph, and David W. Russell. 2006. "Purification of Nucleic Acids by Extraction with
757 Phenol:chloroform." *CSH Protocols* 2006 (1). https://doi.org/10.1101/pdb.prot4455.
758 Shannon, Paul, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage, Nada
759 Amin, Benno Schwikowski, and Trey Ideker. 2003. "Cytoscape: A Software Environment for
760 Integrated Models of Biomolecular Interaction Networks." *Genome Research* 13 (11): 2498–2504.
761 Smit, Afa, R. Hubley, and P. Green. 2013-2015. "RepeatMasker Open-4.0."
762 http://www.repeatmasker.org.
763 Soshnev, Alexey A., Ryan M. Baxley, J. Robert Manak, Kai Tan, and Pamela K. Geyer. 2013. "The
764 Insulator Protein Suppressor of Hairy-Wing Is an Essential Transcriptional Repressor in the
765 Drosophila Ovary." *Development* 140 (17): 3613–23.
766 Stork, Tobias, Amy Sheehan, Ozge E. Tasdemir-Yilmaz, and Marc R. Freeman. 2014. "Neuron-Glia
767 Interactions through the Heartless FGF Receptor Signaling Pathway Mediate Morphogenesis of
768 Drosophila Astrocytes." *Neuron* 83 (2): 388–403.
769 Teixeira, Felipe Karam, Martyna Okuniewska, Colin D. Malone, Rémi-Xavier Coux, Donald C. Rio, and
770 Ruth Lehmann. 2017. "piRNA-Mediated Regulation of Transposon Alternative Splicing in the Soma
771 and Germ Line." *Nature* 552 (7684): 268–72.
772 Venken, Koen J. T., Joseph W. Carlson, Karen L. Schulze, Hongling Pan, Yuchun He, Rebecca
773 Spokony, Kenneth H. Wan, et al. 2009. "Versatile P[acman] BAC Libraries for Transgenesis
774 Studies in Drosophila Melanogaster." *Nature Methods* 6 (6): 431–34.
775 Venken, Koen J. T., Karen L. Schulze, Nele A. Haelterman, Hongling Pan, Yuchun He, Martha Evans-
776 Holm, Joseph W. Carlson, et al. 2011. "MiMIC: A Highly Versatile Transposon Insertion Resource
777 for Engineering Drosophila Melanogaster Genes." *Nature Methods* 8 (9): 737–43.
778 Vermaak, Danielle, and Harmit S. Malik. 2009. "Multiple Roles for Heterochromatin Protein 1 Genes in
779 Drosophila." *Annual Review of Genetics* 43: 467–92.

780 Wang, Sidney H., and Sarah C. R. Elgin. 2011. “Drosophila Piwi Functions Downstream of piRNA
781 Production Mediating a Chromatin-Based Transposon Silencing Mechanism in Female Germ Line.”
782 *Proceedings of the National Academy of Sciences of the United States of America* 108 (52):
783 21164–69.
784 Wang, Xiaoxi, Lei Pan, Su Wang, Jian Zhou, William McDowell, Jungeun Park, Jeff Haug, Karen
785 Staehling, Hong Tang, and Ting Xie. 2011. “Histone H3K9 Trimethylase Eggless Controls
786 Germline Stem Cell Maintenance and Differentiation.” *PLoS Genetics* 7 (12): e1002426.
787 Warming, Søren, Nina Costantino, Donald L. Court, Nancy A. Jenkins, and Neal G. Copeland. 2005.
788 “Simple and Highly Efficient BAC Recombineering Using galK Selection.” *Nucleic Acids Research*
789 33 (4): e36.
790 Wayne, S., K. Liggett, J. Pettus, and R. N. Nagoshi. 1995. “Genetic Characterization of Small Ovaries,
791 a Gene Required in the Soma for the Development of the Drosophila Ovary and the Female
792 Germline.” *Genetics* 139 (3): 1309–20.
793 Weiler, K. S., and B. T. Wakimoto. 1995. “Heterochromatin and Gene Expression in Drosophila.”
794 *Annual Review of Genetics* 29: 577–605.
795 Yang, Peng, Yixuan Wang, and Todd S. Macfarlan. 2017. “The Role of KRAB-ZFPs in Transposable
796 Element Repression and Mammalian Evolution.” *Trends in Genetics: TIG* 33 (11): 871–81.
797 Yasuhara, Jiro C., and Barbara T. Wakimoto. 2006. “Oxymoron No More: The Expanding World of
798 Heterochromatic Genes.” *Trends in Genetics: TIG* 22 (6): 330–38.
799 Yin, Hang, and Haifan Lin. 2007. “An Epigenetic Activation Role of Piwi and a Piwi-Associated piRNA in
800 Drosophila Melanogaster.” *Nature* 450 (7167): 304–8.
801 Yoshioka, K., H. Honma, M. Zushi, S. Kondo, S. Togashi, T. Miyake, and T. Shiba. 1990. “Virus-like
802 Particle Formation of Drosophila Copia through Autocatalytic Processing.” *The EMBO Journal* 9
803 (2): 535–41.
804 Yuan, Kai, and Patrick H. O’Farrell. 2016. “TALE-Light Imaging Reveals Maternally Guided,
805 H3K9me2/3-Independent Emergence of Functional Heterochromatin in Drosophila Embryos.”

*Genes & Development* 30 (5): 579–93.

Zook, Justin M., Daniel Samarov, Jennifer McDaniel, Shurjo K. Sen, and Marc Salit. 2012. “Synthetic Spike-in Standards Improve Run-Specific Systematic Error Analysis for DNA and RNA Sequencing.” *PloS One* 7 (7): e41356.

Zhimulev, Igor F., and Elena S. Belyaeva. 2003. “Intercalary Heterochromatin and Genetic Silencing.” *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 25 (11): 1040–51.