Supplemental Figures and Tables for:

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

Benner_FigS1

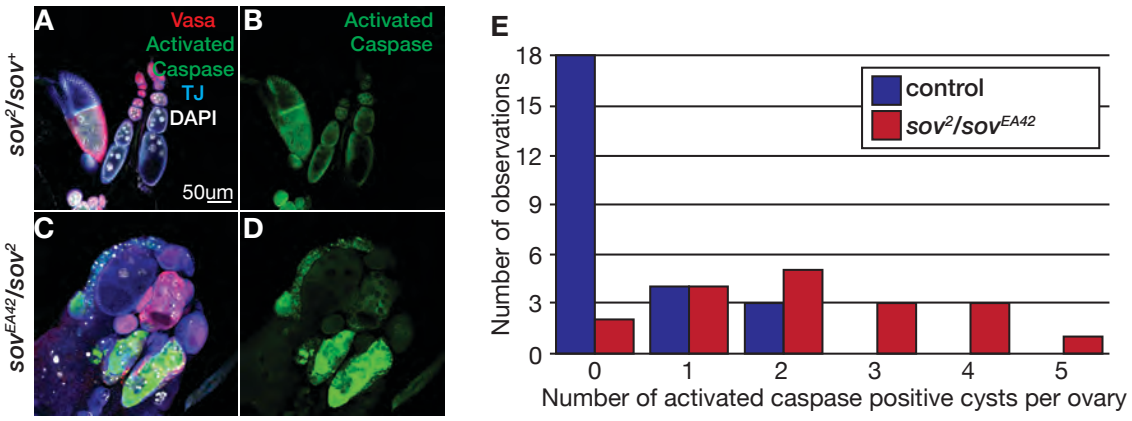

Figure S1. Loss of *sov* leads to apoptosis of cysts in the ovary. (A–D) Ovaries (5 days posteclosion) of the indicated genotypes were stained for anti-Vas (red), -Tj (blue), -cleaved Caspase-3 (green), and DAPI (white). *sov* mutant ovaries had more cysts with activated Caspase-3 than did controls, indicating that loss of *sov* leads to apoptosis in the ovary. (E) Quantification of activated Caspase-3 cysts in *sov* mutants and controls.

Benner_FigS2

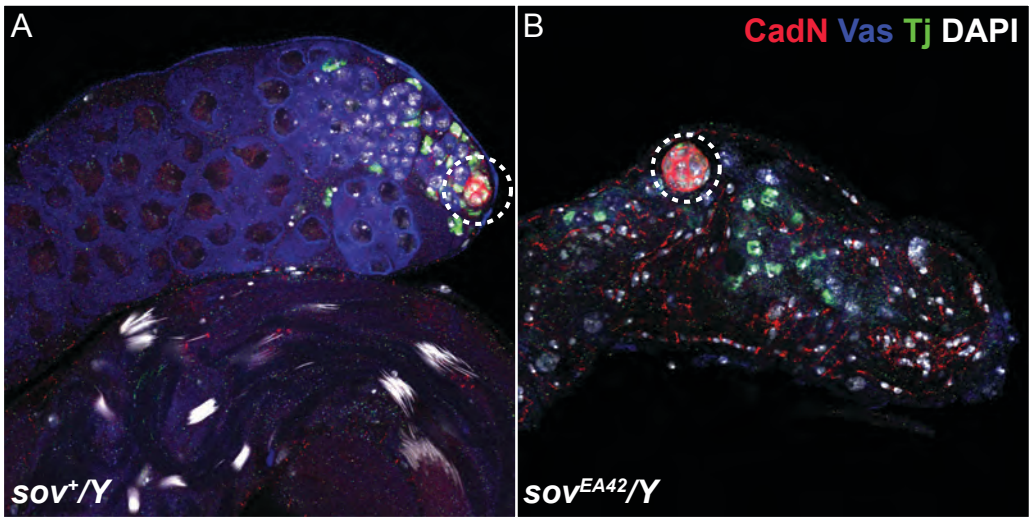

Figure S2. Loss of germ cells in males escaping lethality. (A, B) Testes (5 days posteclosion) of the indicated genotypes were stained for anti-Vas (blue), -Tj (green), -CadN (red), and DAPI (white). (A) Fertile *sov*$^{+}$*/Y* testis contain developing spermatogonia (blue) posterior to hub cells (red, dashed circle). (B) Escaper, sterile, *sov*^{EA42}*/Y* males that survive to adulthood contain rudimentary testes that are agametic and therefore lack anti-Vas staining. The hub is usually present, but displaced from the anterior end of the testes (red, dashed circle).

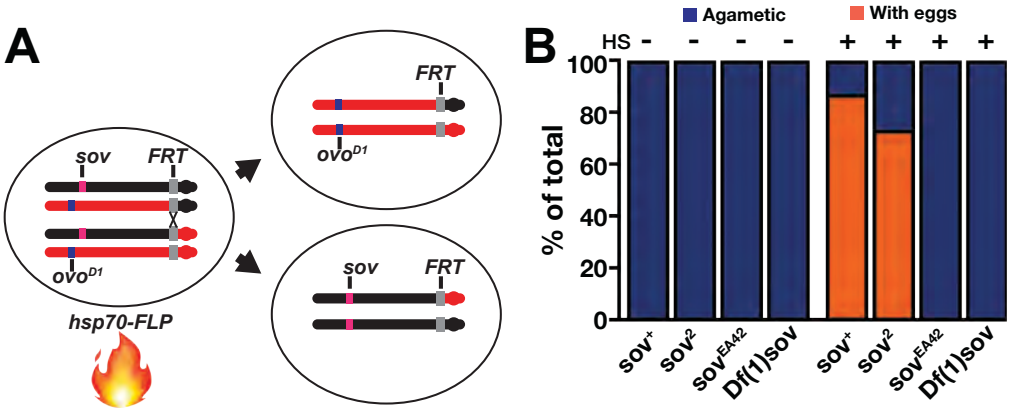

Figure S3. *sov* germline clonal analysis. (A) Cartoon of dominant female sterile germline clonal analysis technique. Recombination occurs between homologs at FRT sequences only in the presence of heat-shock induced FLP expression. (B) Quantification of germline production from *sov* germline mutant clones. Heat shock (HS) was used (+) to induce expression of FLP, but was omitted (–, no HS) in controls.

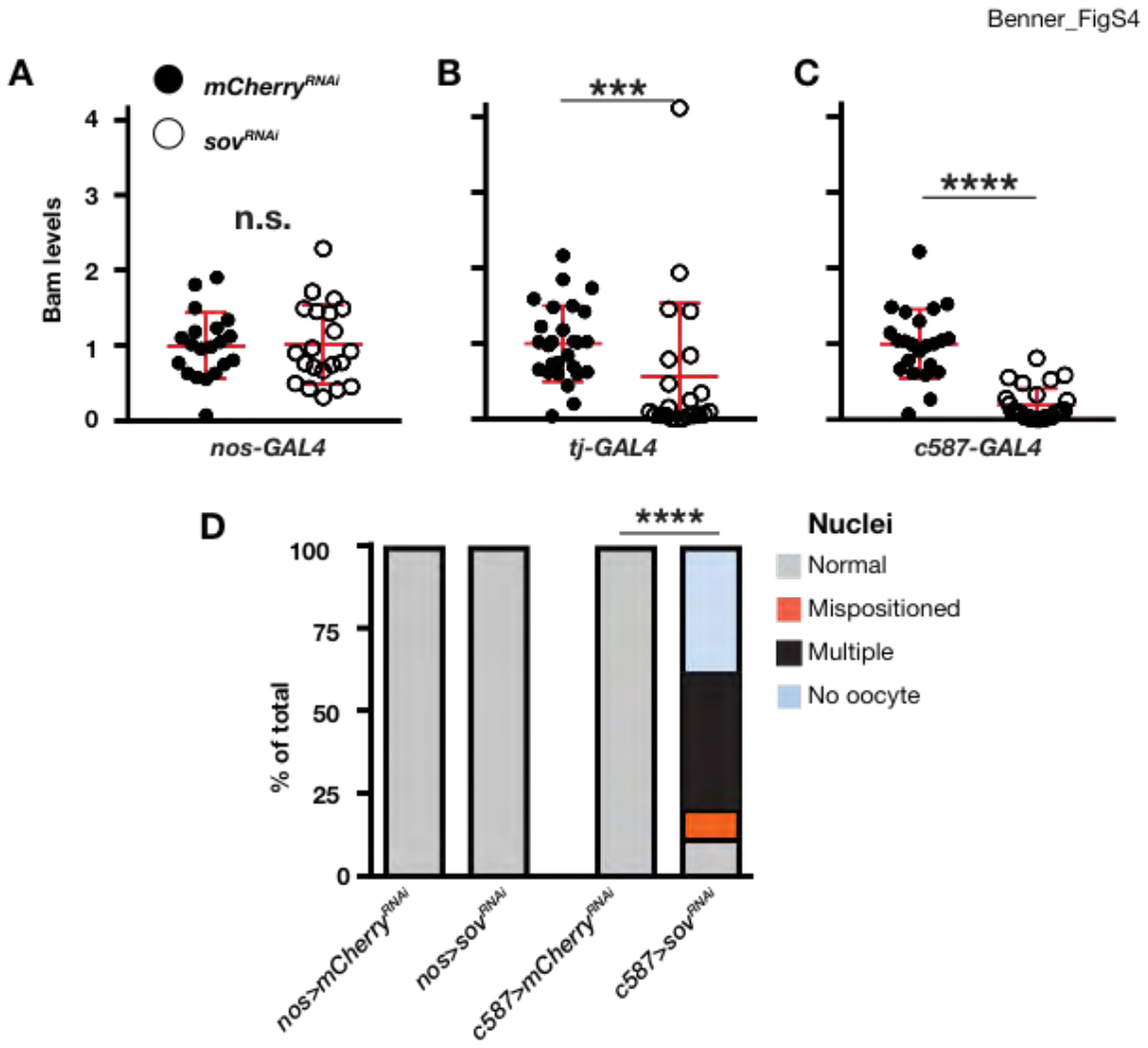

Figure S4. Quantification of Bam staining intensity and Orb positive nuclei. (A) Knockdown of *sov* in germline cells does not alter Bam intensity levels. (B, C) In contrast, knockdown of *sov* in somatic cells leads to significant reduction in Bam intensity levels within region 1. ***, $p<0.001$; ****, $p<0.0001$, n.s., not significant by two-tailed t-test or Mann-Whitney test. Each dot represents a measurement from an individual germarium. Data are mean +/- SD. Bam signals were measured from individual ovarioles as integrated densities from maximum-intensity projections from the region showing peak Bam signal intensity defined by the ROI tool (ImageJ) minus the integrated density measured from a neighboring ROI to subtract background fluorescence. Bam intensity was measured only from ovarioles containing distinguishable germaria and/or egg chambers. (D) Quantification of Orb positive nuclei in the indicated genotypes. Loss of *sov* in somatic cells leads to defects in Orb localization. ****, $p<0.0001$, Fisher's exact test. Shown are representative data from one experiment following 2–3 biological replicates.

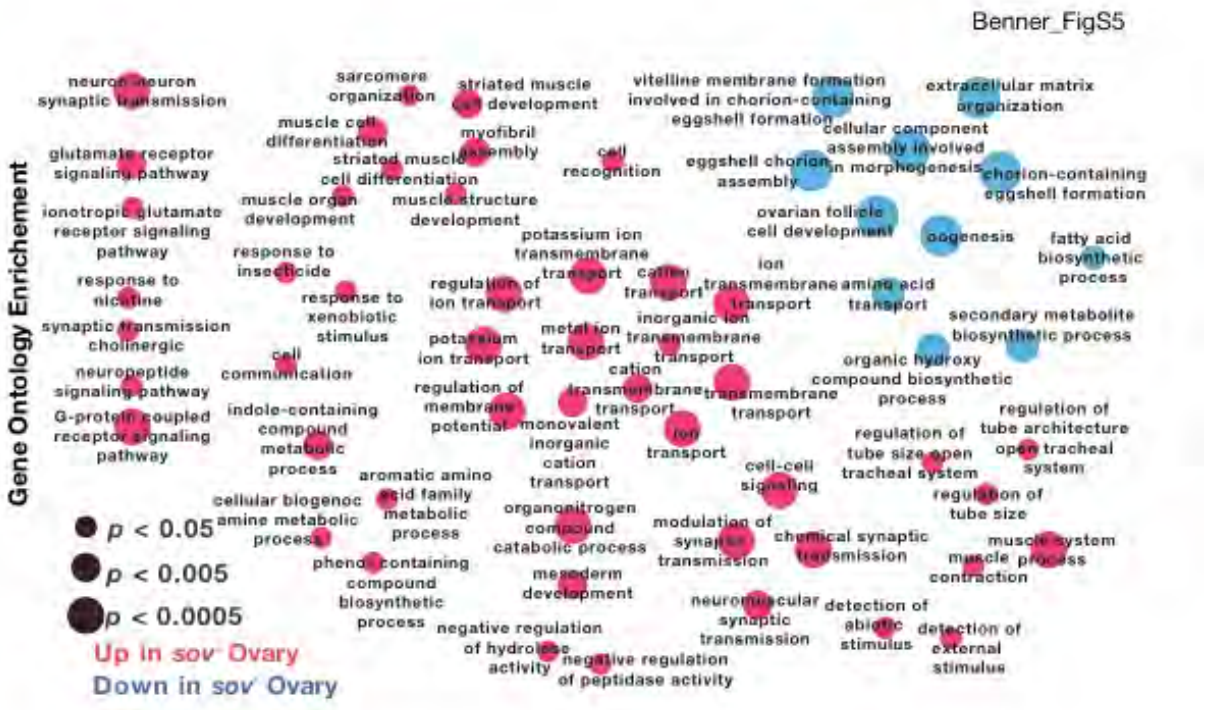

Figure S5. Gene ontology terms of differentially expressed genes in *sov* mutant ovaries. One sided GO enrichment analysis of gene lists that are differentially expressed in *sov* mutant ovaries. Red and blue circles show GO terms associated with derepressed and repressed genes respectively. Circle size denotes Bonferroni Step Down corrected *p*-values.

Table S1. DNA sequencing results. Modified output from snpEFF analysis. 'X Location' is the base that is modified from the reference. 'Reference' is the reference nucleotide at that position. 'Polymorphism' is the nucleotide change from the reference. 'Gene Effected' is the annotated gene that contains the polymorphism. 'Type' is the attribute given to the polymorphism change. 'Effect' denotes the predicted severity of the polymorphism. 'Codon Change' specifies if the polymorphism results in an amino acid change. The 'Found in Genotypes' states which genotypes contained the polymorphism. 'Allele specificity' is if the polymorphism can be specified to a single *sov* allele. Green rows indicate the ncRNA polymorphisms and yellow denote the predicted *sov* allele specific mutations.

Table S2. Sov RNA-seq sample correlations. The Pearsons correlation coefficient of the mean RNA-seq RPKM values for each genotype (row x column).

Table S3. Differentially expressed genes in *sov* mutants *vs.* controls. (A) Total number of genes differentially expressed in sov mutants for each genotype comparison. 'Genotype' is the sov mutant/RNAi line that is compared to the respective control. '> 2 FC ALL' are the number of genes and transposons with a greater than 2 $\log_2$ fold change and *padj* <0.05 according to DESeq2 analysis. '< -2 FC ALL' are the number of

genes and transposons with a less than -2 $\log_2$ fold change and *padj* value <0.05 according to DESeq2 analysis. The same description applies to the rest of the columns but represent 'Genes' and 'Transposons' only. (B) DESeq2 gene expression analysis output for all *sov* mutants vs. all controls. 'FBgn' corresponds to the FlyBase gene identifier. 'BaseMean' is the mean normalized read count for *sov* and control samples. 'log2FoldChange' is the $\log_2$ normalized read count fold change in *sov* mutants *vs.* controls. 'padj' is the fdr adjusted *p-values* for differential expression. 'gene' refers to the common gene name.

Table S4. GO term enrichment analysis. Significantly enriched GO terms for gene sets differentially expressed in *sov* mutants with ClueGO analysis (Bindea et al. 2009). (A) Significantly enriched GO terms for *sov* mutant derepressed gene set. (B) Significantly enriched GO terms for *sov* mutant repressed gene set. 'GOID' is the identification number of the associated GO term. 'GOTerm' is the biological description of the GO term. 'Ontology Source' is the reference set the GO term is derived from. 'GOLevels' is the hierarchical GO levels the terms are grouped from. '% Associated Genes' is the number of genes in the enrichment set over all the genes in the respective GO term. 'Nr. Genes' is the total number of genes in the enrichment set associated with the respective GO term. 'Associated Genes Found' are all the gene in the enrichment set that are part of that respective GO term.

Table S5. FlyBase ART file. Table containing all experimental materials used in the study.

Table S6. Tissue Expression of derepressed genes in *sov* mutants. $\log_2$ fold change values from DESeq2 analysis of the derepressed genes in *sov* mutants that were expressed in all tissues. Each tissue comparison is to the sex-respective whole organism tissue expression. 'FBgn' corresponds to the FlyBase gene identifier. 'gene' refers to the common gene name. The following columns are the sex and tissue $\log_2$ fold change values.